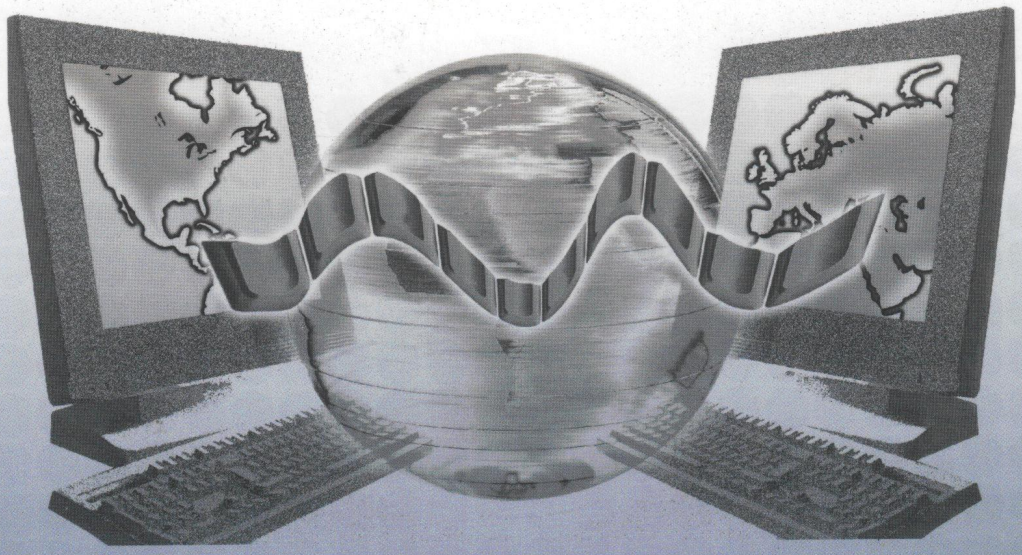
2006 International Symposium
on Distributed Computing and Applications
to Business, Engineering and Science

DCABES 2006 PROCEEDINGS

Volume II

Editor in Chief: Xu Wenbo
Associate Editor: Wang Guangming



SHANGHAI UNIVERSITY PRESS

2006 International Symposium
on Distributed Computing and Applications
to Business, Engineering and Science

DCABES 2006 PROCEEDINGS

Volume II

Editor in Chief: Xu Wenbo
Associate Editor: Wang Guangming

SHANGHAI UNIVERSITY PRESS

图书在版编目（CIP）数据

2006年国际电子、工程及科学领域的分布式计算应用学术研讨会论文集 / 须文波主编. —上海：上海大学出版社，2006.10
ISBN 7-81118-023-5

Ⅰ.2… Ⅱ.须… Ⅲ.电子计算机—国际学术会议—文集—英文 Ⅳ.TP3-53

中国版本图书馆CIP数据核字(2006)第120341号

责任编辑 姚铁军
封面设计 孙 敏

DCABES2006 PROCEEDINGS

2006年国际电子、工程及科学领域的分布式计算和应用学术研讨会论文集

须文波 主编

王光明 副主编

会议承办单位：江南大学
会议协办单位：浙江工商大学



上海大学出版社出版发行
（上海市上大路99号 邮政编码 200444）
（http://www.shangdapress.com 发行热线 66135110）

出版人：姚铁军

*

宜兴市德胜印刷有限公司印刷
开本 1/16 印张 89.5 字数 1 878千
2006年10月第一版 2006年10月第一次印刷
印数：001-600册
ISBN 7-81118-023-5/TP.037

定价：300.00元

DCABES 2006 PROCEEDINGS

Volume II

责任编辑 姚铁军
封面设计 孙 敏

ISBN 7-81118-023-5

ISBN 7-81118-023-5/TP · 037

Price: 300 RMB

Volume II

8. Natural Network Computing

9. E-commerce and business applications

10. Computer Control and Applications

11. Network and Web-based Computing

12. System Architecture and applications

Preface

The series of meetings, International Symposium on Distributed Computing and Applications to Business, Engineering and Science (DCABES), is now becoming an important international event on various applications and the related computing environments of distributed and grid computing. The first meeting was held at Wuhan University of Technology, Wuhan, and the second meeting was held at Southern Yangtze University, Wuxi, the third meeting was held at Wuhan University of Technology, Wuhan, the fourth meeting was held at Greenwich University, Greenwich. In this year, the fifth meeting will be organized by Southern Yangtze University and Zhejiang GongShang University and hold at Hangzhou. The conference theme include not only its traditional theme such as parallel and distributed computing, but also intelligent and high performance computing that will be described as follows.

It was my pleasure that the DCABES2006 conference had received a great number of papers submitted cover a wide range of topics, such as Parallel/Distributed Algorithms, Distributed System and Distributed Computing, Grid Computing and Parallel Processing, Network and Applications, Database and Engineering Applications.

Papers submitting to the conference come from over 16 countries and regions. All papers contained in this Proceeding are peer-reviewed and carefully chosen by members of Scientific Committee, proceeding editorial board and external reviewers. Papers accepted or rejected are based on majority opinions of the referee's. All papers contained in this proceeding give us a glimpse of what future technology and applications are being researched in the distributed computing area in the world.

I would like to thank all members of the Scientific Committee, the local organizer committee, the proceedings editorial board and external reviewers for selecting the papers. Special thanks are due to Dr. Choi-Hong LAI, Prof. Qingping Guo and Prof. GuangMing Wang, who co-chaired the Scientific Committee with me. It is indeed a pleasure to work with them and obtain their suggestions. I am also grateful to Professor H. Power, Professor Peter M.A. Sloot, Professor Aoying Zhou, Professor Yao Zhen, for their contributions of keynote speeches in the conference.

Sincerely thanks should be forwarded to the China Ministry of Science and Technology (MOST), the China Ministry of Education (MOE), the Natural Science Foundation of China (NSFC) and Southern Yangtze University and Zhejiang GongShang University.

Finally I should also thank Professor Yun Ling, WeiMing Wang and Mrs Li Liu, for their efforts in conference organizing activities, my postgraduate students, such as Mr. Jun Xu, Mr. Peng Wang, and Mr. ZhaoKuo Nan, for their time and help. Without their time and efforts this conference cannot be organized smoothly.

Enjoy your stay in Hangzhou. Hope to meet you again at the DCABES 2007.

Professor Wenbo Xu,
Chair of the DCABES2006
School of Information Technology
Southern Yangtze University
Jiangsu, China

Chair of Scientific Committee

Xu, Prof. W. B., Southern Yangtze University, Wuxi, China

Co-Chair of Scientific Committee

Guo, Prof. Q. P., Wuhan University of Technology, Wuhan, China
Lai, Prof. C. -H., University of Greenwich, UK
Wang, Professor G. -M., Zhejiang Gongshang University, China

Chair of Organizeing Committtee

Xu, Prof. W. B., Southern Yangtze University, Wuxi, China

Scientific Committee (in alphabetical order)

Cai,Prof. X. C., University of Colorado, Boulder, U.S.A

Cai,Professor Jiamei, Zhejiang Industry University, Hangzhou, China

Cao,Prof. J.W., Research and Development Centre for Parallel Algorithms and Software, Beijing, China

Chi, Prof. X.B., Academia Sinica, Beijing, China

Guo, Professor Q.P., Wuhan University of Technology, Wuhan, China

Ho, Dr. P. T., University of Hong Kong, Hong Kong, China

Jesshope, Professor C., University of Amsterdam, the Netherlands

Kang, Professor L.S., Wuhan University, China

Keyes, Professor D.E., Columbia University, USA

Lai, Prof. C. H., University of Greenwich, UK

Lee, Dr. John, Hong Kong Polytechnic, Hong Kong, China

Liddell, Professor H. M., Queen Mary, University of London, UK

Lin, Dr. H.X., Delft University of Technology, Delft, the Netherlands

Lin, Dr. P., National University of Singapore, Singapore

Loo, Dr. Alfred, Hong Kong Lingnan University, Hong Kong, China

Sloot, Professor P.M.A., University of Amsterdam, Amsterdam the Netherlands

Sun, Professor J., Academia Sinica, Beijing, China

Tsui, Mr. Thomas, Chinese University of Hong Kong, Hong Kong, China

Wang, Professor G. M., Zhejiang Gongshang University, China

Wang, Professor Shitong, S.T Southern Yangtze University, Wuxi, China

Wang, Professor WeiMing, Zhejiang Gongshang University, Hangzhou, China

Xu, Professor W.B., Southern Yangtze University, Wuxi, China

Zhang, Professor J., University of Kentucky, USA

Zhuang Professor, Yueting, Zhejiang University,Hangzhou, China

Local Organizing Committee (in alphabetical order)

Chen, Professor Jian, Southern Yangtze University, Wuxi, China

Feng, Professor Bin, Southern Yangtze University, Wuxi, China

Ling, Professor Yun, Zhejiang Gongshang University, Hangzhou, China

Liu, Professor Yuan, Southern Yangtze University, Wuxi, China

Sun, Professor Shouqian, Zhejiang University, Hangzhou, China

Editorial Board

Underwater Bottom Mine Shells Target Classification Based on Relevance Vector Machine*

Qiang Wang
College of Metrological Technology & Engineering, China Jiliang University
Hangzhou, Zhejiang 310018, China
Email: qiangwang@cjlu.edu.cn

And

Heyun Huang
Department of Information Science & Electronic Engineering, Zhejiang University
Hangzhou, Zhejiang 310027, China
Email: wangtulip@yahoo.com.cn

ABSTRACT

The problem of classifying underwater bottom mines from acoustic backscattered signals is addressed here. Standard short-time fourier transform (STFT) is applied to convert the echo signal into the time-frequency plane to precisely depict the echo spectrogram, then time-frequency feature extraction scheme based on modified STFT is introduced to deal with mine shell echo signals corrupted by impulse, non-Gaussian noise. The scheme provides a robust estimation of STFT in the reverberation and suppresses it. The target echo features are extracted to reflect different target strengths of two mine shell types influenced by reverberation. The overall system classification performance is benchmarked on two mine shell echo data sets with 25 kHz-50 kHz bandwidth. Echo features are sent to relevance vector machine (RVM) classifier which represents a Bayesian extension of support vector machine (SVM). Compared with SVM, the case study shows RVM yields a much sparse solution and improves classification accuracy. The lake experiment exploits the robustness of feature extraction scheme and effectiveness of classifier with the analysis of the echoes from the two shells underwater bottom.

Keywords: Underwater Mine Classification, STFT, Feature Extraction, Relevance Vector Machine.

1. INTRODUCTION

The problem of detecting and classifying underwater bottom mines remains a challenge. The objective is to discover what information contained in the backscattered signal that might be exploited for the sea bottom target classification. Mine classification in shallow water is dominated by the needs to discriminate mines in reverberation and to separate them from competing clutter [1, 2].Active sonar transmits acoustic wave to discriminate underwater targets of diverse sizes, shapes and scattering properties. The acoustic backscattered signals are mainly applied in the active sonar system to perform the classification task. This task involves stable extraction of characteristics of the targets from backscattered signals in the underwater environment and the successive pattern recognition from relative small samples. The theoretical treatments of resonant and scattering for under- water objects have been found in [3, 4, 5]. When an object is excited under continuous force excitation, it will undergo a forced vibration which is the sum of many excited modes of vibration of objects. Though the returning echoes contain abundant characteristic information of targets [1], target echoes are corrupted by many other irrelevant objects returning signals, i.e., reverberation. The solution of echo signal processing is complex for some reasons: 1) difficulty lies in how to effectively characterize the target and accurately estimate its parameters. 2) active sonar often receives echoes with strong reverberation, which brings a large disturbance to the echoes and leads to imprecision of describing targets' properties. Also, the influences of the channel on the target's signature can not be ignored. 3) collected samples of echoes always have only a small set which results in poor generalization of classifier.

Detecting and classifying underwater target objects is challenging by using acoustic backscattered signals. Feature selection scheme is a key point in this problem. The pioneer paper used simple spectral features as input to the neural network classifier in order to distinguish a cylindrical target from rock with similar shape. The G-Transform is developed to represent the resonance or the modulation on the frequency spectrum of acoustic backscattered signal [2]. Several methods were tested for finding a good data representation scheme via optimal decomposition with wavelet basis functions [3]. Wavelet and neural networks classification scheme is developed to discriminate mine like and non-mine like objects from the acoustic backscattered signals, and using adaptive feature mapping in changing environments[4]. The filter bank model which consists of several band pass filters each to tune to select certain frequency information, is proposed to process and encode information in echoes [12]. Robinson and Azimi-Sadjadi demonstrated that the classification accuracy can be achieved largely when the backscattered signals' feature can be fused by some linear or non-linear fusion schemes [13]. Transient sonar signal classification using hidden Markov models and neural network is in [14]. Recently, time reversal imaging (TRI) is applied to classifying underwater target by relevant vector machine [5].

It is not possible to extract statistics-based features from the time-series as the number of samples is too small. To classify underwater targets, we need to get a trade-off between classification efficiency and computation complexity. Aim at effectively classifying underwater mine shells, backscattered signals feature extraction scheme based on short-time fourier transform(STFT) and relevance vector machine classifier are revealed in this paper.

2. ECHOES FEATURE EXTRACTION

* This work is support by China Jiliang University. (XZ0513)

2.1 Echoes Signal Model

The echo signal model of the targets will be investigated first in this section. As the transmitted signal is linear frequency modulated (LFM), it is suitable to represent this signal $s_t(t)$ as Gaussian enveloped[1]. The point scatterer echo return is modeled as Eq. (1).

$$s_r(t) = \sqrt{\beta E_s} s_t(\beta(t-\tau)) \quad (1)$$

Where $s_t(t)$ means transmitted signal, underwater single target received back scattered signal is $s_r(t)$, β is dilation parameter, τ is delay, E_s is echo energy. The parameters β,τ,E_s reflect different target's properties.For simplicity, Eq.(1) can be denoted as:

$$s_r(t) = \sqrt{\beta E_s} s_t(\beta t) \quad (2)$$

Based on echo energy cycle [1], it is clear that the energy of the received echo is influenced by transmitter, projector, spreading channel, target strength and receive array. The target strength effect on the backscattered signal is considered as:

$$E_s = T_s(\omega) \quad (3)$$

An introduction of STFT is given in [8]. Implementation of STFT on the echo $s_r(t)$ can acquire the Eq.(4).

$$Sp_r(t,\omega) = \frac{1}{\sqrt{2\pi}} \int s_r(u) h(u-t) e^{-j\omega u} du \quad (4)$$

To different underwater targets, parameters β,τ,E_s contain important information of returning echo. The β,τ depend on moving condition of target and transmitter. In case study,the mine shells is static on the lake bottom, so transmitter moving information can not reflect target features. Therefore, just considering echo signal amplitude, the Eq.(4) is simplified as:

$$Sp_r(t,\omega) = \frac{1}{\sqrt{2\pi}} \int \sqrt{E_s} s_t(u) h(u-t) e^{-j\omega u} du = E_s Sp_t(t,\omega) \quad (5)$$

The backscattered signal is influenced by target strength T_s, so E_s is function of frequency ω, $E_s=T_s(\omega)$.

$$Sp_r(t,\omega) = T_s(\omega)\, Sp_t(t,\omega) \quad (6)$$

To integrate with t, the target strength T_s is measured as:

$$F(\omega) = \int Sp_r(t,\omega) dt = \int T_s(\omega) Sp_t(t,\omega) dt \quad (7)$$

Where $F(\omega)$ means echo signal features, which includes $Sp_t(t,\omega)$ (transmitted signal spectrogram) and T_s. So varying T_s value means different underwater target features.

2.2 Robust Echo Feature Extraction

In underwater varying environment, target echo features will not obvious because of reverberation disturbance, while STFT can map time domain signal into the time-frequency domain and localize the time and frequency simultaneously. It is used to determine the resonance spectrum of submerged elastic cylindrical wires in water [6]. STFT divides the frequency axis into equal-bandwidth components obviously yields a linear division of the spectrum. STFT owns quick computation speed because it uses fast Fourier transform to produce spectrogram. The echo signal's energy is distributed in the joint time-frequency domain. The LFM echo features extraction process is described in Fig. 1..

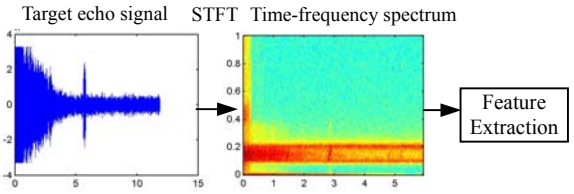


Fig. 1. Time-frequency features extraction process

The received echo owns important frequency features information reflecting the target strength, the key point of underwater mine shell discrimination is robust echo features extraction [4,5]. Though time-frequency analyses, such as Wigner-Ville distribution, STFT are powerful tool for non-stationary signal, it produces poor results in an impulse noise environment [16]. Underwater mine echoes often corrupted by non-Gaussian tailed noise, therefore, standard reduced interference distributions will not adequately reduce this kind noise. In addition, some signal-dependent distributions will not be efficient here since they recognize these strong noise components as parts of the signal [15].

Modified STFT scheme is proposed. Given a frequency ω, the total energy of this frequency band, 'frequency margin', measures the target strength T_s. The Fig.2 shows any frequency only has a finite duration. To integrate whole time t to calculate frequency feature is unreasonable as the noise adding into features. It obvious contains no necessary reverberation component. The following method avoids this drawback, just to integrate the region between two black triangle in Fig. 2..

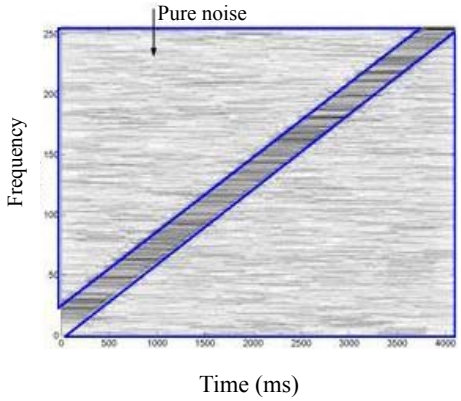


Fig. 2. Spectrum of target signal corrupted by noise (the area in the black triangle just contain pure noise

Fig. 3 shows that any frequency ω only has a finite duration t. In active sonar detection problem, the existence of uncertain noise is inevitable, so there is only noise except in the signal duration. The following feature avoids this disadvantage [17].

$$F(\omega) = \int_T T_s(\omega) Sp_t(t,\omega) dt \ , \ T = \{t, Sp_t(t,\omega) > 0\} \quad (8)$$

At a frequency ω, different target strength T_s results in different features. The discrete model is applied to calculate features.

$$F(\omega) = \sum_{i \in T} Sp_r(i,\omega) \ , \ T = \{i, Sp_t(i,\omega) > 0\} \quad (9)$$

3. RVM CLASSIFIER

3.1 Support Vector Machine

The various classification methods have been proposed to classify underwater targets, such as *K*-nearest neighbor, multivariate Gaussian, probabilistic neural networks, SVM [4]. Due to the complexity and uncertainty in underwater environment, many transmitted signals can not return to the receivers, so the echo signals collected in practical conditions is limited. With the goal of eliminating variables that are redundant, so underwater target classifier generalization is more important in data feature level.

SVM has emerged as one powerful tool for face recognition, signal processing, fault diagnosis by Vapnik[7]. Unlike most classification methods, SVM can be used when only a small amount of training sample is available. It combines excellent generalization properties with a sparse kernel representation. It makes decision based on a function of the form.

$$y(x)=\sum_{n=1}^{N} wK(x,x_n)+w_0 \tag{10}$$

Where $\{w_n\}$ are the weights and $K(x,x_n)$ is kernel function. However, it does suffer from a number of disadvantages. The standard SVM solution is more complicated and depends on the conditional distribution function and various conditional expectations. Note that in most cases, it is neither possible nor necessary to classify all the possible future samples. In fact, only a particular set of the testing data is of real interest [9, 10].

Although relatively sparse, only using SVs(support vector),SVM make unnecessarily liberal use of basis function since the number of SVs required typically grows linearly with the size of training set[11].

3.2 Relevance Vector Machine

Relevance vector machine(RVM) is proposed by M.Tipping in 2001, which represents a Bayesian extension of SVM. It inherits SVM merits and overcome its drawbacks [9].Till now, the application of RVM in underwater target classification just restricted in TRI [5]. To underwater target recognition, RVM classifier owns obvious advantages. 1) it typically utilizes dramatically fewer kernel functions guarantee hyperplane robustness, and helps to reduce classification system computation complexity. 2) RVM is followed standard probabilistic formulation and assumes that the targets are samples from the model with some additive noise.

The RVM classification is desired to predict the posterior probability of class membership given the input. Similar to SVM, we generalize the linear model by applying the logistic sigmoid function.

$$\sigma(y)=1/(1+e^{-y}) \tag{11}$$

where y is inner product of weight vector w and input samples x. We assume kernel is corrupted by Gaussian noise.

$$t_n=y_n+e_n=y(x_n;w)+e_n \tag{12}$$

Due to the assumption of independence of the t_n, the likelihood of the complete data can be written as:

$$p(t|w,\sigma^2)=(2\pi\sigma^2)^{-N/2}\exp\{-\|t-\Psi w\|^2/2\sigma^2\} \tag{13}$$

Where t is vector, our aim is to get $p(w|t)$ expression, and then according to Bayes criterion to select w. assuming sigmoid function follows Bernoulli distribution. Therefore, to deterministic weight vector w, its distribution function is：

$$p(t|w)=\prod_{n=1}^{N}\sigma\{y(x_n;w)\}^{t_n}[1-\sigma\{y(x_n;w)\}]^{1-t_n} \tag{14}$$

Assumption vector weight w follows multi-dimension normal distribution [10], to w_i (i =1…N) then

$$p(w|v)=\prod_{i=0}^{N} N(w_i|0,v_i^{-1}) \tag{15}$$

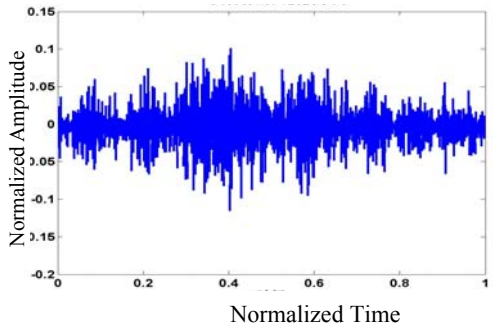
so $p(w|t,v)$ is determined by $p(t|w)p(w|v)$, based on Eq.(14)and Eq.(15), we get a estimation.

$$\log[p(t|w)p(w|v)]=\sum_{n=1}^{N}[t_n\log y_n+(1-t_n)\log(1-y_n)]-\frac{1}{2}w^T Aw \tag{16}$$

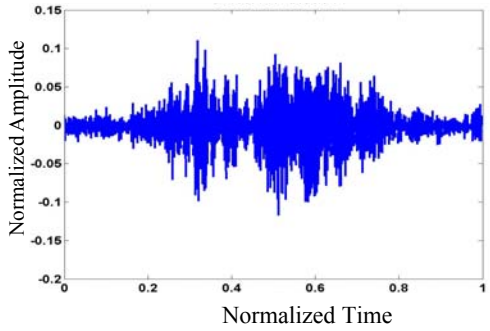
Then weight w is determined by above Eq. (16). The total algorithm is presented in [10].

4. LAKE EXPERIMENT RESULTS

The sonar data set used in this study is a subset of acoustic backscattered data set collected at a lake test,and data sets are 608 samples[17]. The two kinds mine shells are placed on the bottom of lake, about 20meters below the lake surface. The broadband LFM signals are transmitted from the bank of this lake. LFM duration is 4ms and bandwidth in the range of 25 kHz−50kHz.The typical echoes of two kinds mine shells are shown as Fig. 3. (a) and Fig. 3. (b). After the targets backscatter the transmitted signal, the acquired data from sensors is revealed in the Fig. 3., which includes two waveforms of different mine shell[17].The STFT spectrums of echoes are in Fig. 4 .(a) and Fig. 4. (b).

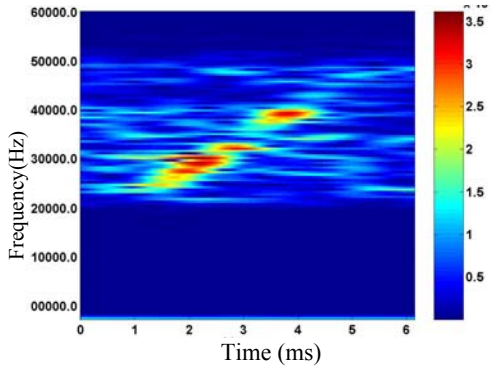


(a)Target 1

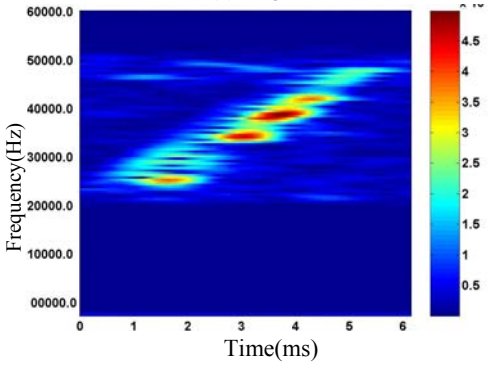


(b)Target 2

Fig. 3. Two types of mine shells echo signals



(a)Target 1



(b)Target 2

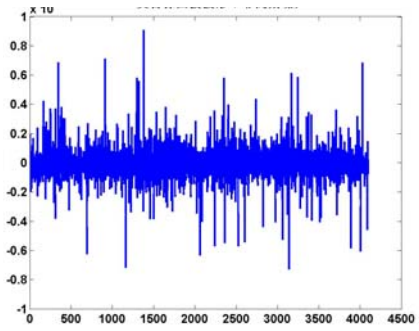
Fig. 4. STFT spectrum of two kind targets echo signals

The classification results of SVM with Gaussian noise are shown as Table 1.. Classifier owns good generalization and classification performance as the number of samples is too small. But classifying correctness rate will obviously decrease, less than 80%, when adding non-Gaussian noise.

Table 1. SVM classification results

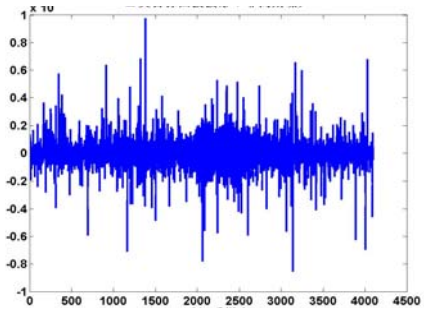
Name	total samples	correct samples	accuracy target 1	accuracy target 2	total accuracy
Training samples	32	31	93.25%	99.75%	96.50%
Testing samples	608	580	90.38%	96.38%	93.38%

To test the robustness of the features and the ability of classifier to adapt to environmental variations, certain signal noise rate (SNR) non-Gaussian noise is added to original signals. SNR is varied from 0 to 20dB. When SNR drop to -3dB, the echo signals of two mine shell is shown in Fig. 5.



Time (ms)

(a)Target 1



Time(ms)

(b)Target 2

Fig. 5. Target 1 and 2 echo signals adding non-Gaussian noise

Adding non-Gaussian noise is enough strong, features effectivity is worsening. As a result, the correctness rate less than 90%.If continuing to enhance no-gaussian noise energy, SNR drop from −3dB to −12dB, the correct rate is reduce from 90% to 85%. When SNR is high than −12dB (the SNR reference value is from [4]), the results is worse, and only 75% classification correct rate.

To the same two kind mine shells acoustic backscattered signal performing feature extraction scheme, two dimensions vectors are get and sent to RVM classifier.

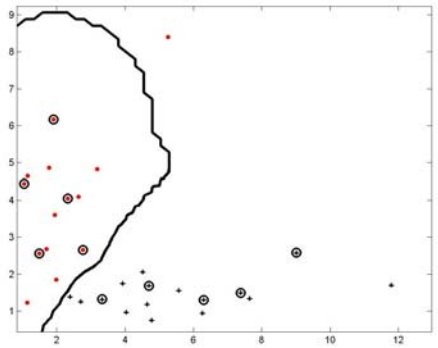


Fig. 6. Two dimension SV vector RVM Classification

From Fig. 6, symbols circle point and cross stand for two kind echo signal features vector, bold black line is classification decision hyperplane. The circle points in black circle are RVM support vector (SV). The RVM classifier advantages exists SV numbers are obviously reduced. As Table 2., the correct rate is high than 90%,when signal reverberation ratio (SNR) is -3dB.Correct classification rate is also high than 85%,when SNR is −12dB. A classification result is encouraging and proves feature extraction method and classifier owns robustness.

Table 2. RVM classification results

Name	total samples	correct samples	accuracy target 1	accuracy target 2	total accuracy
Training samples	32	31	93.75%	100%	96.87%
Testing Samples	608	595	96.38%	99.34%	97.86%

The RVM classification result vs. SNR is presented in Fig. 7.

When adding non-Gaussian noise energy is enough strong, an extracted feature effectivity is worsen, and causes the correctness rate drop to 90.55%. Continuing to add no-Gaussian noise, SNR drop from −3dB to −18dB. The correct rate remains above 80% when SNR is high than −12dB. The encouraging classification results prove feature extraction scheme and classifier owns robustness.

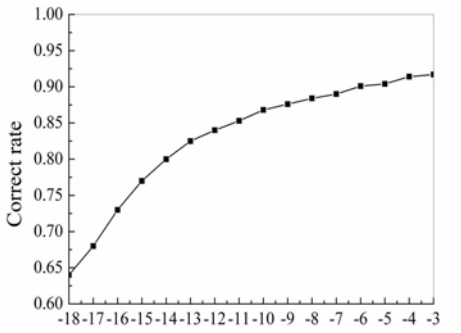


Fig. 7. Classification correct rate vs. SNR

5. CONCLUSIONS

A robust time-frequency feature extraction scheme is proposed, which are specially applied to suppress the influence of reverberation. It helps to produce more robust echoes feature in the case of echoes corrupted with impulse, heavy-tailed noise. The features based on modified STFT reflecting different target strength of two type of underwater mine shells. Then RVM classifier is developed for classifying underwater bottom mine shells. Even, on condition that the environment becomes worse by adding non-Gaussian noise, the performance of the algorithm is also acceptable. When SNR is −3dB, the correct rate is above 93%. Classification correct rate is above 85%, when SNR is −12dB. Compared with SVM, the test results on the acoustic backscattered data collected from two mine shells showed that RVM classifier using robust time-frequency features offered good performance in changing underwater situations, and achieve much sparse solution for classification. RVM uses fewer bases and yields a higher accuracy in this example. But much room for improvement, the first improvement will be to identify features that are more robust to environmental changes and multi targets. Another future work is to apply more kinds of window time-frequency distributions to represent backscattered signal feature.

6. REFERENCES

[1] D.W. Ricker, Echo Signal Processing, Boston:Kluwer Academic Publishers, 2003.

[2] M.R Azimi-Sadjadi, D. Yao, et al, “Underwater target classification using wavelet packets and neural networks”, IEEE Trans. Neural Networks, May 2000，11（3）：784-794.

[3] M.R Azimi-Sadjadi, D. Yao, et al, “Underwater target classification in changing environments using an adaptive feature mapping,” IEEE Trans. Neural Networks, May 2002， 13（5）: 1099-1111.

[4] D.H. Li, M.R. Azimi-Sadjadi, et al, “Comparison of different classification algorithms for underwater target discrimination”, IEEE Trans. Neural Networks, Jan 2004, 15（1）: 189-194.

[5] N. Dasgupta, L. Carin, “Time-reversal imaging for classification of submerged elastic targets via Gibbs sampling and the relevance vector machine,” Journal of Acoustical Society of America, April 2005, 117(4): 1999-2011.

[6] M.D. Billy, Determination of the resonance spectrum of elastic bodies via the use of short pulses and Fourier transform theory”, Journal of Acoustical Society of America, Feb 1986, 79（2）: 219-221.

[7] V. Vapnik, Statistical Learning Theory, New York: John Wiley, Pub., 1998.

[8] Q. Shie, Introduction to Time-frequency and Wavelet Transforms,Pearson Education,Inc., New Jersey,2002.

[9] M.E. Tipping, “Sparse Bayesian Learning and the Relevance Vector Machine”, Journal of Machine Learning Research, June 2001,Vol.1, 211-244.

[10] O. Williams, A. Blake, “Sparse Bayesian Learning for Efficient Visual Tracking”, IEEE Trans.on Pattern Analysis and Machine Intelligence, August 2005, 27（8）: 1292-1304.

[11] N. Cristianni, J. Shawe-Taylor, “An Introduction to Support Vector Machine and Other Kernel-based Learning Methods”, Cambridge University Press, Cambridge, 2000.

[12] P. A. Saillant, J. A.Simmons, et al, “A computational model of echo processing and acoustic imaging in frequency-modulated echolocating bats: The spectrogram and transformation receiver”, Journal of Acoustical Society of America, 94（5）: 2691-2712.

[13] M. Robinson, M.R. Azimi-Sadjadi, et al, “Multi-aspect target discrimination using hidden Markov models and neural networks”, IEEE Trans Neural Networks, Vol. 16, No.2, March 2005,pp. 447- 459.

[14] D. Yao, M.R. Azimi-Sadjadi, et al, “A Study of Effects of Sonar Bandwidth for Underwater Target Classification”, IEEE Journal of Oceanic Engineering, July 2002, 27（3）: 619-627.

[15] I. Djurovic, L. Stankovic, “Robust Wigner Distribution with Application to Instantaneous Frequency Estimation”, IEEE Trans Signal Processing, December 2001, 49（12）: 2985-2993.

[16] S. Barbarossa, O. Lemoine, “Analysis of nonlinear FM signals by pattern recognition of their time-frequency representations”, IEEE Signal Processing Letters, April 1996, 3（4）: 112-115.

[17] H.Y. Huang, “Underwater target classification based on time-frequency extraction and support vector machine”, Master Thesis, Hangzhou, Zhejiang University, China, June 2006, 43-46.

AN INTELLIGENT NEURAL NETWORK RULE EXTRACTION TECHNIQUE

Yunling Liu
College of Information and Electrical Engineering, China Agricultural University
Beijing, 100083, China
E-mail: lyunling@163.com
Jianjun Lu
Graduate School of Economics, Kyushu University
Fukuoka, 812-8581, Japan

ABSTRACT

Most of decision making studies focus primarily on developing classification models with high perspective accuracy without any attention to explaining how the classifications being made. We deals with the neural network rule extraction techniques based on the Genetic Programming (GP) to build intelligent and explanatory evaluation systems. GP is utilized to automatize the rule extraction process in the trained neural networks where the decision basically boils down to a binary classification problems. Simple and relevant classification rules are obtained by pruning weight among neurons to obtain simple but substantial binary expressions.

Keywords: Neural Network, Rule Extraction, Genetic Programming, Classification Rules.

1. INTRODUCTION

Numerous methods have been proposed in the literature to develop decision making and decision tables models. These models include traditional statistical methods such as multivariate discriminant function, logistic regression and even neural network techniques [1,2].There are also methods categorized to classification tree such as the entropy-based decision tree and statistical procedure to generate linguistic if-then rules (inductive learning) [3].In most of these studies, how to develop classification models with high perspective accuracy is mainly considered, whereas how the classifications being made is ignored.

This paper utilized the GP in the neural network rule extraction techniques to build intelligent and explanatory evaluation systems [4-7]. We used GP to automatize the rule extraction process in the trained neural networks where the decision basically boils down to a binary classification problems.

GP have been successfully applied to approximate the chaotic dynamics, where the approximated expression of dynamics is also utilized to control chaotic behavior [6,7]. Among recent developments in algorithms that extract rules from trained neural networks, popular techniques such as Neurorule, Trepan and Nefclass are known [4,5]. Even though Neurorule is looking at the internal structure of networks, the process for extracting the rules is complicated. Techniques using fuzzy inference and fuzzy sets prevent us to represent rules in ordinary usable forms like decision tables. Then, we focus on the capability of trained neural networks for the decomposition of input variables to reduce simplified rules. The processes to combine the input and output units are carried out by using the GP. After training neural networks using the discretized input variables, the pruning process of weight among neurons is applied to obtain simple but substantial binary expressions which are used as statements is classification rules. Then, the GP is applied to generate ultimate rules.

2. NEURAL NETWORKS AND RULE EXTRACTION

2.1 Algorithm of Neural Networks

Neural networks are mathematical representations inspired by the functioning of the human brain [4-7]. Especially, the multilayer neural network is typically composed of an input layer, one or more hidden layers, and an output layer, each consisting of several neurons. Each neuron processes its input and generates one output value which is transmitted to the neurons in the subsequent layer. All neurons and layers are arranged in a feed forward manner, and no feedback connections are allowed. The output of the neuron i is computed by processing the weighted inputs and its bias term as follows.

In the formula, the $w_{i,j}^{n,n-1}$ denotes the weight connecting the *jth* unit in layer *n* with the *ith* unit in layer $n-1$. The value z_j^{n-1} is the output of *j th* unit in layer $n-1$. Then, the input to i *th* neuron in layer n is obtained by

$$u_i^n = \sum w_{i,j}^{n,n-1} z_j^{n-1} \quad (1)$$

The output of neurons is calculated by using the transfer function emulating threshold logic which is called sigmoid function.

$$z_i^n = f(u_i^n - h_i^n) \quad (2)$$

Where h_i^n is the threshold value for the neuron. Usually, we use the sigmoid function for the function $f(.)$.

$$f(y) = 1/(1+\exp(a-y)) \quad (3)$$

The weight $w_{i,j}^{n,n-1}$ and bias h_i^n are the critical parameters of neural networks and need to be estimated during a training process which is usually based on gradient descent learning to minimize some kind of error function over a set of training observations.

The signal which is backpropagated from i *th* neuron in output layer N is obtained by using the difference between the output of output layer N and the prescribed observation d_i^N as follows.

$$\delta_i^N = (d_i^N - z_i^N)g(u_i^N), g(y) = f(y)(1-f(y)) \quad (4)$$

Similarly, the signal backpropagated from the i th neuron in layer n to neurons in layer $n-1$ is given as

$$\delta_i^N = df(u_i^n)/du \sum_k \delta_k^{n+1} w_{k,i}^{n+1,n} \quad (5)$$

In summary, the update for the weight and the threshold value is given as follows where $t, t+1$ are the time steps in the update.

$$\Delta w_{i,j}^{n,n-1}(t) = \eta \delta_i^n x_j^{n-1} \alpha \Delta w_{i,j}^{n,n-1}(t-1) \quad (6)$$

$$w_{i,j}^{n,n-1}(t+1) = w_{i,j}^{n,n-1}(t) + \Delta w_{i,j}^{n,n-1}(t) \quad (7)$$

$$\Delta h_i^n(t) = \eta \delta_i^n + \alpha \Delta h_i^n(t-1) \quad (8)$$

$$h_i^n(t+1) = h_i^n(t) + \Delta h_i^n(t) \quad (9)$$

Where η, α are constants to confirm the convergence.

2.2 Neural Network Rule Extraction

As universal approximators, neural networks can achieve significantly better predictive accuracy compared to models that are linear in the input variables. However, their complex mathematical internal workings prevent them from being used as effective management tools in real-life situations where besides having accurate models, explanation of the predictions being made is essential.

In the literature, the problem of explaining the neural network predictions has been tackled by techniques that extract symbolic rules from the trained networks. These neural network rule extraction techniques attempt to open the neural network black box and generate symbolic rules ith the same predictive power as the neural network itself.
In the decomposition algorithm such as the Neurorule, we start to extract rules at the level of the individual hidden and output units by analyzing the activation values, weights, and biases. Decompositional approaches then typically treat the hidden units as threshold units.

For example, the algorithm of the Neurorule is summarized as follows [4].

(Step 1) Train a neural network to meet the prespecified accuracy requirement.

(Step 2) Remove the redundant connections in the network by pruning while maintaining its accuracy.

(Step 3) Discretize the hidden unit activation values of the pruned network by clustering.

(Step 4) Extract rules that describe the discretized hidden unit activation values in terms of the networks inputs.

(Step 5) Generate rules describe the discretized hidden unit activation values in terms of the network inputs.

(Step 6) Merge the two sets of rules generated in Step4 and 5 to obtain a set of rules that relates the inputs and outputs of the network.

2.3 Neural Network Rule Extraction Based On the GP

We introduce the GP procedure in place of Step 4 through Step 6 in Neurorule algorithm. The overview of the GP procedure for generating rules is summarized as follows.

(Step 1) Discretize input variables Similar to Neurorule, the input variables are descretized by using threshold variables and are represented in binary forms (inputs).

(Step 2) Train a neural network using the discretized input variables to meet the prespecified accuracy requirement.

(Step 3) Remove the redundant connections in the network by pruning while maintaining its accuracy. Then, we obtain substantially important statements represented in binary forms.

(Step 4) Generate rules based on the GP using the binary representation as the terminal variables (statements) in the logical expression.

2.4 Comparison of Computational Complexity

In the design of multi-layer neural networks, we start from an oversized networks, and then gradually remove the irrelevant connections. When all connections to a hidden neuron have been pruned, this neuron can be removed from the network. The selection of for pruning is achieved by inspecting the magnitude of their weights. A connection with sufficiently small weigh can be pruned from the network without affecting the network's classification.

In Neurorule, once a trained and pruned network has been obtained, the activation values of all hidden neurons are clustered to simplify the rule extraction process. In the case of hyperbolic tangent hidden neurons, the activation values lie in interval [-1, 1]. A greedy clustering algorithm then starts by sorting all these hidden activation values in increasing order. Adjacent values are then merged into a unique discretized value as long as the class labels of corresponding observations do not conflict. The merging process hereby first considers the pair of hidden activation values with the shortest distance in between. We merge the hidden activation values by equivalent process using the χ-square test statistics

In Neurorule, the rule extraction process is three-hold, and is very complicated. At first, we obtain activation values of hidden unit by clustering (Step 3), and then describe tentative rules by using the activation values and outputs (Step 4). Even more, we must also generate tentative rules to describe the relation between activation values and inputs (Step 5). Then, finally we merge two corresponding rules obtained in Step 4 and 5 to generate the relations between inputs and outputs to get ultimate production rules.

We find two important drawbacks in the algorithm of Neurorules compared with the method proposed in the paper. Firstly, in the process for generating tentative rules in step 4 and 5, we can use conventional method to reduce induction rules based on the entropy.

Secondly, the reduction process is mainly based on finding the relations between activation values and inputs as well as outputs. We can have no statistical measure to assess the relevancy of combinations. In contrast to these drawbacks, the GP procedure provides us a simple and comprehensive algorithm for generating rules. Once we obtain discretized inputs, we can generate production rules in a simple way using the GP, which improves the fitness of rules based on the genetic operations. Even more, the relevancy of generated rules is directly obtained by observing the maximum fitness of individuals. Comparing the time necessary to optimize these parameters in neural networks, the time to execute GP procedure is very small.

Table 1 shows an example of computational complexity of Neurorule and GP method proposed in the paper in the execution time and steps of programs. We also show the computational complexity of Trepan system which is worse than Neurorule, and only the performance is listed in Table 1. without explaining the details of algorithm. As is seen from Table 1. The GP method of the paper is better than

Neurorule both in the complexity and computation time to obtain production rules.

Table 1. Comparison of computational complexity

	Neurorule	Trepan	GP
Program Size	1865 steps	3270 steps	385steps
Execution time	300	780	90

3. Production rules and the GP

3.1 Tree Representation of Production Rules

The prefix representation follows traditional representation by using the Lisp syntax [6-8].

For checking the validity of underlying parse tree, the so-called stack count (denoted as *StackCount* in the paper) is useful [8]. The *StackCount* is the number of arguments it places on minus the number of arguments it takes off from the stack.

Any two loci on the two parent genomes can serve as crossover points as long as the ongoing *StackCount* just before those points is the same. The crossover operation creates new offspring by exchanging sub-trees between two parents.

(Step 1) Generate an initial population of random composition of possible functions and terminals for the problem at hand.

(Step 2) Execute each individual (evaluation of system equation) in population by applying the optimization of the constants included in the individual. Then, assign it a fitness S_i for individual i giving partial credit for getting close to the correct output.

(Step 3) Select a pair of individuals chosen with a probability p_i based on the fitness.

(Step 4) Then, create new individuals (offspring) from the selected pair by genetically recombining randomly chosen parts of two existing individuals using the crossover operation applied at a randomly chosen crossover point.

(Step 5) If the result designation is obtained by the GP become larger than the prescribed value, then terminate the algorithm, otherwise return to Step 2.

3.2 Applying the GP to Rule Generation

In our system based on the neural networks, the input variables are discretized based on threshold values. Even more, by pruning the connections (weights) in neural networks, we have finally several binary expressions used for the input to the simplified neural networks. Then, these binary expressions are used as the propositions included in the logical expressions in the GP. For example, we define new logical variables x_i represented by input variable v_i such as

$$X_{kj} = \begin{cases} True, \ if \ v_i = j; \\ False, otherwise \end{cases} \tag{10}$$

where α_i is a constant reduced from selecting of threshold values.

Then, we find that the logical expressions included in the production rules are the same as the arithmetic expression using prefix representation by replacing the operands by the propositions, and the arithmetic operators by the logical operators. For example, numerical variables $x_i \rightarrow$logical variables x_i, arithmetic operators $+,\times \rightarrow$OR, AND.

We also define the fitness of individuals as the accuracy of rule generated by the rule corresponds to the underlying individual. To improve the fitness of individuals, we apply the GP operations to the logical expressions.

The fitness of individuals is evaluated as follows:

(1) calculation of arithmetic expressions

By substituting the value of input variables v_i, we can evaluate the binary arithmetic expressions included in propositions. For example, a set of specific values for v_1, $v2 \ldots v_n$ are used to obtain numerical values.

(2) interpretation of propositions

Since we know the values of arithmetic expressions, the logical value of logical variables x_i are obtained by combining the propositions and comparative operators such as equation (10).

(3) interpretation of logical formula

Finally, we can know the logical value of the whole logical formula (individual) by applying the logical operations among propositions.

As already mentioned, the value is compared with the prescribed observation *r* to calculate the fitness.

4. Conclusion

This paper used the neural network and binary classification to obtain simple and relevant classification rules by pruning weight among neurons to obtain simple but substantial binary expressions which are used as statements is classification rules generated by the Genetic Programming. For further works, the method will be applied to extract rules to predict specific actual problems.

5. References

[1] Y.H. Pao, Adaptive Pattern Recognition and Neural Networks, Addison-Wesley Publishing Co., Inc., 1989.

[2] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, U.K., 1995.

[3] R.Andrews, J.Diederich, et al, "A survey and critique of techniques for extracting rules from trained neural networks," Knowledge Based Systems, 1995, 8（6）: 373-389.

[4] B. Baesens, R. Setiono, et al, "Using neural network rule extraction and decision tables for credit-risk evaluation", Management Science, 2003,49（3）: 313-329.

[5] M.W. Craven, J.W. Shavlik, "Extracting tree-structured representations of trained networks", D.Touretzky, M.Mozer, Advances in Neural Information Processing Systems (NIPS), Vol.8, MIT Press, Cambridge, MA, 1998,24-30.

[6] J. Lu, S. Tokinaga, et al, "Explanatory Rule Extraction Based on the Trained Neural Network and the Genetic Programming", Journal of the Operations Research Society of Japan, 2006, 49（1）: 66-82.

[7] S. Tokinaga, J. Lu, et al, "Neural network rule extraction by using the Genetic Programming and its applications to explanatory classifications," IEICE

Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2005, E88-A（10）: 2627-2635.

[8] J.R. Koza, Genetic Programming, MIT Press, 1992.

The Application of BP Neural Networks based on Chaotic Analysis algorithm in Short-term power Load Forecasting

Jia Shen
Colleges of Mech. and Elec. Engineering, Agricultural University of Hebei
Baoding, Hebei province, China
Email:shen-jia@163.com
And
Qing Zhang*
Colleges of Mech. and Elec. Engineering, Agricultural University of Hebei
Baoding, Hebei province, China
Email: auhzhangq@163.com

ABSTRACT

Using chaotic theory, the paper analyzes the complicated time series of power load, and then concludes that the time series of power load is chaotic series. At this point a new neural network learning algorithm, chaotic analysis algorithm, is proposed in this paper. With chaotic analysis method, the paper reconfigures load series in the phase space, calculates the Branch dimension and the largest Lyapunov exponents, testes its uncertainty, and then gets its largest forecasting time scale. As a result, the number of input points in NN is determined, as well as BP neural network based on chaotic analysis algorithm is established. In this way, the algorithm overcomes the natural problem of the BP learning algorithm in feedforward NN, so that it has a good performance in convergence、Velocity、Error during NN training. The result shows that the BP Neural Network model based on chaotic analysis algorithm has a higher forecasting precision.

Keywords: Short-term Load Forecasting (STLF), BP Neural Network, chaotic analysis, Phase Space Reconfiguration Theory (PSRT), Lyapunov exponent.

1. INTRODUCTION

Short-term load forecasting(STLF) is a basal work in power system, and is the basic of production arrangement、maintenance、Operation scheme. Its forecast precision has a direct impact on economic benefit. For a long time, Electricity operators in home and abroad do a lot of researches on the method of STLF, propose a great deal of algorithm of load forecasting , such as time series method、NN method、Weighted compounding method、expert system method、Obscuring NN method[1,2]. But its astringency and adaptability are both limited in different degree, and the complexity and randomness of factor which can affect load change is hard to be depicted by mathematical model, so people devote themselves to exploration of some new algorithm.

On the basis of calculating the largest Lyapunov exponents to load series, the paper proposes that BP NN model should be built in the phase space for forecasting the power system, and test whether the forecasting method is effective through the operational parameter.

*Corresponding author: Qing Zhang

2. LARGEST LYAPUNOV EXPONENTS

If the chaotic time series is:

$$\{x(t_j)\} \quad j = 1, 2, \cdots, d \tag{1}$$

Based on phase space reconfiguration theory (PSRT) [3], the multidimensional phase space is established, and the substantial character of chaotic time series is distilled. If the intermission of time series is $\triangle t$, the time series of Eq. (1) is taken into phase model distribution of n-dimension phase space. Then the jth n-dimension phase point is:

$$Y(t_j) = (x(t_j),\ x(t_j + \tau), \cdots, x(t_j + (n-1))) \tag{2}$$

Here, $\tau = k \triangle t$, k=1,2,⋯,The τ is delay time, and n is embedded dimensionality. Eq. (2) can be unfolded into $p=d-(n-1)\ \tau$ phase points in n-dimension phase space, and it describes evolution locus of the system in phase space. These phase points train the weight of network as input specimen of NN.

Chaotic system is very sensitive to the initial value. As long as there are small changes in initial conditions, the system will track the evolution over time to index times the speed of the original track separation, and conceals the real state of the system after some time. This indicates the unpredictability of the long-term behavior of the system is unforecastable. For this reason, the part of pull-up particle of chaotic system is not steady, and finally the locus drops onto the same chaotic pull-up particle of phase space. When the largest Lyapunov exponent is bigger than zero(λ_1>0), it show that there is chaotic pull-up particle, and that chaotic degree is measured by the size of Lyapunov exponents[4]. Moreover, the reciprocal(1/ λ_1)of largest Lyapunov exponents also show the largest forecasted time scale. If the time scale is exceeded, and error won't be controlled, the real information of the system will be concealed.

Man such as Wolf and so on [5] propose that the largest Lyapunov exponents should be estimated by the way based on phase locus evolvement. The main step is:

1). Reconfigure the phase space.

2). If the initial phase point is $Y(t_0)$, and the nearest point is $Y'(t_0)$, the distance $L(t_0)$ can be calculated. If delay time of evolution is τ, it will get evolution state vector $Y(t_0+\tau)$ of $Y(t_0)$and evolution state vector $Y'(t_0+\tau)$ of $Y'(t_0)$, then distance $L'(t_0)$ should be calculated; find $Y'(t_1)$ in $t_1=t_0+\tau$, then trace and calculate $L(t_1)$,$L'(t_1)$; repeat this process n times, and then the largest Lyapunov exponent λ_1 is calculated:

$$\lambda_1 = \frac{1}{t_M - t_0} \sum_{i=1}^{N} \log_2 \frac{L'(t_i)}{L(t_i)} \qquad (3)$$

3. BP NEURAL NETWORK MODEL BASED ON CHAOTIC ANALYSIS

3.1 Network model

Feedforward ANN（BP Network）would be regarded as high nonlinear mapping from input to output, namely $f(x)=y$, $f: Rn \rightarrow Rm$. The structure of the network is agile, and it doesn't follow fixed pattern. In 1987 Hecht-Nielsen demonstrated that a 3-tier BP network could achieve the random mapping that from n dimension to m dimension [6]. The requested problem determines the number of the nerve cell which in input and output tier of the network. Some scholars propose the empirical formula about the number of nerve cell of concealing cell, so the paper make sure the number of concealing cell by testing wrong way. On the other word, it quest the smallest structure (the least concealing cell) in the training network, and the network has a higher fit accuracy and forecasting precision. [7]

There is a 3-tier NN structure in the paper. If the dimension of input tier is n in the network, if the number of input nerve cell is n which is the best inset dimension in the chaotic phase space, and if its output dimension is one (forecasting output), the dimension of the concealing cell is fixed by testing wrong way during learning and training on network.

If there are two random phase points—$Y(t_i)$ and $Y(t_i+T)$ in the chaotic phase space (T is forecasting period), their relation of evolutive function $f: Y(t_i+T) = f(Y(t_i))$ is fitted by BP Neural Network. Forecasting point $Y(t_i+T)$ is n dimensional vector, but in reality former $(n-1)$ dimension is historical data. It is simple that network output is the nth component of $Y(t_i+T)$.

To a random nerve cell j of nonlinear cell, there is a input and output relation—

$$\left.\begin{aligned} u_j &= \sum \omega_i x_i - \theta_j \\ y_j &= f(u_j) \end{aligned}\right\} \qquad (4)$$

here, x is the input of nerve cell; ω is the connecting weight; θ is the threshold value; y is the output of the nerve cell; f is the excitation function. We usually use Sigmoid function, namely $f(u_j) = 1/(1+e^{-u_j})$.

3.2 Algorithm

The study of the NN connecting weight and threshold value reversely spreads algorithm (BP algorithm) according to the error, which is a study algorithm having teacher. The study processes consists of positive、negative process. Firstly, every network output is calculated in the positive direction according to Eq. (4). Secondly, the error between output and teacher value is calculated. Thirdly, the connecting weight and threshold value are reversely corrected with error. Finally, the two process are repeated until the error between output and stylebook value is small enough to some set-points ε.

If k is the number of input learning specimen; $x^k=(x_1^k, x_2^k, \cdots, x_n^k)$, y^k is forecasting value of output; t^k is relevant teacher value; the object function is Eq.(5)

$$E = \sum_k E_k = \frac{1}{2} \sum_k (t^k - y^k)^2 \qquad (5)$$

Connecting weight and threshold value are corrected in the way of gradient descent. The algorithm of additional momentum term can avoid shocking and accelerating convergence:

$$\omega_{sq}(n_0+1) = \omega_{sq}(n_0) - \eta \frac{\partial E}{\partial \omega_{sq}(n_0)} + \alpha \Delta \omega_{sq}(n_0) \qquad (6)$$

Here, ω_{sq} is a random connecting weight or threshold value; n_0 is the time of repeating; η is training velocity; ω_{sq} is correcting value; α is momentum factor.

3.3 Forecasting steps

BP Neural Network model is used to forecast. The detailed steps like this:

1) The paper selects the best delay time and inset dimension by chaotic analysis, establishes multi-dimensional phase space of chaotic time series, and constitutes the learning specimen and teacher value;

2) Determination of the NN structure. The inset dimension determines the number of the nerve cell of input, and the number of concealing nerve cell is preferentially selected by testing wrong way in the learning of network;

3) Learning of network. The paper randomly puts up the original value of the connecting weigh of the network, and calculates the output and the correcting weigh value through formulas Eq. (4) and formulas Eq. (6) until the error is in the allowed range, or until the time of repeat reaches the advance value;

4) Forecast model. If the foregone phase point is input to NN, the forecasting value is the output of network.

4. ANALYSIS OF EXAMPLE

Using STLF of BP neural network based on chaotic analysis [8], the paper compare forecasting data with actual number of selling electricity .the data is from 2005-6-7 to 2005-8-29 in a power bureau Baoding. Limited in the length, the paper only gives the data on the 2005-7-13(refer to table 1)。 According to the Table 1, the forecasting error is usually less than 2%, and the 80% error of data is less than 1%. It has high precision.

5. CONCLUSIONS

The NN model based on chaotic analysis is applied to STLF of power system. The paper presents the largest scale of forecasting time through calculating the saturated inset dimension and the largest Lyapunov exponent of data time series in Baoding. On the basis of that, the paper determines the number of input points in the NN, establishes STLF based on BP neural network. When the BP forecasting model based on chaotic analysis is applied to the forecasting of nonlinear time series, using EP evolutionary algorithm, the paper selects suited nonlinear feedback terms through the random nonlinear feedback terms, which has a good performance in nonlinear forecasting. Through testing, it has

a satisfying result and a good performance in self-adaptive ability、 robustness、 precision and currency. Meanwhile, the paper uses the forecasting way based on BP neural network, which is good for establishing forecasting system of real-time load.

Table 1. Actual forecasting result on 2005-7-13

T /h	Actual Value /MW	forecasting value/MW	Relative Error /%
0	105.89	106.39	0.47
1	113.47	114.13	0.58
2	107.12	107.49	0.35
3	104.78	106.04	1.20
4	110.21	111.19	-0.89
5	112.79	113.67	0.78
6	108.17	109.90	1.60
7	106.54	105.58	-0.95
8	105.04	104.91	-0.12
9	109.30	108.46	-0.77
10	107.41	109.74	2.17
11	110.91	109.80	-1.00
12	112.79	112.19	-0.53
13	110.76	114.86	0.37
14	111.17	110.41	-0.68
15	106.86	107.88	0.95
16	107.47	106.35	-1.04
17	104.86	105.66	0.76
18	103.71	103.17	-0.52
19	108.57	108.86	0.27
20	102.89	102.48	-0.40
21	112.47	113.46	0.88
22	109.47	108.72	-0.69
23	105.59	103.93	-1.57

6. REFERENCES

[1] J.H. Lu, Y. Zhan et al, Nonlinear Chaotic Improved Model on Short-Term Load Forecasting in Power System [J]. Proceedings of the CSEE, 2000, 20(12): 80-83.

[2] Z.S. Liang, L.M. Wang, et al, short-term load forecasting using chaotic theory in power system [J]. Control and decision-making, 1998, 13(1): 87-90.

[3] N.H. Packard, J.P. Crutchfield, et al, Geometry from a time series [J]. Phys Rev Lett, 1980, 45(9):712-716.

[4] F.S. Chen, Chaos and Its Application. Beijing: China Electric Power Press, 1998.

[5] A. Wolf, J.B. Wsift, et al, Determining Lyapunov Exponents from a Time Series. Physica D 16,1985，285-317.

[6] R. Hecht-Nielsen, Komogrov's mapping neural network existence theorem [J]. Proceedings of the international conference on Neural Networks, IEEE Press, 1987, (3): 11-13.

[7] G.d. Liu, J. Ding, Discussion on problems of BP neural networks applied to hydrological prediction [J]. Journal of HydraulicEngineering, 1999,(1):65-69.

[8] J.H. Lu, J.A. Lu, et al, Analysis of Chaotic Time Series and its Application [M] Wuhan: Wuhan University press, 2002.

An Optimal Fuzzy Controller for Linear Motor Drive Based on Genetic Algorithm

HaiJun Wang, YiMin Chen, ShaoLiang Wei
School of Computer Engineering and Science, Shanghai University
Shanghai, 200072, China
Email: whjcc9939@163.com

ABSTRACT

This paper presents a method to optimize fuzzy PID controller's scaling factors by utilizing improved Genetic Algorithm (GA), the scaling factors can regulate the outputs of the controller in terms of the system outputs feedback, thereby to improve the inferential rules and membership functions. Nowadays many fuzzy PID controllers are largely used in industrial automation, but its inferential rules are mainly determined by some technical experts or operators, so the rules take on much individual subjective experience, and it will not be suitable for the control systems under some circumstances, especially with great disturbance and more unexpected conditions. The simulated experiment demonstrates the proposed optimal controller's effectiveness and feasibility.

Keywords: Fuzzy Control, Genetic Algorithm, PID Control, Scaling Factor, Linear Motor.

1. INTRODUCTION

Conventional Proportional–Integral–Derivative (PID) controllers have been well developed and applied for about half a century, and are extensively used for industrial automation and process control today. The main reason is due to its simplicity of operation, ease of design, inexpensive maintenance, low cost, and effectiveness for most linear systems, the continuous equation of PID is as below:

$$u(t) = Kp[e(t) + (1/T_i)\int_0^t e(t)dt + T_D de(t)/dt] \quad (1)$$

Where: Kp is the proportional parameter.
Ki is the integral parameter, $Ki = (KP*T)/T_i$.
Kd is the differential parameter, $Kd = (Kp*T_D)/T$.
T is the sample time.

However, it is known that conventional PID controllers generally can not work well enough to the nonlinear systems, higher order and time-delayed linear systems, especially to the particularly complex and vague systems which have no any precise mathematical models [5, 7]. To overcome these difficulties, various types of modified conventional PID controllers such as auto tuning and adaptive PID controllers were developed lately.

As a paradigm to deal with unknown dynamic systems under uncertain environments, Fuzzy PID control has achieved success in many practical systems, such as subway systems, nuclear reactor control, and automobile transmission control.

A fuzzy controller is generally composed of four stages of which are fuzzification, rule base, inference engine, and defuzzification. A fuzzy controller is generally designed in the light of experience and experts' knowledge [4, 6]. However, such knowledge is not always easy to acquire, moreover, even the system is designed by skillful and well experienced experts, whether it is well enough optimal designed, it would be uncertain or unreliable to work normally. To improve this situation, we optimized the fuzzy PID control system utilizing genetic algorithm, more detailed information will be described in section 2.

This paper is organized as follows, in section 2, the basic concepts of fuzzy set theory are reviewed and the scaling factor is discussed. In section 3, we introduced the genetic algorithm, analyzed the linear motor drive's mathematical model, gave the overall control framework and demonstrated the example. Finally in section 4, a conclusion is given.

2. REGULATING FUZZY INFERENTIAL RULES

Most of the fuzzy controllers used in industrial automation are belong to the type with two inputs and one output, the input variables are error (E) and the velocity of error (EC). The fuzzy controller will use the variables of E and EC to infer the fuzzy output in terms of the expert rules, after being defuzzified, the output can be applied to control the plant. In this process, the most effective way to regulate the fuzzy rules is the quantified factors or scaling factors' self-tuning, according to the value of rise time, overshoot, static error, etc., the controller will tune the quantified factors or scaling factors to improve the fuzzy rules.
For example, the fuzzy rules are described as below:

if $E = NB$ and $EC = NB$ then $U = PB$
if $E = NB$ and EC = PS then $U = PM$
……
……
if $E = NM$ and $EC = PM$ then $U = PS$

Assumed the scaling factor is x1, and u is output variable, Then:

$u = f(U) * x1$ ($U = PB$ or PM, PS…,)

$f(U)$ is the function of defuzzification, if we regulate $x1$ to $x2$, ($x2 < x1$), then:

$$u' = f(U) * x2$$

Here $u' < u$, but E and EC do not change any more, the instance of tuning scaling factor is equivalent to regulating the fuzzy rule:

"if $E = NB$ and $EC = NB$ then $U = PB$"

is regulated as below:

"if $E = NB$ and $EC = NB$ then $U = PM$"

With the appropriate scaling factors, the later can be a much more accurate control rule to the real environment conditions than the former, it reduce much more personal subjectiveness, so it can improve the controller's performance and get the desired result. An illustration is shown in Fig. 1..

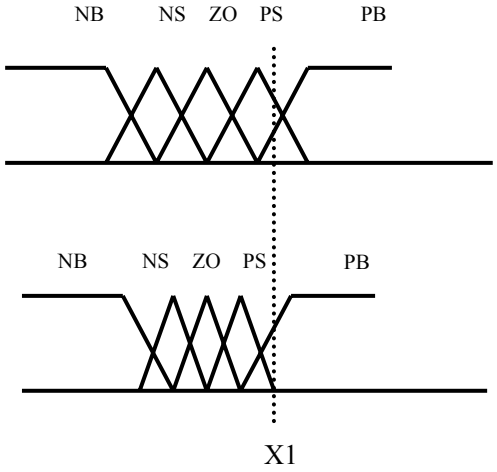


Fig. 1. Membership function based on different quantified factors

3. DESIGN AND SIMULATION

3.1 Genetic Algorithm

Genetic algorithm is general-purpose stochastic search techniques based on natural genetic and evolution mechanisms [1, 2]. During the last two decades, GA has been successfully applied to several complex optimization problems in business, science, and engineering. One of the most interesting aspects of GA is that they do not require any prior knowledge, space limitations, or special properties of the function to be optimized such as smoothness, convexity, inimicality, or existence of derivatives. They only require the evaluation of the so called "fitness function" to assign a quality value to every solution produced. Another interesting feature of GA is that they are inherently parallel. Instead of using a single point and local gradient information, they evolve a population of candidate solutions where each individual represents a specific solution not related to other solutions. Therefore, their application to large-scale optimization problems can be easily implemented on parallel machines resulting in a significant reduction of the required computational time. Assuming a randomly generated initial population, genetic evolution proceeds by means of three basic genetic operators: Parent selection, crossover, and mutation [8, 9, 10].

1) Select operator: based on the evaluation to the individual candidate's fitness, the candidate with higher fitness value will be reserved to next generation after select operation. Assuming *S* is the population size, *Fi* is the fitness value of candidate *i*, *Pi* is the selective probability of the candidate *i*, so we can get the select operation expressions as below:

$$Pi = Fi / \sum_{0}^{s} Fi \qquad i = 1, 2, \ldots, S \qquad (2)$$

2) Crossover operator: it is important to get new gene from exchanging some part of two gene candidates under some certain probability, the new gene maybe have better ability than its parents. Assuming *Xi, Xj* are the parent gene, and the new gene are $X^1 i$, $X^1 j$, so the crossover operation can be expressed as below:

$$\begin{aligned} X^1 i &= \alpha Xi + (1-\alpha) Xj \\ X^1 j &= \alpha Xj + (1-\alpha) Xi \end{aligned} \qquad (3)$$

where α is crossing parameter, its value is defined by the fitness value, assuming $f_.$ is the bigger fitness value between the two gene which is to be crossed, f_{max} is the biggest fitness value among the whole candidates, f_v is the mean value of the whole candidates. When $f > f_{max}$, the parameter α is defined by the equation (4) :

$$\alpha = k_1 (f_{\max} - f) / (f_{\max} - f_v) \qquad (4)$$

While $f < fmax$, α can be assigned a random value less than 1.

3) Mutation operator: utilizing the genetic mutation, we can extend the searching candidates' space. Assuming that $Xi \in [Vmin, Vmax]$,the value after mutation is Xi^1.

Then the value of Xi^1 is expressed as below:

$$Xi^1 = V_{\min} + \alpha (V_{\max} - V_{\min}) \qquad (5)$$

Where: α is the same as mentioned above (see Eq. (4)).

According to the above mentioned three genetic operators, we know that the candidates with smaller fitness value have the bigger probability of crossover and mutation, but the individuals with higher fitness have the bigger probability being reserved into the next generation with less change, so it is essential to maintain the variety of the population size to prevent the prematurely, and ensure the convergence of the algorithm. Genetic algorithm can utilize the clustering way to research the object solution in parallel, it only depend on the value of fitness function, has no request of the function's continuity and derivativeness, it is easy to use in many optimal applications and has a good robustness and reliability.

3.2 Frame Chart of Controller

Firstly we initialize the three basic parameters of PID with a constant value:

Kp = *Kpc*=5.8;
Ki = *Kic*=0.0896;
Kd = *Kdc*=0.0451;

The initial value of proportional parameter is confined to [0~20], the integral parameter is confined to [0~1] and the differential parameter is also confined to [0~1]. Here the initial values of *Kpc*, *Kic*, and *Kdc* have better performance than those parameters which is regulated by Z-N method.

Secondly the fuzzy controller's three outputs are defined as the increment of *Kp*, *Ki*, *Kd*, i.e. △*Kp*, △*Ki*, △*Kd*.

Every chromosome was constructed by the three scaling factors: *Pkp*, *Pki*, *Pkd*, and every one in the population cluster will be evaluated by GA in terms of the fitness function. Finally we can get the three basic parameters:

Kp = *Kpc* +△*Kp* * *Pkp*
Ki = Kic + △*Ki* * *Pki*
Kd = Kdc +△*Kd* * *Pkd*

3.3 Linear Motor Drive System

Our simulated system is the linear motor based on transport system, because the linear motor leaves out the middle cushion parts, such as chain, gear and so on, so it takes on characters as much more nonlinear effects and stronger disturbances, under the consideration of its push power, several friction force in practice, we derived its mathematic model as below:

$$\begin{aligned}(M+\Delta m)a_c &= F_1-\mu_1[(M+\Delta m)g+KF_1]\\ &-(\mu_2-\mu_1)\exp(-av)[(M+\Delta m)g+KF_1]\\ &-\mu_3[M+\Delta m+(KF_1)/g] \end{aligned} \quad (6)$$

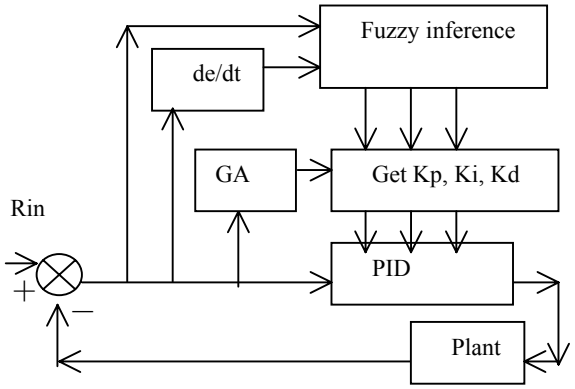


Fig. 2. The frame chart of fuzzy-PID control based on GA

Where in the Eq. (6), the first item is the linear motor's push force (*F1*), the second one is slide friction force, the third one is static friction force, and the fourth item is viscous friction force. The influence of centripetal force is also being taken into account, i.e. the coefficients of *K* and *F1* among all items in the formulation above. Considering the current loop effects of the linear motor, all relevant parameters measurement, we use the two-order transfer function as its mathematic model:

$$G(s)=K/(Ms+d)(Ls+r) \quad (7)$$

Where
M is the loaded weight, *M*= 8(kg),
d is the friction force coefficient, *d*=0.8,
L is the inductance of the motor, *L*=0.0165(L),
r is the resistance, r=1.7(Ω), *K*=30.
In our experiment, we add some random noise on the transfer function (Eq.7) during the simulated process.

3.4 Algorithm Design

The optimized variables are *Pkp,Pki,Pkd*, they're the scaling factors of proportional coefficient, integral coefficient and differential coefficient, the value range of *Pkp* is limited between[0.5,5], *Pki, Pk*d are limited between [0.5,1.5], they are real encoded in genetic algorithm. Because during the control process, the parameter of *Pkp* has more impact to the rise time, overshoot and stabilization time than the others, so its value range is larger than other two parameters. GA needn't much more external information in the search procedure, it only require the evaluation of the so called "Fitness Function" to assign a quality value to every solution produced. So it is very important to select a Fitness Function, this quality value is used as a comparative measure of each solution against other members of the population, its formulation is a key problem for the application of GA to practical optimization problems. In our approach, the Fitness Function is of the form is:

$$K=\int_0^{\infty}(\omega_1|e(t)|+\omega_2 u^2(t)dt+\omega_3\cdot t_u \quad (8)$$

Where:
e (t) is the system error,
u (t) is the output of the controller,
t_u is the rise time,
w1=0.999,
w2=0.001,
w3=2.0.
w1, w2, w3 are the weighted value.

Parent selection is a simple procedure where two individuals are selected from the parent population based on their fitness value. Solutions with high fitness values have a higher probability of contributing new offspring to the next generation [3]. In our approach, a simple roulette-wheel selection rule is employed. Elitism ensures that the best solution found is never lost when evolving from one generation to the next generation. The best solution of each generation replaces a randomly selected individual in the new generation. A uniform crossover operator with a probability of 0.9 was adopted. The mutation probability was chosen equal to 0.008. Initial population size *P* is set to 60, the generations *G* is 100. After optimized by GA, the values of *Pkp*, *Pki*, *Pkd* is:

Pkp =-3.5737;
Pkd = 0.9946;
Pki = 0.5627;

The whole simulated code is omitted, anyone who has interest in it can contact with the author for it. The simulated experimental results can be found at Fig. 3. to Fig. 7., we can see that the rapid response of fuzzy PID controller optimized by GA is better than it without optimized in Fig. 7., it is very important to some transport systems which need to upload or offload goods and materials at exact positions, otherwise the whole system would go to failure.

4. CONCLUSION

In this paper, an optimal fuzzy PID controller for linear motor drive based on genetic algorithm was presented. The scaling factors can be generated randomly by genetic algorithm using real coding, which has advantages related to the convergence time (few generations) and simplicity to assemble the individuals. The initial constant value, i.e. Kpc, Kic, Kdc mentioned in this paper, would be a crucial factor to a fuzzy controller's performance, sometimes improper value will lead the fuzzy controller to failure. But the scaling factor can weaken this adverse case by regulating the parameters more appropriate to real conditions. As compared with fuzzy PID controller, experimental results

verified the effectiveness and robustness of the proposed optimal approach.

5. ACKNOWLEDGEMENTS

This work is supported by the Shanghai Science and Technology Development Fund (015115042). I will thank the anonymous reviewers' suggestions and recommendations

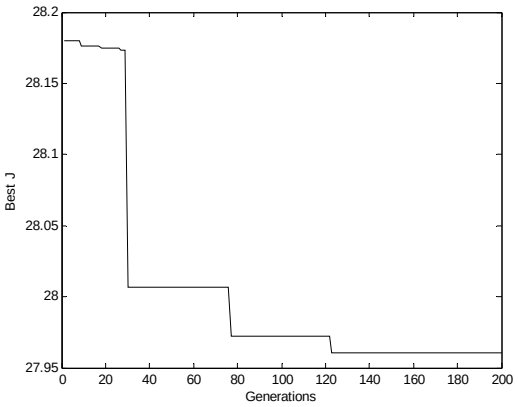


Fig. 3. Trajectory of fitness function based on GA

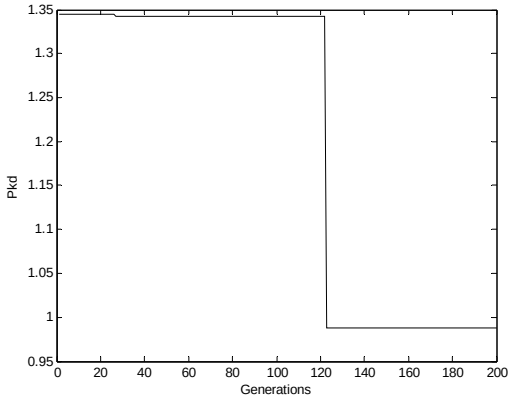


Fig. 4. Trajectory of scaling factor Pkd based on GA

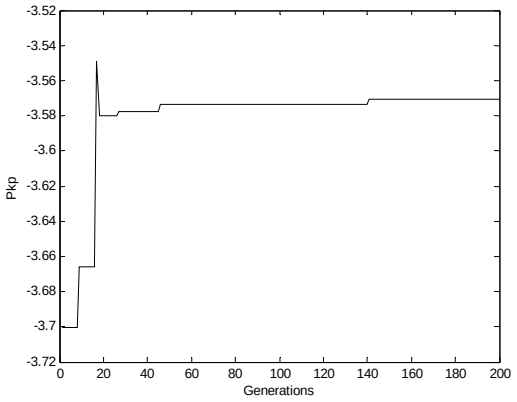


Fig. 5. Trajectory of scaling factor Pkp based on GA

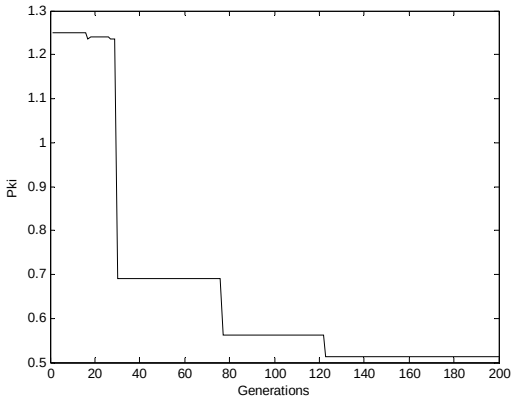


Fig. 6. Trajectory of scaling factor Pki based on GA

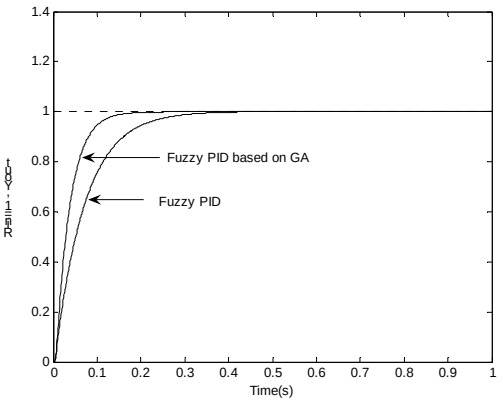


Fig. 7. Comparison of the two algorithms

6. REFERENCES

[1] C.E. Zoumas, A.G. Bakirtzis, et al, "A genetic algorithm solution approach to the hydrothermal coordination problem", IEEE Transactions on Power Systems, Aug. 2004, 19(3): 1356-1364.

[2] G. Golino, "Improved genetic algorithm for the design of the optimal antenna division in sub-arrays: a

multi-objective genetic algorithm", IEEE International Conference on Radar, 9-12 May 2005, 629-634.

[3] J.S. Petko, D.H. Werner, "The Evolution of Optimal Linear Poly fractal Arrays Using Genetic Algorithms", Antennas and Propagation, IEEE Transactions, Nov. 2005, 53(11): 3604-3615.

[4] D. Fatta, G. Hoffmann, et al "A genetic algorithm for the design of a fuzzy controller for active queue management", Systems, Man and Cybernetics, Part C, IEEE Transactions, Aug. 2003, 33(3): 313-324.

[5] K. Natarajan, "Robust PID Controller Design for Hydroturbines", IEEE Transactions on Energy Conversion, sep. 2005, 20(3): 661-667.

[6] L.A. Zedah, "Fuzzy sets", Information Control, July 1965, 8(4): 338-353.

[7] B.G. Hu, G.K.I Mann, et al, "A systematic study of fuzzy PID controllers function-based evaluation approach: An overview", IEEE Transactions on Fuzzy systems, Oct.2001, 9(5): 699-712.

[8] K. Chellapilla, "Combining mutation operators in evolutionary programming", IEEE Trans. Evolve. Computer. Sep.1998, 2(3): 91-96.

[9] L. Jiao, L. Wang, "A novel genetic algorithm based on immune", IEEE Trans. on Systems, Man and Cybernetics, Sep. 2000, 30(5): 552-561.

[10] S.H Ling, F.H.F Leung, et al, "A novel genetic-algorithm-based neural network for short-term load forecasting", IEEE Transactions on Industrial Electronics, Aug.2003, 50(4): 793-799.

On The Modeling and Application Of Rbf Neural Network

Liping Qu
Information Engineering school, BeiHua University, Jilin City,Jilin Province132021,China
Graduate School of Science and Technology, Chiba University, Chiba City,Chiba263-8522,Japan
Email:qulipingmine@sohu.com
Jianming Lu, Takashi Yahagi
Graduate School of Science and Technology, Chiba University, Chiba City,Chiba263-8522,Japan
Email:jmlu@faculty.chiba-u.jp
Email:yahagi@faculty.chiba-u.jp

ABSTRACT

At present, the high power DC graphitizing furnace is widely applied, but the temperature-control accuracy is not perfect. In order to improve the temperature-control, we decided first modeling the high power DC graphitizing furnace and then applying the intelligent control. Here we studied the modeling strategy of RBF neural network. Two approaches for the selection of centers of the RBF neural network are discussed. All the simulated results show that the discussed approaches are effective.

Keywords: RBF neural network, Selection of centers, DC graphitizing Furnace, Orthogonal Forward Regression, Direct Typical-Point Selection.

1. INTRODUCTION

By far, people have found many different neural network models for different objects. All these models work very effectively, but every model has it's own restriction.

Concerning control systems, the multi-layer perceptron (MLP) with back-propagation algorithm is most commonly used, and there have been many modifications for its learning algorithm. But the BP algorithm is based on the steepest descent, and so it has its shortcoming as follow: the convergence speed is very slow and it is easy to converge to the local minimum. From the point of view of "Learn=approximation by mapping"[1], T.Poggio proposed the learning network with symmetrical Gaussian function as typical basis function according to regularization techniques, which is called RBF (Radial Basis Function) network. The general approximation ability of RBF network provides the theoretical basis for representing complex systems and has the advantages of speedy convergence and strong approximation ability for nonlinear modeling and identification[2,3].

The RBF network is two-layer processing structure. The hidden layer consists of an array of nodes. Each node contains a parameter vector called center. The node calculates the Euclidean distance between the center and the network input vector, and passes the result through a non-linear function. The output layer is essentially a set of linear combiner. The structure of RBF network is shown in fig. 1.

The overall input-output of RBF network is a mapping[4,5], f: $R^n \rightarrow R^m$, i.e.

$$fi(x) = \sum_{j=1}^{n} \eta_{ji}\Phi_j = \sum_{j=1}^{n} \eta_{ji}\Phi(\|x - c_j\|, \rho_j), 1 \le i \le m \qquad (1)$$

Where η_{ji} are the weights of the linear combiners, $\|.\|$ denotes the Euclidian norm, ρ_j are some positive scalars called widths, Φ (.,ρ) is a function from $R^+ \rightarrow R$, c_j are the RBF centers, n is the number of node in the hidden layer. Some typical choices of Φ (z,1) are thin-plate –spline function,the multi-quadric function, the inverse multi-quadric function and the Gausian function. Here we adapt the Gausian function.

The RBF network can approximate any non-linear function and can be used to represent accurately an arbitrary continuous function. The key of RBF network is the selection of the center of the hidden layer. If the centers are adequately selected, only a small number of nodes can gives very accurate approximation. At present, for practical applications some researchers use expert system for the selection of centers. In this paper, we study the following two methods: the orthogonal forward regression algorithm and direct selection of typical points.

2. ALGORITHM WITH ORTHOGONAL FORWARD REGRESSION

The general linear regression model is as follow[6]

$$y(t) = \sum_{j=1}^{M} \Phi_j(t)\theta_i + \varepsilon(t) \qquad (2)$$

The RBF (Radial Basis Function) is a special case of the above model, where Φi(t) are known as the regressor which are some non-linear function of the delayed input, output and predicted error. Φi(t) can be written as[7,8]

$$\Phi_i(t) = \Phi_i[x(t)] \qquad (3)$$

$$x(t)=[y(t-1),...,y(t-n_y),u(t-1),...u(t-n_u),\varepsilon(t-1),...\varepsilon(t-\varepsilon_l)] \qquad (4)$$

$y(t)$ can be called the related variable. Suppose $\varepsilon(t)$ is independent with regressors $\Phi i(t)$, then the set of $\Phi i(t)$ can be considered as the expanded model set. Although model(6) is in linear form, it can represent non-linear system, obviously the centers related to the fixed non –linear function $\Phi i(t)$ correspond to the regressors $\Phi i(t)$. Hence, the RBF expanded equation can be considered as the special case of the expanded model set, and the problem of selecting centers subset from candidated centers can be considered as selecting significant regressors from given expanded model set. The approach of selecting model subset from equation(6) is called the orthogonal forward regression. The following gives the detail.

Eg.(6) can be written in matrix form[9]

$$Y = \Phi\theta + E \qquad (5)$$

Where

* This paper format should be strictly adhered.

$$\left.\begin{aligned} Y &= [Y(1),...,Y(N)]^T \\ \Phi &= [\Phi_1,...,\Phi_M] \\ \theta &= [\theta_1,...,\theta_N]^T \\ E &= [\varepsilon(1),...,\varepsilon(N)]^T \end{aligned}\right\} \quad (6)$$

$$\Phi_i = [\Phi_i(1),...,\Phi_i(N)]^T, 1 \le i \le m \quad (7)$$

Φ is considered as the regression matrix and can be decomposed as

$$\Phi = WA \quad (8)$$

Where

$$A = \begin{Bmatrix} 1 & \alpha_{12} & \alpha_{23} & ... & \alpha_{1M} \\ & 1 & \alpha_{23} & ... & \alpha_{2M} \\ & & ... & ... & \\ & O & & ... & \alpha_{M-1,M} \\ & & & ... & 1 \end{Bmatrix} \quad (9)$$

$$W = [w_1,...,w_M] \quad (10)$$

W is N×M matrix with orthogonal column satisfying

$$W^TW = H \quad (11)$$

H is positive diagonal matrix

$$H = \begin{Bmatrix} h_1 & ... & O \\ & ... & . \\ O & ... & h_M \end{Bmatrix} \quad (12)$$

Where

$$h_i = w_i^T w_i, 1 \le i \le M \quad (13)$$

The equation(9) can be written as

$$Y = (\Phi A^{-1})(A\theta) + E = Wg + E \quad (14)$$

Where

$$A\theta = g \quad (15)$$

As ε*(t)* is independent with Φi*(t)*,hence

$$g = H^{-1}W^TY \quad (16)$$

Or

$$gi = \frac{w_i^T Y}{w_i^T w_i}, 1 \le i \le M \quad (17)$$

In general, the number of all the candidate regressors *M* is very large. But an adequate model may only requires *Ms*(«*M*) significant regressors. These regressors can be selected using the orthogonal forward regression. From (18) ,we have

$$Y^TY = \sum_{i=1}^{M} g_i^2 w_i^T w_i + E^TE \quad (18)$$

It is seen, each Wi may be considered as a part of relative variable square error. The error reduction ratio due to w_i can be defined as

$$(err)i = \frac{g_i^2 w_i^T w_i}{Y^TY} \quad (19)$$

Based on this ratio, a simple and effective procedure can be derived for selecting a subset of significant regressors by using forward regression approach. If at the jth step, the maximum [err]j is generated, then this regressor is selected. The procedure is terminated when

$$1 - \sum_{j=1}^{M} [err]_j < \rho \quad (20)$$

Where $0<\rho<1$ is a chosen tolerance. And we say, the optimal number of RBF network centers is *Ms*.The related regressor selection procedure is summarized as follows: At 1st step, for $1 \le i \le M$, Compute

$$w_1^{[i]} = \Phi_i \quad (21)$$

$$\gamma_{1q}^{[i]} = (w_1^{(i)})^T y_q \Big/ (w_1^{(i)})^T (w_1^{(i)}), 1 \le q \le m \quad (22)$$

$$[err]_1^{(i)} = (\sum_{q=1}^{m} (g_1^{(i)})^2)(w_1^{(i)})^T w_1^{(i)} \Big/ trace(y^T y) \quad (23)$$

Find

$$[err]_1 = [err]_1^{(i_1)} = \max([err]_1^{(i)}, 1 \le i \le M) \quad (24)$$

Select

$$w_i = w_1^{(i_1)} = \Phi_{i1} \quad (25)$$

At the kth step where k≥2, $1 \le i \le M$, $i \ne i_1$, $i \ne i_{k-1}$, Compute

$$\beta_{jk}^{(i)} = w_j^T \Phi_i \Big/ (w_j^T w_j), 1 \le j \le k \quad (26)$$

$$w_k^{(i)} = \phi_i - \sum_{j=1}^{k-1} \beta_{jk} w_j \quad (27)$$

$$\gamma_{kq}^{(i)} = (W_k^{(i)})^T y_q \Big/ (W_k^{(i)})^T W_k^{(i)}, 1 \le q \le m \quad (28)$$

$$[err]_k^{(i)} = (\sum_{q=1}^{m} (\gamma_{kq}^{(i)})^2)(W_k^{(i)})^T W_k^{(i)} \Big/ trace(y^t y) \quad (29)$$

Find

$$[err]_k = [err]_k^{(ik)} = \max([err]_k^{(i)}, 1 \le i \le m, i \ne i_1, ..i \ne i_{k-1}) \quad (30)$$

And select

$$w_k = w_k^{(ik)} = \Phi ik - \sum_{j=1}^{k-1} \beta_{jk} w_j \quad (31)$$

Where

$$\beta_{jk} = \beta_{jk}^{ik}, 1 \le j < k \quad (32)$$

At the *Ms*th step when

$$1 - \sum_{j=1}^{Ms} [err]_j < \rho \quad (33)$$

the procedure is terminated.

3. DIRECT TYPICAL-POINT SELECTION

This approach is based on the principle of selecting typical practical operating points. Firstly, let us look over all the operation points in one industrial process, secondly, from all these points we select the typical points which basically represent the evolution feature of the process. We take these typical points as the RBF network centers.

When selecting the typical points, we must take into consideration of the corner points, or the edge points where the process changes greatly. Hence the variation feature of process can be described with small number of typical points and the model constructed will be simple and accurate.

4. SIMULATION OF RBF NETWORK MODEL

High power DC furnace for graphitization is an important plant for production of graphitic products. The product quality is mostly dependent on the control of the furnace temperature, but up to now no accurate model of the process of the furnace has been constructed. By use of direct selection of typical points, we constructed the RBF network model for the DC graphitizing furnace and the accuracy of the model is satisfying. Fig. 2. gives the measured temperature curve of a certain furnace. Fig. 3. shows the simulated curve with the RBF model constructed by use of direct selection of typical points as centers. Fig. 4. shows the error between the measured and simulated curves. It can be seen that the RBF model constructed gives satisfying results.

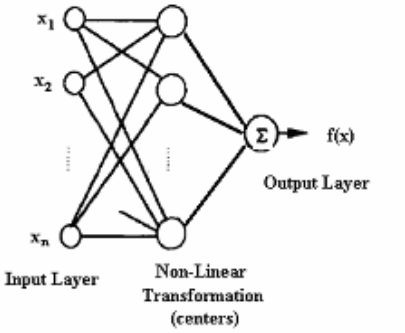


Fig. 1. Radial Basis Function Network

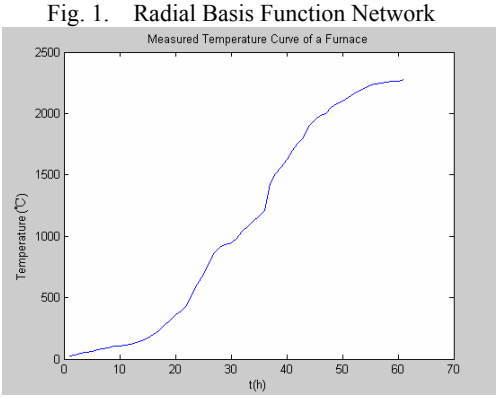


Fig. 2. Measured Temperature Curve

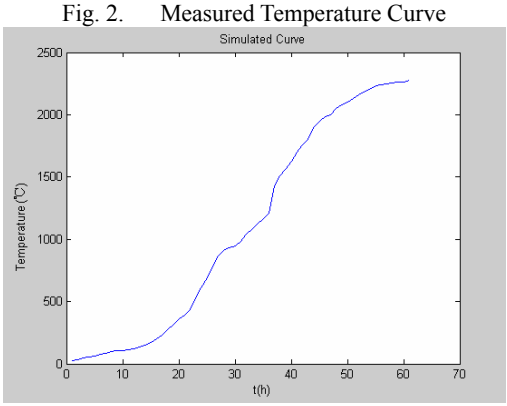


Fig. 3. Simulated Curve

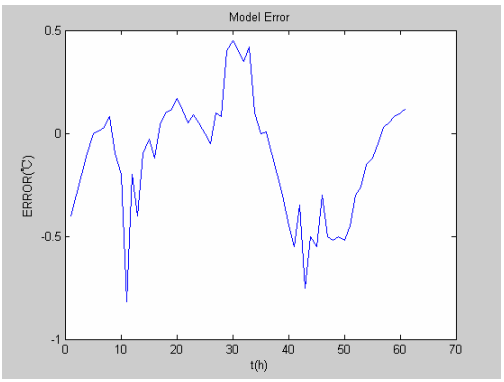


Fig. 4. Error

5. CONCLUSION

This paper discusses two approaches of selecting RBF network centers. In practical modeling, the two approaches are all effective and feasible. Practical experiments show that training the RBF net by MATLAB language, the algorithm converges within 2 min and an adequate model can be constructed.

REFERENCES

[1] S. Chen, S.A. Billings, et al, “Practical identification of NARMAX model using radial basis function”, Int.J.Control, 1990,52:1327-1350.

[2] S. Chen, S.A. Billings, et al, “Non-linear systems identification using radial basis functions, Int.J.Syst.Sci., 1990, 21(12):2513-2539.

[3] S. Chen, S.A. Billings, et al, “Reconstruction of binary signals using as adaptive radial basis function equalizer”, Signal Process, 1991, 22(1).

[4] S. Chen, S.A. Billings, “Neural networks for nonlinear dynamic system modeling and identification”, Int.J.Control, 1992, 56(2): 319-346.

[5] T. Chen, H. Chen, “Approximation capability to functions of several variables, nonlinear functionals, and operators by radial basis function neural networks,” IEEE. Trans”,Neural Networks, 1995, 6:904-910.

[6] N. B. Karayiannis, “Gradient descent learning of radial basis neural net-works,” in Proc. 1997 IEEE Int. Conf. Neural Networks, vol.3, Houston, TX, 1997, 1815-1820.

[7] “Learning algorithms for reformulated radial basis neural net-works,”in Proc.1998 Int. JointConf. NeuralNetworks, Anchorage, AK, 1998, 2230-2235.

[8] “New developments in the theory and training of reformulated radial basis neural networks,” in Proc. 2000 Int. Joint Conf. Neural Net-works, 2000, 3:614-619.

[9] N. B. Karayiannis, S. Behnke, “New radial basis neural networks and their application in a large-scale handwritten digit recognition problem,” in Recent Advances in Artificial Neural Networks: Design and Applica-tions, L. C. Jain and A. M. Fanelli, Eds. Boca Raton, FL: CRC, 2000, 39-94.

Optimization of SVM Kernels and Application to Down Category Recognition

Xiaoyan Yang[1], Hongwei Ge[2]
The School of Information Technology, Southern YangTze University
Wuxi, Jiangshu, People's Republic of China
[1]Email: yangyan1911@yahoo.com.cn
[2]Email: ghwjm@sina.com.cn

ABSTRACT

In recent years, the use of support vector machines (SVMs) on various classifications has been increasingly popular. However, the results of classification usually depend on the parameters of the model. These parameters are usually picked by experience, experimental compare, and large-scale search or cross-validation provided by software package. Our scheme to optimize SVM hyper-parameters is to minimize an empirical error estimate using a Quasi-Newton optimization method on the validation set. The method has been used successfully in our down category recognition system.

Keywords: Support Vector Machines (SVM), Hyper-parameters, Quasi-Newton Optimization Method, Empirical Error Estimate.

1. INTRODUCTION

In support vector machines, kernel parameters and trade-off parameter (C) are hyper-parameters which one needs to find their optimal values. A bad choice for these hyper-parameters will deteriorate the performance of SVM. For any classification task, picking the best values for these parameters is a non- trivial model selection problem that needs either an exhaustive search over the space of hyper-parameters or optimized procedures that explore only a finite subset of the possible values, or cross-validation function provided by software package. Until now, many practitioners select these parameters empirically by trying a finite number of values and keeping those that provides the least test error. Our interest is to optimize the SVM classifier for any kernel function used. More specifically, we seek an automatic method for model selection that optimizes the kernel profile dependently of the data. This allows better performance for the classifier and rigorous comparison among different kernel results as well.

We use an optimization method to achieve our goals. The optimization algorithm consists of two parts: one part is to use probabilistic outputs for the SVM by fitting a two-parameter sigmoid to the unbiased output of the classifier. Another part is to minimize an empirical error estimate using a Quasi-Newton optimization method [1,2]. The optimization algorithm has been used in our down recognition system based on SVM. And the experimental results show a better recognition performance.

2. INTRODUCTION OF SVM

This section gives a very brief introduction to SVM theory [3,4]. The basic idea of SVM is that we can transform the data points to another high dimensional space such that the data points will be linearly separable through a nonlinear map. The dot product in that high dimensional space is equivalent to a kernel function of the input space. So we need not be explicit about the transformation as long as we know that the kernel function is equivalent to the dot product of some other high dimensional space. Selecting different kernel function can produce different SVM, and then there is different optimal hyper-plane in a high dimension feature space. There are three commonly used kernel functions as follow:

(1) Polynomial kernel function:

$$K(x,x_i)=[(x\cdot x_i)+1]^q$$

(2) RBF kernel function:

$$K(x,x_i)=\exp(-\frac{|x-x_i|^2}{2\sigma^2})$$

(3) Sigmoid kernel function:

$$K(x,x_i)=\tanh(v(x\cdot x_i)+c)$$

3. OPTIMIZATION OF SVM KERNELS

Let us assume any kernel K depends on one or several parameters encoded into a vector $\theta=(\theta_1,\cdots,\theta_n)$. The trade-off parameter C is also encoded into the vector θ through modifying kernel K. Support vector machines consider a class of functions parameterized by α,b and θ as

$$f_{\alpha,b,\theta}=\sum_i \alpha_i y_i K_\theta(x,x_i)+b \quad (1)$$

Where the parameters α_i are found by maximizing a quadratic function and y_i represents the target of support vectors x_i.

Optimizing the SVM hyper-parameters is a model selection problem that needs adapting multiple parameter values at the same time. The parameters to tune are those that embed any kernel function as the σ parameter in an RBF kernel or the q parameter in a polynomial kernel. In addition, another parameter the optimization may consider is the trade-off parameter C, which may has a strong effect on the SVM behavior. Usually, minimizing an estimate of the generalization error does model selection.

We consider probability estimation in SVM as an estimate of the generalization error. Now, we use a sigmoid function given by

$$P(y=1|f)=\frac{1}{1+\exp(Af(x)+B)} \quad (2)$$

as probability estimation in SVM, where $f(x)$ is the SVM output and $y=1$ represents the target of the data example x. The two-parameter contracting function allows mapping the SVM output values to the corresponding posterior probability. The method is easy to implement and requires a nonlinear optimization of the couple of parameters (A,B) such a way the negative log-likelihood

$$\sum_i t_i \log(p_i)+(1-t_i)\log(1-p_i) \quad (3)$$

through the validation data points is minimized, where $p_i=\frac{1}{1+\exp(Af_i+B)}$ is the inferred posterior probability and $t_i=\frac{y_i+1}{2}$ is the binary target coding for the data point (x_i,y_i). During the experiments, we tried out the above-mentioned algorithm using a Newton method to adapt the sigmoid in Eq.2.

Let us consider now, a target-coding scheme for which $t_i=1$ if the input vector x_i belongs to class C_1 and $t_i=0$ if it belongs to class C_2. We can formulate the misclassification probability estimate of observing either target value for a given data example x_i as

$$E_i=P(y_i\neq z_i)=p_i^{1-t_i}(1-p_i)^{t_i} \quad (4)$$

where $z_i=sign(f_i), f_i=f(x_i)$ is the corresponding SVM output value and p_i is the estimated posterior probability. For a validation set of size N, the average estimate of the error could be written as:

$$E=\frac{1}{N}\sum_{i=1}^{N}E_i=\frac{1}{N}\sum_{i=1}^{N}p_i^{1-t_i}(1-p_i)^{t_i} \quad (5)$$

To optimize the SVM kernel parameters we should minimize the average estimate of the error E. We will use a Quasi-Newton optimization method to complete the optimization process because the method has faster convergence. So computing the gradient of the error is very important. It follows that we can approximate the gradient of the error as

$$\frac{\partial E}{\partial\theta}=\left.\frac{\partial E}{\partial\theta}\right|_{\alpha=\alpha_0}+\left.\frac{\partial E}{\partial\alpha}\right|_{\theta=\theta_0}\cdot\frac{\partial\alpha}{\partial\theta} \quad (6)$$

where $\alpha=(\alpha_1,\ldots,\alpha_k)$ represents the vector of multiplicative parameters and k equals to the number of support vectors. α_0 represents the support vector found by maximizing a quadratic function. θ_0 is the initial value of the hyper-parameters. To compute the gradient of the error we shall assume that the current vector of kernel parameters θ_0 to optimize is near a local minimum so the first derivative of the error w.r.t. θ at the minima vanishes. So to obtain global optimal value, we should give the initial value of hyper-parameters by experience in optimization method.

The components $\left.\frac{\partial E}{\partial\theta}\right|_{\alpha=\alpha_0}$ could be written as

$$\left.\frac{\partial E}{\partial\theta}\right|_{\alpha=\alpha_0}=\frac{1}{N}\sum_{i=1}^{N}\frac{\partial E_i(\alpha_0,\theta)}{\partial p_i}\frac{\partial p_i}{\partial f_i}\frac{\partial f_i(\alpha_0)}{\partial\theta} \quad (7)$$

where $\frac{\partial f_i(\alpha_0)}{\partial\theta}=\sum_{j=1}^{K}\alpha_j y_j\frac{\partial K_\theta(x_j,x_i)}{\partial\theta}$, x_j, y_j and α_j represent respectively the jth support vector, its label and its corresponding multiplier. Moreover, in equation (6), the gradient of the error w.r.t. the model parameters α is a k-dimensional vector given by

$$\left.\frac{\partial E}{\partial\alpha}\right|_{\theta=\theta_0}=\left(\frac{\partial}{\partial\alpha_1}E(\alpha,\theta_0),\ldots,\frac{\partial}{\partial\alpha_k}E(\alpha,\theta_0)\right) \quad (8)$$

the gradient $\frac{\partial}{\partial\alpha_j}E(\alpha,\theta_0)$ given by

$$\frac{\partial}{\partial\alpha_j}E(\alpha,\theta_0)=\frac{1}{N}\sum_{i=1}^{N}\frac{\partial}{\partial p_i}E_i(\alpha,\theta_0)\frac{\partial p_i}{\partial f_i}\frac{\partial f_i}{\partial\alpha_j} \quad (9)$$

where

$$\frac{\partial E_i}{\partial p_i}=-p_i^{1-t_i}t_i(1-p_i)^{t_i-1}+(1-t_i)(1-p_i)^{t_i}p_i^{-t_i}$$

$$\frac{\partial p_i}{\partial f_i}=-Ap_i^2\exp(Af_i+B) \text{ and}$$

$$\frac{\partial f_i}{\partial\alpha_i}=y_iK_\theta(x_j,x_i),\quad y_i=\pm1.$$

Once the error derivative w.r.t. multipliers vector α is computed, the next step consists of estimating the derivative of α w.r.t. θ. Notice that we may include the SVM bias b in the vector α as $\alpha=(\alpha_1,\ldots\alpha_k,b)$. It is shown that

$$\frac{\partial\alpha}{\partial\theta}=-H^{-1}\frac{\partial H}{\partial\theta}\alpha^T \quad (10)$$

where $H=\begin{pmatrix}K^Y & Y\\ Y^T & 0\end{pmatrix}$, $K_{ij}^Y=y_iy_jK(x_i,x_j)$. The

vector Y is the target vector corresponding to the support vectors set. Y^T is its transpose. H is a $(\#sv+1)\times(\#sv+1)$ matrix, $\#sv$ being the number of support vectors. Next, we shall refer to the matrix H as the kernel's Hessian. This matrix is different from the Quasi-Newton related Hessian. We shall refer to the latter as H'. We give herein the optimization algorithm using a quasi-Newton scheme:

1) Initialize θ to θ_0, and $\theta_0 = \arg\min_\theta E(\alpha_0, \theta)$.
2) Train the SVM with fixed θ_0 on training set.
3) Infer the sigmoid parameters A and B.
4) Estimate the probability of error on the validation set.
5) Calculate the gradient of that error $\frac{\partial}{\partial\theta}E(\alpha,\theta)$.
6) Calculate the Hessian H' over the kernel's parameter space.
7) Update using $\Delta\theta = -\lambda H' \frac{\partial}{\partial\theta}E(\alpha,\theta)$.

where λ is the amplitude of the step along the search direction and H' is a $n\times n$ matrix; n being the dimension of vector θ.

4. APPLICATION OF OPTIMIZATION METHOD IN DOWN CATEGORY RECOGNITION

Today, there are mainly two kinds of down products in the market, that is the goose down and the duck down. In the aspect of performance of keeping warm, the goose down is nearly one time higher than the duck down, and their prices also differ very greatly. So, some factories and companies replace the goose down with the duck down or mix goose down into duck down. Therefore, accurate recognition of goose down and duck down becomes more and more important.

So far, the automatic equipments on down category recognition have not been reported in the world. Down category recognition are often done by man with a microscope. Identifications of that kind demand a great deal of training and practice experiences, at the same time, many artificial factors in the process may result in different examination results with the same sample. Obviously, developing the automatic system on down category recognition has the realistic significance.

Here, we first analyze the structure of goose down and duck down under the microscope carefully, and get the picture of the down. Accordingly, after the image processing, the triangular nodes in the binary images of down is recognized by SVM, then the triangular nodes that have been recognized will be matched and the distance between the matched triangular nodes is calculated. Finally, we determine the down category according to the distance [5].

The structure analysis of the goose down and duck down: Goose down and duck down looks like a huge tree under the microscope (see Fig 1, the left image is goose down's image, the right one is duck down's image). We can find some nodes different in size in the tree, which are called triangular nodes, and the distance between two triangular nodes next to each other in the same branch is called nodes distance. Not all of the branches have the triangular nodes, and the triangular nodes often distribute in the end of the branches.

The shape and size of the triangular nodes of the goose down and duck down are different. The triangular nodes of goose down are isosceles triangles, and the nodes distance is longer. The triangular nodes of duck down are equilateral triangles, and the nodes distance is shorter. After the analysis of the structures of goose down and duck down, we find out that we can not recognize goose down and duck down according to the shapes of the triangular nodes, because they are quite similar. But the difference of the distance of the matched triangular nodes between goose down and duck down is remarkable. So we can recognize the down category according to the nodes distance.

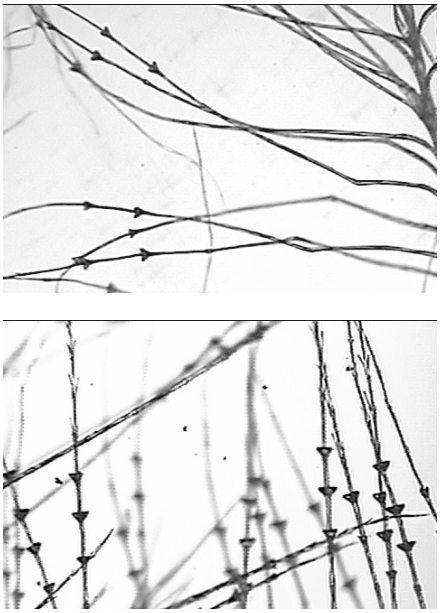


Fig .1. Original goose down image and original duck down image

Fig.2. Image containing a triangular node and image containing no triangular node

Table 1. The triangular nodes recognition results before and after optimization of Polynomial Kernel was done

Initial (q, C)	Misclassification rate	Tuned (q, C)	Misclassification rate
q=2,C=10	13.75%	q=3,C=0.1542	6.05%

Table 2. The triangular nodes recognition results before and after optimization of RBF Kernel was done

Initial (σ, C)	Misclassification rate	Tuned (σ, C)	Misclassification rate
σ=50,C=100	7.5%	σ=43.1448,C=69	5%

Table 3. The triangular nodes recognition results before and after optimization of Sigmoid Kernel was done

Initial (v, c, C)	Misclassification rate	Tuned (v, c, C)	Misclassification rate
v=-0.01,c=5,C=10	26.25%	v=-0.0165,c=5,C=15	15%

Application of optimization method: The recognition of triangular nodes is very important in the system. Here we use SVM to recognize the triangular nodes. To improve the recognition rate, we use the optimization method to adapt the hyper-parameters. In the rest of the paper we will describe the process and give the experimental results.

First, we shoot some images of goose down and duck down. After image processing, we get some binary images. Then from the binary images, we can extract the images of 16×16 (pixels), which contain triangular node in the center (see the left image in Fig 2). Using the same method, we extract 100 images of goose down which contain triangular nodes, 60 images of duck down which contain triangular nodes in the center, and 40 images which contain something that is not a triangular node but similar to it (see the right image in Fig 2). We take them as the training samples set of SVM. In the same way, we extract another 80 images of 16×16 (pixels) with the same scale of samples as validation samples set. Then we choose polynomial kernel function whose kernel parameters vector $\theta=(q,C)$ is initialized to $\theta_0=(2,10)$. The optimization process can be described as:

1) Get a model of SVM after training SVM on training samples set.
2) Get a set of output values $f_{i(i=1,\cdots 80)}$ when classify on validation samples set using the model of SVM.
3) Infer the sigmoid parameters A and B using Newton optimization method based the above obtained f_i, then calculate error estimation E.
4) Calculate the gradient of that error $\frac{\partial}{\partial\theta}E(\alpha,\theta)$, minimize an empirical error estimate E using Quasi-Newton optimization method on the validation set and get the optimal value of θ. Before calculate the gradient of the error $\frac{\partial}{\partial\theta}\mathrm{E}(\alpha,\theta)$, we add a very small constant to the diagonal elements of Hessian H, so H is guaranteed to be semi-definite positive.
5) Calculate respectively misclassification rate before and after optimization of kernel was done on validation samples set.

The triangular nodes recognition results before and after optimization of Polynomial Kernel was done is shown in Table 1.

If choose RBF kernel function and Sigmoid kernel function, we can get the triangular nodes recognition results before and after optimization of Kernel was done (See Table 2 and Table 3).

From the tables above we can find that misclassification rates of triangular nodes fall rapidly after optimization is done when the hyper-parameters values are far from the optimal values in fixed scale, as shown in Table 1 and Table 3. Besides, the performance of sigmoid kernel function is the worst in three kernels.

Now, extract 100 original images of 480×320 (pixels), including 50 duck down images and 50 gooses

Table 4. results compare before and after optimization of kernel was done

Kernel	polynomial	RBF
Initial parameters	q=2,C=10	σ=50,C=100
Node recognition rate	86.08%	91.94%
Down recognition rate	81%	87%
Tuned parameters	q=3,C=0.1542	σ=43.1448,C=69
Node recognition rate	92.12%	93.32%
Down recognition rate	87%	89%

down images. We can get the binary images after these images are processed. We take these binary images as the training set. There are 546 triangular nodes in these binary images. In the recognition stage, we scan the binary image with a window of 16×16 (pixels), by horizontal step 2 and vertical step 3, and recognize every window with SVM. We choose Polynomial Kernel Function and RBF Kernel Function in experiments. The result is shown in Table 4.

From Table 4 we can see that the recognition rate of down is improved through the optimization of SVM kernels with Quasi-Newton method.

5. CONCLUSION

Choosing the best hyper-parameters is very important in SVM classification task. It is easy when we process all kinds of classification task if we can automatically choose hyper-parameters. In this paper, we use a Quasi-Newton method to adapt the hyper-parameters of SVM kernels and improve our down category recognition system. The optimization method can automatically tune multiple parameters for a SVM, so this approach makes it possible for the use of highly complex and tunable kernels in practical application. In the future, we will use Quasi-Newton method and retrain method to improve our recognition system.

6. REFERENCES

[1] N.E. Ayat, M. Cheriet, et al, Empirical Error Based Optimization of SVM Kernels：Application to Digit Image Recognition. Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition, 2002

[2] O. Chapelle, V. Vapnik, et al, Choosing Multiple Parameters for Support Vector Machines. Machine Learning Journa, 2002 l, 46(1): 131-159

[3] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines. Cambridge University Press, 2000.

[4] V. Vapnik, Statistical Learning Theory. Wiley, New York, 1998.

[5] G. Hongwei, A down Category Recognition System Based on SVM. Computer Engineering and Applications, 2005, 41(27): 221-224

Moving Vehicle License Plate Recognition Based on Parallel Integrated Multi-classifier

Tongneng He, Nengneng Liu
Department of Information Engineering, Zhejiang University of Technology, Hangzhou, 310014, China
Email: htn@zjut.edu.cn
Defu Chen
School of Electronic, Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, 200030, China
Email: defu.ren.ren@163.com

ABSTRACT

In the process of moving vehicle license plate recognition, the sequential images get degraded generally, due to the movement and influence of lacking illumination and smut. It is difficult to recognize the number of vehicle license plate correctly by those common classifiers. A new integrated classifier proposed here based on multi-feature and multi-structure is applied to recognize moving vehicles license plates. The recognition correction rate of license plate is exceeded 94% in experiments. Statistical recognition results show, compared with common single-structured classifiers, this new integrated scheme is able to improve the recognition performance of single character and license plate number.

Keywords: Character Recognition, License Plate Recognition, Neural Network, Classifier Integration.

1. INTRODUCTION

Vehicle license plate automatic recognition (LPAR) is one of the most important research issue in the ITS (Intelligent Transport Systems), and it has high application value on traffic management, public security, and so on. A full LPAR system mainly consists of vehicle image capture, license plate location, character extraction and character recognition, but character recognition is the sticking point. The rate of correct recognition of license plate finally decides the system's performance.

At present, most methods of character recognition are based on template matching or neural network. A recognition method of multi-expert fusion integrated with multi-method is becoming the hot spot in research and application [1,2,3]. ZHANG Pei-feng et al. [1] test with 100 thousand samples in printed characters recognition system by adopting a multilevel serial structure connecting neural network classifier with statistic classifier in series and got 96.5% recognition rate. CAO Di-min et al. [2] connect BP neural network classifier with linear perception classifier in series, applying to license plate character recognition, gained 100% recognition rate with letter and 96% recognition rate with letter & number. Aiming at number recognition, F Kirma et al. [3] employed the approach combined with statistic classifier and structural classifier, which is realized with improved square decision function and decision tree separately, for testing the combined scheme which has two parallel units and four serial units and finally gained the recognition rate of 95.8%. In theory, the performance of integrated classifier is better than single-structured classifier [4]. Generic classifier usually designed simply which can't obtain perfect performance or designed for non-real time applications, so they are not suitable for moving vehicle license plate recognition. On the bases of research on how to convert a problem of designing a single classifier which has good performance into designing multi-classifiers which has better performance and on how to convert higher dimension space's partition into lower, we present a new method of classifier integration aiming at moving vehicle license plate recognition and got excellent result in experiments.

2. PRE-PROCESSION OF MOVING VEHICLE LICENSE PLATE RECOGNITION

The typical structure of moving vehicle LPAR system is consisting of moving vehicle detection and image capture, license plate location, character extraction and character recognition. When the vehicle entered surveillance area, the system detects the moving vehicle and then captures the video or static image (contained license plate) of the vehicle through high-speed camera or charge couple device (CCD) camera. Vehicle can be detected by hardware or software. Hardware detection method mainly use sensor which emplaced at vehicle detection spot for detection; Soft detection method detects the movement of vehicle from video sequences directly based on the principle of motion detection. Soft detection need no complex hardware requirements and easy for operating, so we adopt it.

For convenience of following processing, we removed the background of vehicle image. Fig. 1. gives the captured images of the vehicle from video sequences and it's background removed.

Locating the plate area follows the step of image capture. License plate location is an important step in LAPR system. An integrated method use different feature of the license plate can get better effect. The edge and region characters are most remarkable signals of the object for human vision. They have important meaning for image analysis and image recognition. This paper adopt the method which combined with character edge information and gray-scale projection: first improve the image's quality with homomorphic filter and strengthen the image's edge with Sobel operator, then get the texture image through horizontal difference operator. Using the rich texture in plate region, we can find several license plate suspect regions through detecting $\{0,1\cdots1,0\}$ and $\{1,0\cdots0,1\}$ sequences in texture image. Because the width-height ratio of license plate is fixed and existing "peak-valley-peak" status in license plate's projection image, we can remove pseudo license plate region and locate the true plate region through vertical projection.

License plate location effect is shown as Fig. 2.

a) foreground image of moving vehicle 1 b) foreground image of moving vehicle 1 c) background removed and gray-scale transform

Fig. 1. Image capture and background expunction

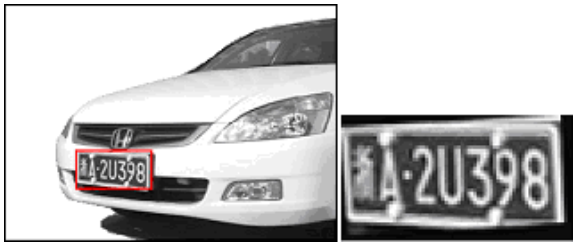


Fig. 2. License plate location

Fig. 3. Character extraction

The next step is character segmenting. Theoretically, projection of character string have local minimum values between characters or clearances between characters, so the correct segment place is near the minimum value and satisfying arrangement format, size limits and other intrinsic conditions. But because of motion blur, lacking illumination and mud, the character easy to be part and conglutination after binarization[5], simple method can't get satisfied segmentation effect. The paper introduced region split-combining algorithm [6] for single character segmentation, Fig. 3. Shows the segmentation effect after incline correction.

3. LICENSE PLATE CHARACTER RECOGNITION BASED ON INTEGRATED CLASSIFIER

Traditional pattern recognition system commonly use single sample feature and specific classifier, but it is empirically confirmed that this will get poor classification effect if the class number is big or input samples have noise. For different classifier, the algorithm and pattern feature is different, so the interface of classes is not same, at last each classifier will put up different performance. People also found that different feature and different classifier has complementary effect in classification performance, so the recognition method with integrated classifier which has multi-feature and multi-classifier can get better performance.

3.1 Multi-classifier and its integration

Traditional classifier integration structure mainly has two forms: parallel connection and serial connection. In parallel structure, each classifier is designed separately and its decision-making results are given independently, then all the decision-makings are combined and by certain integration rule the integration system's decision-making result is obtained. According to the different output of each classifier which participate in the integration, multi-classifier integration can be divided into three levels: (1) abstract layer: The output of classifier is the class number or the set of class numbers of the pattern; (2) sort layer: Classifier sorts the output class number and the first one is the preferred selection; (3) measurement layer: Output of classifier is the possibility of each class which the input pattern belongs to. These three different level integrations need different requirements, which are improved gradually for single classifier, and use more information of each classifier too. So the possibility of good recognition effect also increases in turn. A simple method for the integration of multi-classifier is voting, such as majority voting rule and consistent rule. Parallel system which adopt voting scheme can be described as followings [7]:

Suppose pattern space P is composed of M non-intersectant subclasses: $P = C_1 \cap C_2 \cap \cdots \cap C_M$, for C_i, $i = \Omega = \{1,2,\cdots,M\}$, is a class of one pattern. For each pattern in P, the task of classifier E is determining the class which it belong to and express it with a class number $j \in \Omega \cup \{M+1\}$, where $j \in \Omega$, which shows that X is one of the class of M; $j = M+1$ shows the classifier refuse to recognize X. No matter what principle classifier E worked with, it can be attributed to one of the three-classifier levels mentioned before. For given N separated classifier $(E_k, k = 1,2,\cdots N)$, if don't consider its inner structure, a Boolean function can be defined aiming at classification results $E_k(x) = i$,

$$T_k(x \in C_i) = \begin{cases} 1 & E_k(x) = i, \quad i \in \Omega \\ 0 & other \end{cases}, \tag{1}$$

And the general forms of voting is shown as Eq (2), where $\lambda \cdot N$ is voting threshold, N is the total number of classifiers.

$$E(x) = \begin{cases} i & \sum_{k=1}^{N} T_k(x \in C_i) > \lambda \cdot N, 0 < \lambda < 1 \\ M+1 & other \end{cases} \tag{2}$$

For serial structure, each sub-system is correlative, the fore-layer provides classification information for the back-layer, the back-layer decides with external input and fore-layer's output, and the whole rejection recognition is the output of the last layer. Serial system can be described as followings:

Considering a serial system, which has n layers, the recognition rate of i layer is R_i, error rate is W_i, rejection rate is P_i, and classification feature of each step is independent. $R_i + W_i + P_i = 1, \quad R_i, W_i P_i \in [0,1]$. System error rate is

$$W_{SYS} = \sum_{i=1}^{n} \left[W_i \prod_{k=1,\, k\neq 0}^{i-1} P_k \right], \tag{3}$$

system rejection rate is

$$P_{SYS} = \prod_{k=1}^{n} P_k = \prod_{k=1}^{n} (1 - R_k - W_k), \tag{4}$$

and system recognition rate is

$$R_{SYS} = \sum_{i=1}^{n} \left[R_i \prod_{k=1,\, k\neq 0}^{i-1} P_k \right] = \sum_{i=1}^{n} \left[R_i \prod_{k=1,\, k\neq 0}^{i-1} (1 - R_k - W_k) \right], \tag{5}$$

From the above we can see, the final recognition rate will go to 100% along with the adding of member classifier even the correct recognition rate of some sub-class is low. Certainly, it's impossible for each W_i to be zero. At the same time, serial structure is more complex than parallel structure, and especial when the feature and classifier can be selected in a small range the efficiency of serial connection is not effective and affects factual performance. So each sub-classifier in serial structure usually is constructed depending on prior knowledge.

Obviously, parallel connection classifier or serial connection classifier excel single-structure classifier in performance and each has its strong point. Experiment proved that, on base of the basic fusion of parallel and serial, the mixed integration classifier could use more information of each sub-classifier and get better classification performance.

3.2 Scheme of license plate recognition based on integrated classifier

In LPAR system, recognition rate of character directly reflect system's performance, so character recognition is especially important. At present, character of Chinese license plate can be divided into three classes: Chinese characters, English characters and numbers. The first character on the plate is Chinese character and it's the name of province or municipality for short. The second character is an A~Z English character in capitalization, the third and fourth is English character or number, and the last three characters is number between 0~9. It is obvious that LPAR is a classification problem about big pattern sets. And because of some disturbance such as noise, lacking illumination and license plate distortion exists, people often recognize by error. This paper designed a new serial-parallel integrated classifier and Fig. 4. described its working flow.

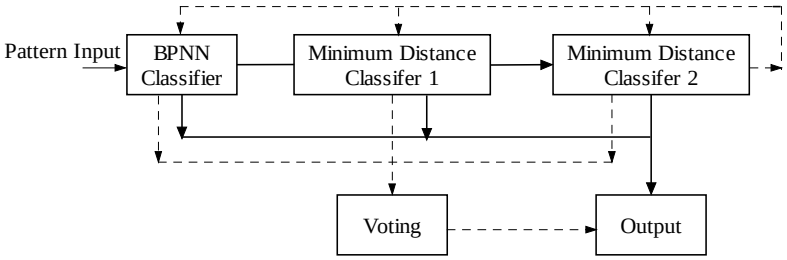


Fig. 4. Flow chart of character recognition with integrated classifier

The classifier adopt the recognition strategy that work in series (show in real line) first then in parallel connection (show in dashed). The work principle is described as follows: extract the features of the character after normalization ((1) meshing feature; (2) inner couture feature; (3) external couture feature; (4) stroke direction consistency feature). First input feature (1) into BP neural network and get the recognition result and confidence through BP network classifier, if the confidence exceed threshold T1, output the result directly; or else send the character into template matching classifier 1 (base on (2)) which based on min-distance. The reason of this arrangement is that Neural Network classifier has nicer fault-tolerance and can classify seriously disturbed characters cursorily, then make fine classification through template matching. According to threshold T2 the current classifier decide if output the result or send the character to next layer classifier, i.e. minimum distance classifier 2(base on (3) and (4)). If the character is still refused after the three classifiers, which connected in serial, the character will be sent to parallel layer and

recognized by each individual classifier. At last, combined with the confidence got in serial process, the integrator adjudicates the parallel output through voting method and outputs the final result.

3.3 Experimental result

Base on the integrated classifier, we test 200 license plates and 1100 single license plate image characters which consist of 200 Chinese characters (province names for short), 400 English characters, 300 alphabetic- numerical mixed characters (200 English characters and 100 numbers) and 200 numbers which got from various road condition, and got the following results as showing in Table 1.

Statistical results of recognition show that integrated classifier has more high performance in recognition of single character or license plate number compared with the sub-classifier.

Table 1. Statistics of recognition results with different classifier

Classifier / Character	Serial classifier		Parallel classifier		Integrated classifier	
	Error or reject number	Recognition rate	Error or reject number	Recognition Rate	Error or reject number	Recognition rate
Chinese character	9	95.5%	10	95%	7	96.5%
Letter	6	98.5%	8	98%	5	98.9%
Letter、Number	7	97.7%	9	97%	6	98%
Number	0	100%	0	100%	0	100%
Number of plate	-	92.8%	-	91.5%	-	94.3%

4. CONCLUSION

A good LPAR system should do relatively excellent at each step, but the result of character recognition finally reflects the system's performance. During the process of moving vehicle license plate recognition the sequential images are usually degraded. It is often resulted from movement and the influence of lacking illumination and smut. So it is difficult to recognize the number of vehicles license plate correctly with common classifiers. We first preprocessed the video sequence image before character recognition, aiming at the shortage of character classifier in existence, a new integrated classifier based on multi-feature and multi-structure is proposed in the paper for recognition of moving vehicles license plates. By employing the new classifier, the rate of correct recognition of license plate exceeded 94% in experiments. Statistical result of recognition show the new scheme has more high performance for single character or license plate number compared with the sub-classifier.

5. REFERENCE

[1] P.F. Zhang, W. Li, et al, on the hybrid mode of pattern recognition system. Information and Control [J]. 1997, 26(2): 121-128.

[2] D.M. Cao, G.R. Xuan, Neural network applied in character recognition of vehicle license plate. Microcomputer Applications [J], 2000, 16(6): 5-9.

[3] F. Kirma, M. Shridhar, Handwritten numerical recognition based on multiple algorithms. Pattern Recognition [J], 1991, 24(10): 969-983.

[4] J. Kittler, M. Hatef, et al. On combining classifiers. IEEE Trans, on PAMI, 1998, 20(3): 226-239.

[5] Y. Zhang, License plate binarization algorithm based on analysis of the spatial distribution and maximum variance between clusters [J]. Journal of Zhejiang University (Engineering Science), 2001, 35(3): 272-276.

[6] D. M. Cao, G. R. Xuan, Abruption and mergence in character area segmentation of license plate. Computer Engineering [J], 2000, 26(10): 18-21.

[7] S. P. Abhijit, B. M. Robert, Pattern recognition with neural network in C++[M], Translate by Xu Yong, Jing Tao. Beijing: Publishing House of Electronics Industry, 1999.

A Study of the Genetic-based Partner Evaluation and Selection Models*

WuLing Ren[1], JianTing Jin[2], ZhiYu Yang[3]
Computer Science and Information Engineering, Zhejiang Gongshang University
Hangzhou, Zhejiang, China
[1]Email:rwl@mail.zjgsu.edu.cn
[2]Email:jjtking@163.com
[3]Email:yzy635029@163.com

ABSTRACT
To significantly improve the performance of partner selection model, a genetic-based partner selection model with fuzzy reasoning is presented. In the model, the genetic algorithm is used to determine an optimal set of parameters due to its good behaviors on global search and quick convergence in early stage. Using heuristic information a genetic-based new fuzzy reasoning mechanism is used to adjust the eigenvalues of the partner's features due to its self-adaptability. In the genetic algorithm, a new genetic operation called changing colonies is defined, in which premature colony is replaced by another colony in the search space in order to get the effect of equivalent mandatory mutation for all genes in a specific period This model has been applied to the CRM system in Sanhua Hold Group, a middle-size manufacturing enterprise, and raised the regular delivery rate of raw materials supplied by the selected partners more than 20%.

Key words: Partner selection, Partner evaluation, Feature weight vector, Genetic algorithm, Fuzzy reasoning.

1. INTRODUCTION

The partner selection problem stated in this paper is to select the best partner according to the evaluation of the partners from multiple aspects that is made by definite evaluation method. It is commonly applicable, for example, it can be used to select better supplier of materials, which would significantly reduce cost and enhance the product quality. How to evaluate the partners synthetically is the key factor in partner selection problem. It's easy to estimate monomial aspects of partner, but it's difficult to evaluate partner synthetically. Obviously, to gain the monomial value and then to evaluate partner synthetically according to these monomial values is very important in decision field.

In this paper, a genetic-based partner evaluation and selection model (G-PESM) is presented. First, the parameters of the model are determined according to the experienced data and heuristic information, and the information database of partners is established. Then, each partner is evaluated according to the given constrained conditions. With the evaluation, appropriate partners can be selected, and the cooperated tasks can be assigned to the selected partners. Finally, the behaviors of the partners with cooperated task can be evaluated in turn, which can help to adjust the information database of the partners. In this model, firstly, the genetic algorithm [1] is adopted, which is good at global search and converges quickly in early stage. Therefore, the parameters of model can be determined. Secondly, the genetic-based new fuzzy reasoning model (G-NFRM)[2] is employed, which can perform self-adaptive fuzzy reasoning quickly and accurately using heuristic information. The G-PESM model will be self-adaptive while G-NFRM is used to adjust the eigenvalues for partner description.

In this paper, the description of model is introduced in the section 2. In section 3, the methods of using genetic algorithm to determine the weight-vectors of partner feature are presented. Several adjustment methods for the eigenvalues of partners using the technique of fuzzy reasoning are introduced in the 4th section. The next section is the model application and a brief summarization of the model is introduced in the last section.

2. DESCRIPTION OF THE MODEL

The structure of GA-based partner evaluation and selection model is showed in Fig. 1.

In partner selection model, the general partner descriptions and the partner feature information are stored in the partner library. The feature information may be divided

* Foundation item: Project supported by the Science & Technology Plan of Zhejiang Province, China (2005C21026 and 2004C11009)

into quantitative information and qualitative information. The quantitative feature is quite definite and its eigenvalues may be expressed by numerical value, while the qualitative feature is often described fuzzily and may be denoted by linguistic value or fuzzy number. Quantitative feature and qualitative feature constitute the feature table of partner. Some normalization processing for features should be conducted in order to make them comparable, such as quantifying the qualitative feature (linguistic value) and making them be the fuzzy number 0~10, transforming quantitative feature into standard feature values whose average is 10, and transforming all these values into feature which is the most advantageous in "the maximization" (It is to say the bigger the value be, the more advantageous it be chosen).

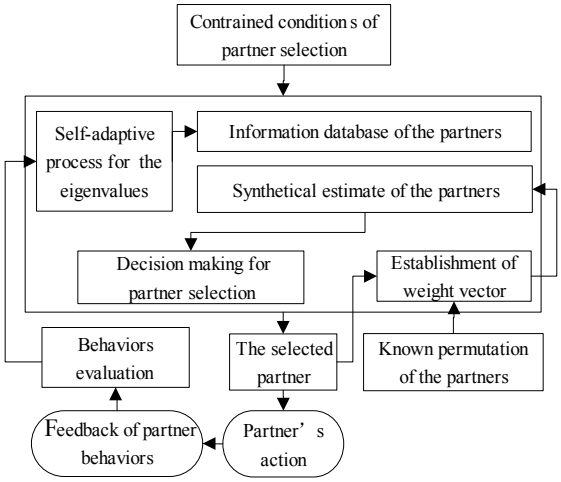


Fig. 1. A genetic-based partner selection model

The partner is chosen according to its synthetically evaluated value. Besides the feature table of the partner, the weight for each feature is needed. Weights of all feature constitute a vector $(\alpha_1, \alpha_2, \cdots, \alpha_n)$ which is called weight vector. Therefore, how to establish feature table of partner and weight vector objectively and accurately is a key factor of the model.

As small change of the weight vector will possibly cause significant difference among the order of partners, it is quite difficult to determine weight vector directly according to experiences. However, it is easier and much more accurate to permute the partners based on their work behaviors. Therefore, practicability and veracity of the model will be enhanced if weight vector can be determined automatically according to the given partner's permutation.

On the other hand, many features of the partner in the model should be determined according to experiences at the beginning, and whether the eigenvalues are accurate or not should be examined through the practice. Therefore, the evaluation on the behaviour of which the partner feed back should be made after the partner is selected. Then the original eigenvalues of model will be revised automatically according to the evaluation, and the features of partner will be reflected more accurately and more objectively by the eigenvalues.

3. ESTABLISHMENT OF THE WEIGHT VECTOR

The mathematical model to get the weight vector in G-PESM is a constrained nonlinear programming (NLP) problem. That is, given a partner feature table and a partner permutation, a weight vector should be solved to make the rank of partners in term of its synthetically evaluation value more accord with the given partner permutation:

$v_1, v_2, \ldots, v_m$ are m n-dimension vectors. v_i is denoted as: $(a_{i1}, a_{i2}, \ldots, a_{in})$, $a_{ij} \in [0, 10]$. According to the n-dimensional weight vector β mapping f_β is defined as: $v_i \rightarrow r_i$, $r_i \in R$, which maps the vector v_i to a real number r_i. Thereupon m vectors $\{v_i\}$ may get a rank according to the values of $f_\beta(v_i)$, denotes the result as Ψ_β: $v_{j1}, v_{j2}, \ldots, v_{jm}$.

Now, we given a permutation of m vectors X: $v_{i1}, v_{i2}, \ldots, v_{im}$, and want to solve out a weight vector β: $(\beta_1, \beta_2, \ldots, \beta_n)$, $\beta_i \in [0,1]$, $\sum \beta_j = 1$, making sure the objective function $\| X - \Psi_\beta \|$ can get the minimum value, where, $\| X - \Psi_\beta \|$ is the distance between the two vector sequences X and Ψ_β (Namely, the permutation difference among the vector sequences).

Obviously, it is quite difficult to work out the nonlinear programming problem in traditional mathematical methods, because the objective function and the obtained vector have no good mathematical nature. The mapping result of f_β is only used for comparing the level. So its numerical value is not needed with high accuracy. And the optimum weight is not sole, but maybe a consecutive series. Therefore, it is more appropriate to work out the weight vector by use of genetic algorithms (GA). After weight vector set gained, it can be further simplified (namely one of the weight vectors is chosen as the weight vector of model), for example, median is chosen as weight vector.

The individual of GA is defined as a chromosome with n-alleles. The value of each gene is ranged in [0, 10]. So, the individual of GA is defined as a string: $\alpha_1, \alpha_2, \cdots, \alpha_n$, $\alpha_i \in [0, 10]$, and it is corresponding to weight vector directly. Fitness function is defined as the difference among the partner permutations calculated by weight vector and the given partner permutations.

In the GA, population size is defined as 50, crossover probability is 0.9, and mutation probability is 0.1. Mutation operation is taken by modular arithmetic with modulus 10 and then plus 5.

In the above matter, the search space is Cartesian product of m consecutive interval [0, 10]. It is huge while the fitness function is a discrete function with limited values. Therefore GA is premature vulnerably. In this case, the problem can't be solved effectively only through adjustment of mutation operation and mutation probability, as showed in Fig.4. Therefore, an operation of changing colonies is defined, in which premature colony is replaced by another colony in the search space in order to get the effects of equivalent mandatory mutation for all genes in a specific period (in operation of changing colonies, all individual gene of colonies is mutated). The experiments show that operation

of changing colonies can enable GA is better converged to the optimum solution.

4. SELF-ADAPTIVE PROCESS OF EIGENVALUES

There are many ways to evaluate the behavior of partners. Correspondingly, there are various methods for the adjustments of eigenvalues.

4.1 Method of direct replacing

If feedback information can be expressed in the form of direct estimated values, they will replace the original eigenvalues of partner after the evaluation. The performance of partner in these features should be stable relatively and incidental diminutively in this adjustment method. Besides, estimated values of the partner should be gained expediently and accurately. Otherwise, it will cause larger error of eigenvalues and result in misplay of the selection of partners.

4.2 Method of fuzzy reasoning

For each feature, feedback information is in the form of estimated value. Assuming that the old eigenvalues of certain feature of the partner is described as v_t and its estimated value is defined as e_t, we hope to work out eigenvalue v_{t+1} through some reasoning or calculation after renewing. If the partner is selected, estimated value e_{t+1} for eigenvalue v_{t+1} will be gained once again. It will be done repeatedly to form a set of numerical 2-tuple: (v_1, e_1), (v_2, e_2), …, (v_t, e_t). For partners with relative stable performance, the next eigenvalues will be decided in light of its recent performance naturally. Therefore, we can get he next point v^* of point sequence with estimated value $\{e_i\}$ of point sequence in the corresponding mathematical model. If point sequence is convergent, the figure represented by point sequence will be a convergent oscillate curve. After several iterations, the fitted curve would be obtained by such curve fitting methods as the least squares fitting principles and the new eigenvalues can be calculated furthermore. However, there lack of the methods. On one hand, the evaluation information can't be used; on the other hand, it is difficult to assure the convergence of point sequence. Therefore, there is much difficulty in adopting method of fitting.

The basic method of fuzzy reasoning is put forward by Zade [3] initially while the new fuzzy reasoning model (NFRM) [4] is presented by Z. Cao and others. D. Park and other researchers introduced genetic operation [2,5] on the basis of NFRM to enable NFRM learn from cases and establish fuzzy relationship matrix and fitness function with good convergence and high accuracy.

Reasoning with the NFRM is a transformation between vectors virtually, in which the imported n-dimension vector e is transformed into the export m-dimension vector a through fuzzy relationship matrix $\boldsymbol{R}_{n\times m}$. The relationship between these two vectors can be expressed as: $\boldsymbol{a}=\boldsymbol{e}\circ\boldsymbol{R}$. The imported n-dimension vector e is the grade vector of membership of estimated value e. It is to say that e is corresponding to the grade vector of membership $\boldsymbol{e}=(e_1, e_2\ldots e_n)$ of each linguistic value (fuzzy subsets) of evaluation E (n is the number of linguistic value E). The export m-dimension vector $\boldsymbol{a}=(a_1, a_2\ldots a_m)$ is the grade vector of membership of adjustment value a (n is the number of linguistic value A). Thus adjustment value a of eigenvalues can be gained by taking method of the weighted average:

$$\mathrm{a}=\sum_j (f_j\times \mathrm{a}_j) \Big/ \sum_j \mathrm{a}_j \qquad (1)$$

Here, f_j is the center value of linguistic value with order j in membership function of adjustment value of eigenvalues.

Thus, after determining the membership function to evaluate eigenvalues, the membership function to adjust eigenvalues and fuzzy relationship matrix R, we list out evaluation value e to calculate out the adjustment value e and gain the new eigenvalues $v_{t+1}=a\times v_t$.

Obviously, whether the model is good or not depends greatly on the choice of model parameters. The model parameters formed in the light of experience in NFRM, such as the membership function to evaluate eigenvalues, the membership function to adjust eigenvalues and fuzzy relationship matrix R, is quite subjective and inaccurate. One of the improving methods is to determine function and reasoning rule on the base of the recent performance of partner in the feature. That is to say, membership function and fuzzy relationship matrix of NFRM is determined according to the recent k sequences (v_1, e_1), (v_2, e_2)… (v_k, e_k) and the new eigenvalues will be worked out. It is clear that the space of searching for model parameters is large and it is feasible to adopt GA. The veracity of model parameters can be expressed by mean-variance, which is defined as follows:

$$d^2=\sum_{i=2}^{k}(v_i-v_i')^2 \Big/ \sum_{i=2}^{k} v_i^2 \qquad (2)$$

Here, v_1, v_2 … v_k, e_1, e_2, …, e_k are the recently used eigenvalues and its corresponding estimated values. v_i' is eigenvalue educed by model after v_{i-1} and e_{i-1} is input.

According to the above requirement, chromosomes of GA should contain three parts: the fuzzy relationship matrix, the membership function to evaluate eigenvalues and the membership function to adjust eigenvalues. Each element of fuzzy relationship matrix is corresponding to a gene for real number between 0 and 1 directly. The triangle function can be used as fuzzy distribution function of membership function. Thus, membership function can be changed through adjustment of the peak position and hemline width of triangle. The three-part gene string together constitute to a chromosome. The fitness of chromosome is defined according to mean-variance of the representing model:

$$J=1\Big/(1+d^2) \qquad (3)$$

Algorithms and the set of parameters of GA is described in the previous subsection. When NFRM begins to work, it is not conformable to adjust fuzzy relationship matrix and

membership function by GA because of the few cases. It will be more appropriate to reason directly through experience data and introduce GA to adjust model parameters when the amount of examples number reach k, which is commonly above 6.

4.3 Method of likelihood reasoning

The method of fuzzy reasoning described in the above subsection 2 can be simplified if eigenvalues are evaluated by fuzzy linguistic value directly. As the evaluating linguistic value is given, reasoning can be executed using fuzzy reasoning rules directly and the adjustment value of eigenvalues can be derived without determining membership function of evaluation linguistic value. Therefore, chromosomes of GA don't need to contain membership function genes. Shortcoming of this method is that there is less linguistic value and the estimate can't be quite definite, which results in larger error.

4.4 Method of mobile center

The above two methods both require the behavior represented by partner should be provided with comparative stability. Otherwise, the causal link between evaluation of eigenvalues and adjustment of eigenvalues will be quite weak. As a result, it is not appropriate to work out new eigenvalues according to original eigenvalues and its evaluation, while it is quite fit to gain new eigenvalues by means of method of mobile center. The main idea of mobile center is to generate new eigenvalues on the base of several used eigenvalues and method of solving weighted mobile center. Each features can be considered as a unit, so can each partner.

5. APPLICATIONS OF G-PESM MODEL (RESULTS OF EXPERIMENT)

Several main feature parameters of material suppliers of Sanhua Hold Group, a middle-size manufacturing enterprise in Zhejiang Province, China and its eigenvalues after standardization are showed in Table 1.

Table 1. Partner features table of material suppliers

	Quantitative features				Qualitative features			
Supplier's Name	Material's Batch	Material's Price	Delivery Time	Service Time	Product Quality	Credit Standing	Service Quality	Product Image
A	16.4	9.9	3.9	12.5	6 ok	8 good	7 well	7 well
B	7.3	10	5.2	6.3	9 excellent	7 well	8 good	8 good
C	9.1	10.1	6.5	12.5	7 well	7 well	10 perfect	8 good
D	5.5	10.2	8.1	8.3	9 excellent	9 excellent	9 excellent	8 good
E	7.3	9.8	21.6	12.5	5 not bad	7 well	6 ok	5 not bad
F	7.3	9.9	16.2	6.3	6 ok	6 ok	6 ok	6 ok
G	9.1	10	12.9	18.8	7 well	5 not bad	7 well	8 good
H	18.2	10.2	5.8	15	9 excellent	8 good	7 well	9 excellent

For a given permutation of the supplies: H, D, B, C, G, A, F, E, the G-PESM gets the result with all individuals gained by random selection which can be seen in Fig. 2 and the result with designated individual（0.02, 0.2, 0, 0.05, 0.33, 0.2, 0.1, 0.1）in the first colony which can be seen in Fig.3. In Fig.2, there are twelve individuals achieving the maximal relative fitness 1 after the 735th generation. In Fig.3, there are twenty-four individuals achieving the maximal relative fitness 1 after the 16th generation. It is easy to see that heuristic information accelerate to speed up convergence pace of GA and distribution of solution is more concentrated while it will be broader with completely random search. Through working out sample median of weight vector corresponding to the individual in Fig.2 and Fig.3, the obtained weight vector is as follows: （0.0126, 0.1939, 0.0089, 0.0460, 0.2828, 0.1030, 0.0585, 0.2941）and(0..133, 0.1333, 0.0000, 0.0333, 0.2200, 0.1333, 0.0667, 0.4000).

It will be failing to reach the maximal relative fitness 1 with completely random search without operation of changing colonies as displayed in Fig.4.

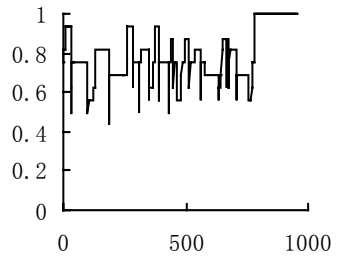


Fig. 2(a). Maximal relative fitness

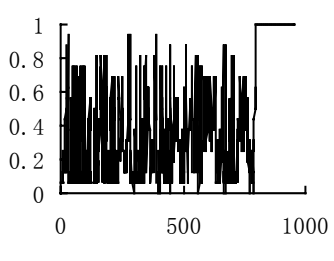


Fig. 2(b). Minimal relative fitness

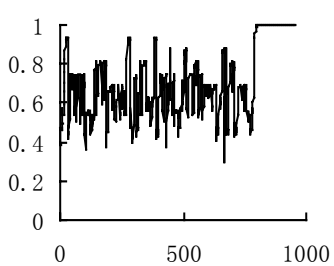


Fig. 2(c). Average relative fitness

Fig. 2. Fitness change on completely random search

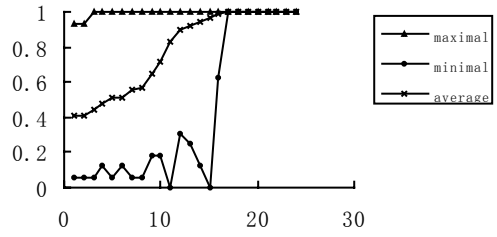


Fig. 3. Fitness changes with heuristic information

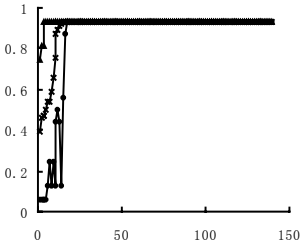


Fig.4. Fitness premature converged to 0.9325 without changing colonies operation

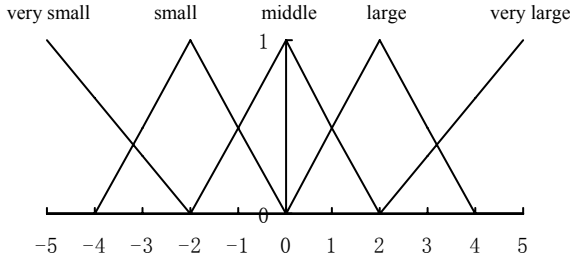


Fig. 5. The membership function of estimate E

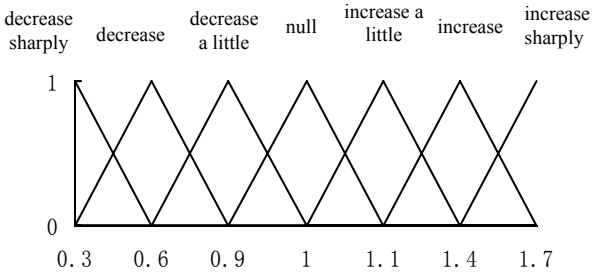


Fig .6. The membership function of adjustment A for eigenvalues

In the partner selection model, by practical experience, the fuzzy reasoning rules can be described by IF-THEN form as follows:

IF Estimate *E*=’very large’
THEN Adjustment *A*=’ decrease sharply’
IF Estimate *E*=’very large’
THEN Adjustment *A*=’ decrease’
IF Estimate *E*=’ large’
THEN Adjustment *A*=’ decrease a little’
IF Estimate *E*=’middle’
THEN Adjustment *A*=’null’
IF Estimate *E*=’small’
THEN Adjustment *A*=’increase a little’
IF Estimate E=’very small’
THEN Adjustment *A*=’ increase’
IF Estimate *E*=’very small’
THEN Adjustment *A*=’ increase sharply’

Here, *A* and *E* are two different linguistic variable, *A* has seven linguistic values: decrease sharply, decrease, decrease a little, null (keep unchanged), increase a little, increase, and increase sharply. While *E* has five linguistic values: very larger, large, middle, small, and very small. Given the range of estimated value *e* is: [-5, 5], the range of adjustment values of eigenvalues is: [0.3, 1.7], i.e. $v_{t+1} = a \times v_t$, then the corresponding membership functions of fuzzy sets *A* and *E* can be defined as the functions shown in Fig.5 and Fig.6. In accordance to the fuzzy reasoning rules, the corresponding fuzzy relationship matrix as shown in table 2:

Table 2. The fuzzy relationship matrix created by experience

Adjustment *A* / Estimate *E*	decrease sharply	decrease	decrease slightly	null	increase slightly	increase	increase sharply
Very large	1.000	1.000	0.000	0.000	0.000	0.000	0.000
large	0.000	0.000	1.000	0.000	0.000	0.000	0.000
middle	0.000	0.000	0.000	1.000	0.000	0.000	0.000
small	0.000	0.000	0.000	0.000	1.000	0.000	0.000
Very small	0.000	0.000	0.000	0.000	0.000	1.000	1.000

Suppose the eigenvalues sequence is: 9.00, 3.90, 3.17, 2.87, 3.14, 2.98, the corresponding adjustment sequence is:

5, 2, 3, -1, 0.5, 0, its running results as shown in Fig.7 and Fig.8. Fig.7 is based on totally random generated chromosome while Fig.8 is the results by designated a chromosome according to the experience information in Table 2, Fig.5, and Fig.6. Table 3 is the fuzzy relationship matrix, and then conducts the updated eigenvalue is: 3.04.

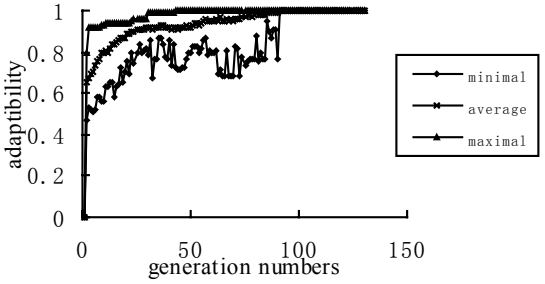


Fig. 7. The generation numbers and adaptability without any heuristic information

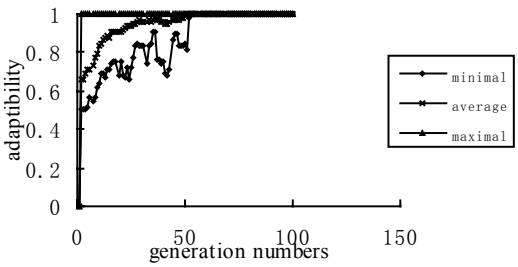


Fig. 8. The generation numbers and adaptability with heuristic information

Table 3. The fuzzy relationship matrix created by GA

Estimate *E* \ Adjustment *A*	decrease sharply	decrease	decrease slightly	null	increase slightly	increase	increase sharply
Very large	1.000	0.824	0.000	0.000	0.000	0.000	0.000
large	0.000	0.000	1.000	0.042	0.000	0.000	0.000
middle	0.000	0.000	0.000	0.227	0.000	0.000	0.000
small	0.000	0.000	0.000	0.000	1.000	0.000	0.000
Very small	0.000	0.000	0.000	0.000	0.000	0.992	1.000

6. SUMMARY

The partner evaluation and selection model presented introduces genetic algorithms and fuzzy reasoning model so that the accuracy and objectivity of model is significantly improved. The model has been applied to the evaluation and selection of raw materials and spare parts suppliers in Sanhua dynamic alliance management system. After the achieves of suppliers, supplied materials and the model of evaluation and selection of suppliers were setup, the selection of material suppliers by experience has shifted into automatic selection according to data model. It receives great achievement in economic benefit and social benefit. There are many details to be improved in the model, such as how to express fuzzy condition and how to enable model accept fuzzy condition when fuzzy condition is chosen by model, by what criteria to choose method of estimating the partner behaviors. Furthermore, the complexity in time and space should be considered in model when partner spaces are larger and partner features are complex.

7. REFERENCES

[1] E. David, “Genetic Algorithms in Search, Optimization, and Machine Learning”, Addison-Wesley, 1989.
[2] P. Daihee, “Genetic-based new fuzzy reasoning models with application to fuzzy control”, IEEE Trans. on Systems, Man, and Cybernetics, 1994, 24(1): 39-47.
[3] L.A. Zedeh, “Outline of a new approach to the analysis of complex systems and decision processes”, IEEE Trans. on SMC, 1973, 3(1): 28-44.
[4] Z.Cao, “A new model of fuzzy reasoning”, Fuzzy Sets and Syst., 1990, Vol.36: 311-325.
[5] P.G. David, “A genetic system for learning models of consumer choice”, Proc. of the 2nd International Conf. on Genetic Algorithms, 1987, 217-223.

Fuzzy PD Controller Application in Microturbine Start-up*

Shijie Yan, Yongkui Man, Huaguang Zhang
School of Information Science and Engineering, Northeastern University
Shenyang, Liaoning 11004, China
Email: yanshijie@mail.neu.edu.cn

ABSTRACT

The research of microturbine generator (MTG) system has drawn a lot of attention these days because they have the advantages of low cost, less pollution and safe distributed power sources. The MTG start-up is the most important stage in a MTG control process. Its model is non-linear, so conventional control approach adopted before cannot achieve desirable result. Based on the research of a 100kW microturbine system, a novel fuzzy PD control algorithm is proposed by combining the nonlinear tracking differentiator and the conventional fuzzy controller in MTG. The main improvement of the controller is that it has high robustness against the noise and is easy for engineering implementation. The algorithm was tested on real prototype system. Experimental results are shown in the paper to verify the effectiveness of the proposed strategy.

Keywords: microturbine, MTG, fuzzy PD control, start-up, Tracking differentiator.

1. INTRODUCTION

Microturbine generator (MTG) is recently getting a lot of attention, because of low cost, clean environment and safe distributed power source, and compared with traditional power generation. The operating period of a MTG consists of the stages such as, soft start-up, igniting, warming up, accelerating, no-loading, loading, shutdown, etc. So, the various studies about the control in each stage from start-up to loading need to be stable and reliable because the control procedure is very complicated. Among the stages, Start-up is the most important one in the whole operating procedure. Meanwhile, the MTG model parameters are variable and uncertain, start-up procedure is the most challenging problems for control applications to develop new control algorithms. A novel fuzzy control strategy was proposed to optimize and enhance performance in a MTG start-up process with double-sided PWM converters. The background of this paper is based on the research of 100KW MTG prototype system, which is named NWDR100 [1]. An extensive simulation results are given to support the control algorithms. Then, a laboratory drive system was designed and implemented to verify the effectiveness of the proposed strategy.

2. NWDR100 MTG SYSTEM OPERATION

NWDR100 MTG is composed of compressor, combustor (combustion chamber), microturbine, fuel system and drive system, and the configuration is shown in Fig.1. The operation of the system is briefly explained as follows:
First, the MTG compressor compresses outside air under atmospheric pressure into microturbine. After the compressed air is heated by using exhaust gas in a recuperator, it is mixed with fuel gas, such as the natural gas. Then, the mixed gas is burned in a combustor and expanded to drive turbine. Finally kinetic energy is generated by the turbine. The turbine output power is utilized for both generator and compressor.

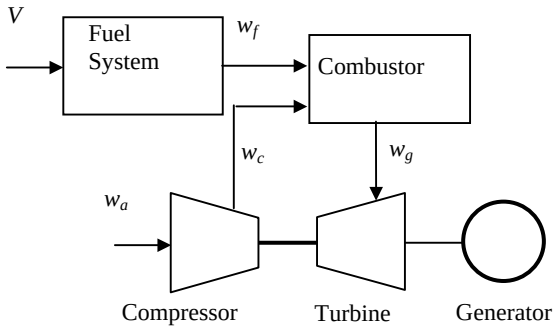


Fig. 1. Microturbine Configuration

The kinetic energy is converted into electric energy with a generator. Since the output frequency of the generator is higher than 1kHz, the output voltage is converted into DC voltage with a rectifier, and then converted into the commercial frequency (50Hz) ac voltage with an inverter.

3. MTG START-UP PROCESS ANALYSIS

Generally, the process that the MTG operates from igniting and warming up to idle state is defined as the start-up stage. Although it is only one of the stages, its control task is the most important in a MTG control process, which ensures a successful running. In [2], a heavy-duty or medium heavy duty gas turbine system model can be made, which can be linearized by expressing them in a total differential equation, but it is impossible to use linear mathematic model to control MTG, because the engine of MTG is the permanent magnet synchronous generator and the MTG power transformation system is a high frequency converter. So we must carefully analyze the whole process of start-up and find out a new control strategy based on expert experience.

During the stage of start-up, the objective of control is that the MTG inlet temperature T_3^* should be within the limit, the accelerating time to reach idle state should be controlled as expected, and the whole operating state should be stable in any disturbance. After igniting, the MTG is required to accelerate and reach idle speed, according to MTG characteristics. Abnormal overshoot in temperature T_3^*, temperature suspended, combustor explosion burn or excessive vibration are not allowed. The high start-up performance is determined by the start-up characteristic curve of the MTG. A reasonable start-up curve as the target of start control is the most important. Following two cases should be taken into account for the choice of the curve [3]:
(1) Fuel feeding during accelerating is the basic control requirements for maintaining the MTG to operate within its safe margin during start-up and output the rated value in the shortest period of time. The start-up curve should not be in the surge domain. The closer to surge line the start-up curve

*Supported by National '863' Plan (2002AA503020) of China.

is, the faster the start-up is. So the start-up curve should be along with the surge line and the time to reach the idle speed is the shortest in many various curve. But, if the temperature of MTG inlet air is high, or acceleration time is too short, it will reduce the service life-span of hot components.

(2) To attain the design life-span, the accumulative stress effects in the highly stressed parts of the machine, i.e. the high temperature and acceleration rate must carefully be controlled. The life-span is mainly determined by the cycles of thermal stress experienced during the start up. Thermal stresses occur due to temperature gradients through the MTG metal. Important variables in the thermal stress equations include exhaust gas temperature, combustion chamber pressure and rates of change of these. Thermal stress equations are developed for a particular plant based upon a finite element grid to represent the inner temperatures of MTG**.** Therefore, to limit to the thermal stress, we should choose optimal start-up curve, that is a gas flow curve, which is obtained by calculating and experimenting in start–up, shown in Fig.2.

During the development period, we chose some expert experience in thermodynamics, vibration, lubrication, fuel control, control system and mechanics, and calculate and analyze the data of designing and experiment. So the start-up curve was obtained. In a MTG, the fuel mass flow signal is independent, its curve is also a start-up curve. Although we got Fig.2 start-up curve, it is difficult to control in start-up stage, as mathematical model of MTG is a highly complex and non-linear system with a strong coupling among the various parameters. Especially in the development stage, a precise mathematical model may not be available. According to requirement of MTG start-up, a novel fuzzy PD control strategy is proposed.

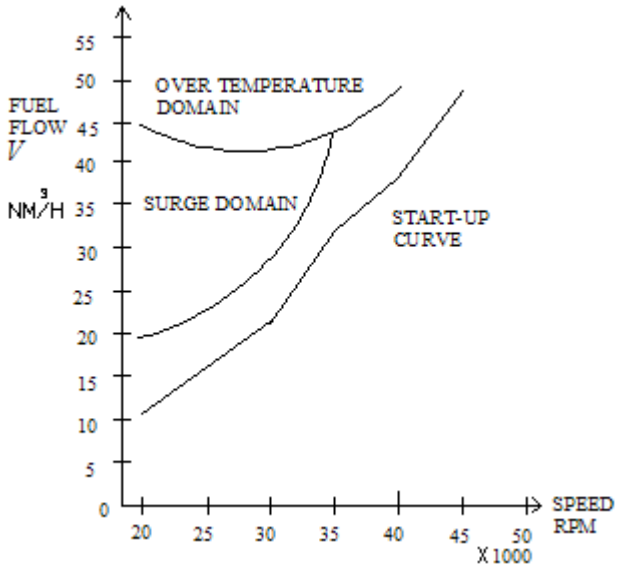


Fig. 2. Reference start-up curve of MTG

4. FUZZY PD CONTROL STRATEGY

PROPORTIONAL-DERIVATIVE (PD) controller in various combinations has been widely used for industrial processes due to their simplicity and effectiveness for linear systems and it can provide high sensitivity and tend to increase the stability of the overall feedback control system. But PD controller usually cannot be used for nonlinear systems. In general, Fuzzy controllers are suitable for many nontraditionally modeled industrial processes that cannot be precisely described by mathematical formulations and they have significant un-modeled effects and uncertainties [4,8,9]. In this section, we designed a fuzzy controller of the implicit type of 'if-then' rules along with their development mechanisms. We first established a framework for the fuzzy PD controller design, and then described the main steps: fuzzification, control-rule base set up and defuzzification.

4.1. Framework of fuzzy PD controller

The conventional PD control law is described by

$$u(t) = K_p e(t) + K_d \frac{de(t)}{dt} \qquad (1)$$

where K_p and K_d are the proportional and derivative gains of the controller, respectively, and $e(t)$ is the error signal defined by $e(t) = x(t) - y(t)$, with $x(t)$ being the reference signal and $y(t)$ the system output. In generally, $\frac{de(t)}{dt}$ can be obtained from difference of $e(k)$ and $e(k-1)$, but this method has a large inaccuracy and it is difficult to improve the performance of the system. HAN [5] presents nonlinear tracking differentiator (TD) to provide high quality differential signal of the measured signal with noise. Based on the TD, a fuzzy PD controller is developed by combining the TD and a conventional fuzzy controller in MTG, Fig.3. The main improvement of the proposed fuzzy PD controller is that it has high robustness against the noise and is easy for engineering implementation.

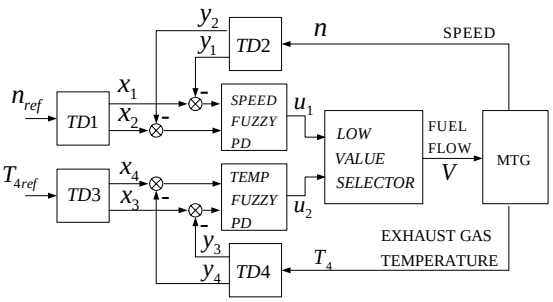


Fig. 3. The framework of fuzzy PD control system

In Fig.3, nomenclature as the follow:

n_{ref} : Reference speed of MTG.

n : Actual speed.

T_{4ref} : Reference maximum exhaust temperature.

T_4 : Actual exhaust temperature.

x_1 : TD1 output signal which tracks reference speed.

x_2 : TD1 output signal, which is the differential signal.

y_1 : TD2 output signal which tracks actual speed n.

y_2 : TD2 output signal, which is the differential signal.

x_3 : TD3 output signal which tracks reference exhaust temperature.

x_4 : TD3 output signal, which is the differential signal.

y_3 : TD4 output signal which tracks actual exhaust temperature.

y_4 : TD4 output signal, which is the differential signal.

u_1 : Speed fuzzy PD controller output signal.

u_2 : Temperature fuzzy PD controller output.

V : Fuel mass flow signal, which is output from low value selector.

Where TD1, TD2, TD3, TD4 is identical structure and expression. SPEED FUZZY PD and TEMP FUZZY PD have identical structure, but the linguistical meaning is not the same. Low value selector is unit that can select minimum value in all input value to output a fuel mass flow signal.

4.2. Nonlinear tracking differentiator

Nonlinear tracking differentiator can be expressed as a second order system [5]:

For the input signal of reference speed n_{ref} , TD1 can produce high quality output signal x_1 , x_2 , and make x_1 track n_{ref} , x_2 track $\frac{dn_{ref}}{dt}$. TD2, TD3, TD4 have the same as TD1 in structure and parameter.

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -Rsat(a,\delta) \end{cases} \quad (2)$$

$$sat(a,\delta) = \begin{cases} sign(a), & |a| > \delta \\ \frac{a}{\delta}, & |a| \le \delta, \delta > 0 \end{cases} \quad (3)$$

$$a = x_1 - r + \frac{hx_2^2}{2R} \quad (4)$$

$$h = \begin{cases} 1, & x_2 > 0 \\ -1, & x_2 < 0 \end{cases} \quad (5)$$

where r is reference input, it may be a kind of variable in n_{ref} , n , T_{4ref} , T_4 . $sat(\bullet)$ is a nonlinear saturation function that protect from oscillation in zero point. $sign(\bullet)$ is a normal sign function. R is amplitude of nonlinear function, δ is transient speed factor. In general, when the tracking signals include 1% noise, R value should be selected between 2.5-50. When the noise signal is further enhanced, the tracking is not affected, but the differential signal will be enlarged and oscillated, so R should be as possible as small in assuring the tracking input signal. δ effects differential signal. According experiment, we define $\delta = 0.00005R$ in Fig. 3 . Design procedure and proving of nonlinear tracking differentiator in detail are described in [5]. In the MTG system, the feedback signal n and T_4 were frequently disturbed, so we used TD to approximate them in front fuzzy controller.

4.3. Fuzzy PD controller design

In this paper, the start-up curve of the MTG is presented in Fig.2. SPEED FUZZY PD controller (SFPC) and TEMP FUZZY PD controller have the same in structure, except for control-rule base. We will only describe the design of the SFPC. The TEMP FUZZY PD is not discussed in this paper. From Fig. 3, the SFPC has two inputs, that are error signal e and its change rate c . They can be expressed as [6]

$$\begin{cases} e = K_e(x_1 - y_1) \\ c = K_c(x_2 - y_2) \end{cases} \quad (6)$$

where k_e is the error gain, k_c is the error change rate gain, which can regulate SFPC precision.

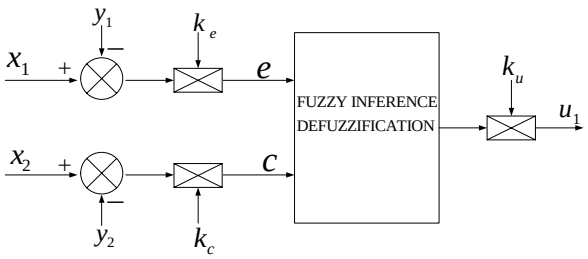


Fig. 4. Block diagram of fuzzy controller

The block diagram of SFPC is shown in Fig.4. Fuzzy control of the MTG speed is a proportional- derivative (PD) control, where the gain factors are nonlinear and are adaptively controlled to get robust response. The speed error e and the error change rate c are converted into per-unit signals by the constant scale factors respectively, and then processed through fuzzy control to generate the signal, as indicated. It is then multiplied by the constant scale factor and the SFPC output signal u_1 is generated. Finally, after LOW VALUE SELECTOR, the fuel flow command signal *V* is created.

In MTG system, we have N control rules. A typical l th control rule can be written as the follows:

R^l : IF e^l is A_l AND c^l is B_l , THEN u_1^l is U_{1l}

where $l = 1,2,...N$, A is fuzzy sets of e , B is fuzzy sets of c .

Membership function of the fuzzy sets utilizes bell function, and it is the expression as follows [6]:

$$\mu(x) = \exp\left(-\frac{|x - x_0|^2}{\sigma^2}\right) \quad (7)$$

In MTG system, the center of gravity defuzzification method is employed. The SFPC output:

$$u_1 = \frac{\sum_{l=1}^{N} w_l U_{1l}}{\sum_{l=1}^{N} w_l} \quad (8)$$

where $w_l = \mu_{A_l} \bullet \mu_{B_l}$.

Table 1. THE CONTRL RULES

e / c	NVL	NL	NM	NS	ZE	PS	PM	PL	PVL
NVL	NVL	NVL	NVL	NVL	NVL	NL	NM	NS	ZE
NL	NL	NL	NL	NL	NL	NM	NS	ZE	PS
NM	NL	NL	NL	NL	NM	NS	ZE	PS	PM
NS	NL	NL	NL	NM	NS	ZE	PS	PM	PL
ZE	NL	NL	NM	NS	ZE	PS	PM	PL	PL
PS	NL	NM	NS	ZE	PS	PM	PL	PL	PL
PM	NM	NS	ZE	PS	PM	PL	PL	PL	PL
PL	NS	ZE	PS	PM	PL	PL	PL	PL	PL
PVL	ZE	PS	PM	PL	PVL	PVL	PVL	PVL	PVL

In the system, nine levels are defined based on (6), (7), (8). The corresponding rule matrix is given in Table 1.
Where PS, NS, PM, NM, PL, NL, PVL, NVL, ZE is linguistic value.

5. EXPERIMENTAL RESULT AND ANALYSIS

In the gas turbine system, the open-loop control approach is generally adopted throughout the start-up, but the uncertain factors exists, during acceleration from igniting to idle, so we applied fuzzy PD control in the MTG system. Through analysis and experiment, we have achieved satisfactory result by utilizing the closed-loop control in start-up process. We use data sets from [7] calculation, which comprised a series of critical operating points. As mentioned above, a complete 100kW MTG prototype system was designed to validate the control laws. The control hardware is based on Texas Instruments TMS320LF2407A-type digital signal processor (DSP) board. The start-up control system is only one subsystem of the whole MTG system. The control parameters are transmitted between this subsystem and the main controller using the CAN field bus. Since SFPC is designed for interpolation purpose at start-up stage, the smaller the error is, the faster the start processing is. By experiment and operator's experience, we should choose the best characteristic parameter in all σ of membership functions, thus the shape of membership functions was adjusted, and then the fuel-speed curve $V-n$ is changed steeply. The dynamic characteristic experimental waveform of the MTG start-up process was shown in Fig.5 - Fig.6 at $\sigma = 0.15$, $\sigma = 0.3$, respectively, While $k_e = 49.3$, $k_c = 5.3$, $k_u = 0.8$.

As shown in the figure, the law of regulation for parameter σ is that the bigger one can eliminate noise signal and the smaller one may enable to approximate the sample data with a given accuracy. The speed is increased in start-up, when the parameter σ in (7) is increased, and the start-up time is shortened. A better value of σ can make the SFPC have the optimum tracking response and small error.

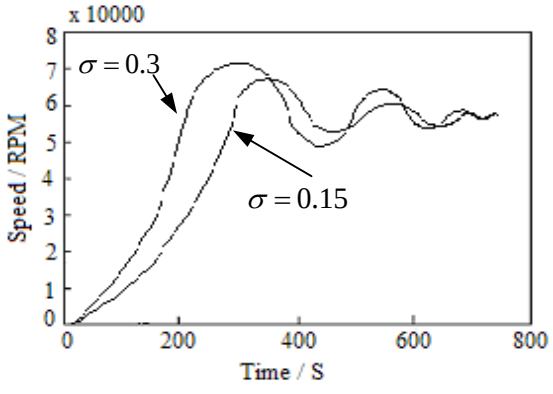


Fig. 5. The curve of rotate speed of the MTG

6. CONCLUSIONS

In this paper, we have described the start-up process of MTG, design principle and tracking performance of the fuzzy proportional-derivative (PD) controller. After establishing a framework for the design, we have given a derivation for the fuzzification, rule base setup, and defuzzification in the design of the fuzzy PD controller. Through tuning the parameter σ of membership functions based on the operating conditions of the MTG, the optimum start-up process of the gas turbine can be achieved, compared with the conventional approach (PID, open-loop). After start-up successfully, another control strategy will be applied to control the MTG running in the whole project, that it is not discussed in this paper. The experimental results of MTG have demonstrated the advantage of the fuzzy PD controller, particularly when the system to be controlled is nonlinear with strong disturbance.

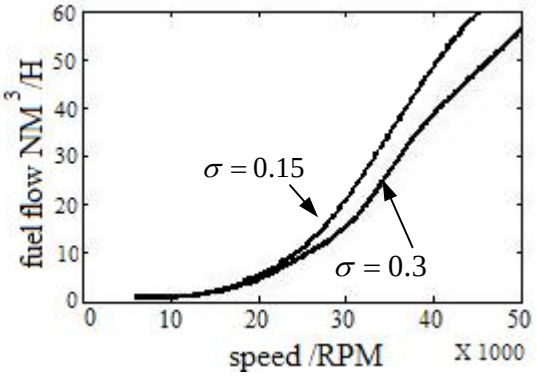


Fig. 6. The curve of start-up process of the MTG

7. REFERENCES

[1] X. X. Wang, C. Y. Zhang, et al, "Design of Microturbine Rotate Speed Control System", Control Engineering, 12(1), 2005,pp.73-76.
[2] W.W. Hung, "Dynamic Turbine Simulation of Gas-turbine Generating Unit", IEE proceeding, vol. 138. (4), 1991,pp.342-350.
[3] J. Wang, G.M. Cui, et al, "Study on Control of Micro Gas Turbine ", Electromechanical Equipment, 4,2004, pp.1-6.
[4] H.A. Malki, H.D. Li, et al, "New Design and Stability Analysis of Fuzzy Proportional-Derivative Control Systems", Fuzzy System, 2(4), 1994,pp. 245-254,
[5] J. Q. Han, W. N. Wang, "Nonlinear Tracking-differentiator", System Science & Mathematics, 14 (2), 1994,pp. 177-183.
[6] L. X. Wang, A Course in Fuzzy System & Control, Prentice-Hall, Inc, 1997.
[7] Z. P. Feng, X. Y. Liu, et al, "Design Analysis of Thermal System of Microturbine", Journal of Engineering Thermophysics, 23(6), 2002,pp.17-20.
[8] M. Lown, E. Swidenbank, et al, "Adaptive fuzzy logic control of a turbine generator system", IEEE trans. Energy Conversion, 12 (4), 1997, pp. 394-399.
[9] J. Talaq, F. A. Basri, "Adaptive fuzzy gain scheduling for load frequency control", IEEE Trans. Power Syst., 14 (1), 1999,pp.145-150.

The application of Particle Swarm Optimization in Training Support Vector Machines

Zhiguo Chen, Yan Shan
School of Information Technology, Southern Yangtze University, Wuxi, Jiangsu, 214122,China
Email: chenzg888@yahoo.com.cn

ABSTRACT

Large number of example vectors brings difficulties for quadratic programming problem with support vector machines, traditional methods may be impossible. The intelligent search technologies, such as genetic algorithms and particle swarm optimization algorithm, can give a similar solve of problems in less time. Particle Swarm Optimization is better than genetic algorithms in convergence and stability of the overall. According to the characters of swarm intelligence and constrained optimization, we propose a method to solve a linearly constrained quadratic optimization problem in training support vector machines with PSO (for short). Testify PSO has determinate applied value in the field of support vector machines, and it is a new way for quadratic programming problem with a large number of example vectors.

Keywords: support vector machines, Quantum-behaved Particle Swarm Optimization, quadratic programming problem.

1. INTORDUCTION

Support Vector Machines (SVMs) is a new machine learning method in the statistical learning theory,. The method is based on structural risk minimization principle theoretical foundation, SVMs have proved their worth - there has been a remarkable growth in both the theory and practice of these learning machines in the last decade. Because of excellent learning performance, SVMs have become a hot research sector in the area of learning machines.

Training a SVM requires solving a linearly constrained quadratic optimization problem. This problem often involves a matrix with an extremely large number of entries, which make quadratic programming problem unsuitable. This directly affects the practical applications of support vector (QP) problem, there is a α_i for each vector in the training set. The value b is a threshold. The decision surface is defined by support vectors that are correspond to the subset of nonzero α_i .

Training the SVM consists of finding the values of α_i . By defining a Hessian matrix Q such that $H_{ij} = y_i y_j k(x_i, x_j)$, training can be expressed as a dual QP problem of solving

$$\text{Min} \quad W(\alpha) = \frac{1}{2}\sum_{i,j}\alpha_i\alpha_j y_i y_j k(x_i, x_j) - \sum_{i=1}^{l}\alpha_i \qquad (2)$$

$$\text{S.T.} \quad \sum_{i=1}^{l} y_i\alpha_i = 0 \qquad (3)$$

$$0 \le \alpha_i \le C \qquad i = 1, 2, 3, \cdots, l \qquad (4)$$

So training a SVM is to solve a linearly constrained quadratic optimization problem.

machines. Several methods have been used to decompose the problem, of which many require numeric packages for solving the smaller sub problems.

Particle Swarm Optimization (PSO) is intuitive and easy-to-implement algorithm from the traditional swarm intelligent algorithm, and is introduced as a way of training a SVM.A PSO is shown to be ideal in optimizing the SVM problem, and find a new way for solving quadratic programming problem in large number of example vectors.

In this paper, the SVM algorithm is described, and the main methodologies for training SVMs is explained as well. PSO is used to solve a SVM's quadratic programming problem. Experimental results on Iris data set show the performances of PSO.

2. SUPPORT VECTOR MACHINES

A SVM is a learning machine for two-class classification problem, which is derived from data classification problem. Given a set of l N-dimensional training vectors $\{ x_i, y_i \}$, $i = 1, 2, 3, \cdots, l$, $\{x_i\} \in Rn$, $y \in \{+1, -1\}$. The algorithm learns a surface that can separate between the two classes. This surface is not created in input space, but rather in a very high-dimensional feature space. The resulting model is accomplished by the use of kernel functions. The kernel function k gives a measure of similarity between a pattern x. The decision boundary that needs to be constructed is of the form

$$f(x_i) = \sum_{i=1}^{l} y_i \alpha^*_i k(x \cdot x_i) + b^* \qquad (1)$$

Where the sign of $f(x)$ determines the class of x. The α_i is Lagrange multipliers from a primal quadratic programming

3. SVM Training Methods

The QP problem is to find the minimum of a constrained bowl-shaped objective function. Due to the definition of the kernel function, the matrix H is positive and is a convex QP problem, so every local solution is also a global solution. The Karush-Kuhn-Tucher conditions give conditions determining whether the constrained maximum has been found. Solving QP problem is to have every vector satisfy the KKT conditions. A solution α of the QP problem is an optimal solution if the following relations holding for each α_i :

$$\begin{aligned} \alpha_i = 0 &\Leftrightarrow y_i f(x_i) \ge 1 \\ 0 < \alpha_i < C &\Leftrightarrow y_i f(x_i) = 1 \\ \alpha_i = C &\Leftrightarrow y_i f(x_i) \le 1 \end{aligned} \qquad (5)$$

In the real application, the number of example vectors is usually large, and solving the QP problem is very difficult. The matrix *H* has a dimension equal to the number of training examples. For large learning tasks, off-the-shelf optimization packages and techniques for general quadratic programming become intractable in the memory and time requirements. So for fast convergence and small memory requirements, a number of other approaches even on large problems have been invented.

Chunking

Chunking was invented by V.Vapnik in [10]. The chunking algorithm is based on the fact that the nonsupport vectors play no role on the SVM decision boundary. So delete the zero Lagrange multiples of the training set of examples, the SVM solution will be the same.

In this algorithm, break the large QP problem down into a number of smaller problems. Optimize the Lagrangian on a subset of data at each iteration, the initial values of the Lagrangian is values at the last iteration. During the optimization, the set of nonzero α_i are retained, and all zero ones are discarded. At every subsequent step, the training examples are vectors of all nonzero α_i, and some of the α_i that violates the KKT conditions. The procedure is iterated until all examples met the KKT conditions, and the margin is maximized. At last chunking has optimized all the nonzero α_i. The overall QP problem is solved.

Decomposition

Decomposition was introduced by E.Osuna in [4][5], is similar to chunking. The large QP problem is broken down into a series of smaller subproblems, and a numeric QP optimizer solves each of these problems. Different form Chunking, decomposition optimizes a subproblem of α_i, and the others are remained the size of subproblems is fixed. Each optimization of subproblem must keep the subproblem meet KKT condition. One vector that violates the KKT condition is added and one removed from the subproblem at each iteration. In this fashion the sequence of QP subproblems will asymptotically converge. Solving each subproblem requires a numeric quadratic optimizer.

SMO

The sequential minimal optimization (SMO) is the extreme case of decomposition. The size of working size is 2; two α_i is chosen to be optimized. The two α_i are directly optimized; the speed of the subproblem is improved. The SMO greatly accelerates the speed of convergence.

SVMlight

The algorithm of SVMlight is an extension of decomposition algorithms. Split Lagrange multiplies of examples into working set and non-working set. Pick α_i for the QP subproblem such that the α_i form the steepest possible direction of ascent on the objective function, where the number of nonzero elements in the direction is equal to the size of the QP subproblem.

4. PARTICLE SWARM OPTIMIZATION

PSO [3] is a method of evolutionary computation. The algorithm derived from the study in acts of birds preying. The intelligence and efficiency lies in the cooperation of an entire flock. The algorithm is introduced in [7].

Different from traditional optimization methods, in PSO the direct fitness information instead of function derivatives or related knowledge is used to guide. Particles collaborate as a swarm to reach a collective goal. Each particle has memory of the best solution that it has found, called its personal best. A particle's traversal of the search space is influenced by its personal best and the best solution found by a neighborhood of particles. So the information is shared. Each particle profits from the discoveries and previous experience of other particles during the exploration and search for lower objective function values. Each particle has personal best, so the whole population has the best performance-the global best-called *gbest*, which connects all the particles in the population to one another.

In the swarm i is the index of a particle. And each particle has a velocity v_i with which the particle fly through the *n*-dimensional searches space R^n. The particle adjusts its position according to its own previous best solution pi and the previous best solution of the whole swarm.

In PSO, the velocity updates are calculated as a linear combination of position x_i and velocity vectors. The particles dynamic evolution and move according to the following equations

$$v_{i,d}(t+1) = wv_{i,d}(t) + c_1 r_{1,d}(t)(pbest_{i,d}(t) - x_{i,d}(t)) + c_2 r_{2,d}(t)(gbest(t) - x_{i,d}(t)) \quad (6)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (7)$$

where *t* is the iteration. The r_1, r_2 ~*UNIF*(0,1) are random numbers between zero and one. These numbers are scaled by acceleration coefficients c_1 and c_2 which is usually between zero and two. And *w* is an inertia weight. The velocity vectors are clamp between upper and lower bounds to avoid moving out of the search space.

The PSO algorithm is summarized below:

1) Set the iteration number *t* to zero. Initialize the position of particles in the swarm, which must meets $Ax_i = b$. With *A* being a *m n* matrix and ***b*** a m-dimensional vector, *X={x | Ax=b}* defines the set of feasible solutions to the constrained problem, each point in P will be a feasible point.
2) Evaluate the objective function $f(x_i)$ value of each particle.
3) Compare the current performance of each particle to its personal best. If the current performance is better than *pbest*, then set the *pbest* to the current value.
4) Set the global best to the position of the particle with the best performance within the swarm.
5) Update the velocity vector for each particle according to equation (6).
6) Change the position of each particle according to equation (7)
7) Let *t:=t*+1
8) Go to step 2,and repeat until convergence.

The box constraints (3)~(4) are easily handled by initializing all particles x_i to lie inside the hypercube defined by the constraints, and restricting their movement to this hypercube. When a particle is moving outside the boundary of the hypercube, its velocity is scaled.

5. TRAINING THE SVM

Solving the SVM QP problem with PSO involves criteria for optimality, how to decompose the QP problem, and how to extend PSO to optimize the SVM subproblem.

The Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient for optimality. Because *H* is a semi-definite matrix and the QP problem is optimal problem of a convex matrix, so the solution that meets the KKT conditions is the best one.

Decomposing the QP problem involves select q sub-optimal values from all α_i as the working set called set B. The working set is optimized while the remaining ones called set N are kept constant. The decomposition algorithm is summered as follows:

1) While the KKT conditions for optimality are violated:
   a) Select q variables for the working set α_B. Keep the other *l-q* variables (set N) constant, and fix them at current value.
   b) Use PSO to optimize $W(\alpha_B)$ on B, get the solution of α_B^{K+1}.
   c) Return the optimized α_B^{K+1} from B to the original set of values.
2) $\alpha = \begin{pmatrix} \alpha_B^{K+1} \\ \alpha_N^{K+1} \end{pmatrix}$, Definite the function *f(x)* of the decision boundary.

The decomposition method above is to find the steepest feasible directions *d* on the object function W as defined in equation (2). Finding an approximation to d is equivalent to solving:

$$min \quad g(a^{(t)})'d \tag{8}$$

S.T.
$$y^T d = 0$$
$$d_i \geq 0 \quad if \ \alpha_i = 0 \tag{9}$$
$$d_i \leq 0 \quad if \ \alpha_i = C$$
$$d_i \in \{-1, 0, 1\}$$
$$|\{d_i : d_i \neq 0\}| = q$$

the function $g(a^{(t)})$ is the derivative of equating (2) at $a^{(t)}$, *d* is the solving direction. The first constraint equation requires the number of elements with sign matches between d_i and y_i must be equal to the number of elements with sign mismatches between d_i and y_i. The selection of *d* is to get the smallest value of the object function. The forth equation requires the feasible direction criteria. The last one requires the solution of the direction *d* includes *q* Lagrange multipliers. Sort the training vectors in increasing order according to $y_i g(a^{(t)})$. Assuming *q* to be even, a "forward pass" selects *q*/2 examples from the front of sorted list, and the "back pass" selects *q*/2 examples from the back. This method is given P.Lskov in [2].

α is separated into α_B and α_N

$$\alpha = \begin{pmatrix} \alpha_B \\ \alpha_N \end{pmatrix}, \quad Y = \begin{pmatrix} Y_B \\ Y_N \end{pmatrix}, \quad H = \begin{pmatrix} H_{BB} & H_{BN} \\ H_{NB} & H_{NN} \end{pmatrix}$$

H is symmetric, so $H_{BN} = H_{NB}$, and the object function is changed:

$$\min_{\alpha_B} W(\alpha_B) = 1/2\alpha_B^T H_{BB}\alpha_B - \alpha_B^T (e - H_{BN}\alpha_N) \tag{10}$$

S.T.
$$\alpha_B^T y_B + \alpha_N^T = 0$$
$$\alpha_B \geq 0 \tag{11}$$
$$C1 - \alpha_B \geq 0$$

Implementation of PSO

The initial solution is constructed in the following way in order to meet the constraint equations (3)~(4):

Randomly select c between 0 and C. And pick γ positive examples and γ negative examples, initialize corresponding α_i of 2 γ examples to *c*. The other α_i are zero. γ is small, less than both the number of positive examples and the number of negative examples. The initial solution is feasible.

In PSO, all particles are initialized such that $\alpha_B^T y_B + \alpha_N^T = 0$ is met. This is done as follows:

1) Initialize each particles in the swarm to the *q*-dimensional vector α_B.
2) Add a random *q*-dimensional vector *v* satisfying $y^T{}_B v = 0$ to each particle, under the condition that the particle will still lie in the hypercube $[0, C]^q$.

The initial swarm lies in the set of feasible solutions $P = \{p \mid Ap = -\alpha_N^T y_N\}$ in the initialization. Then optimize the *q*-dimensional vector α_B with PSO algorithm.

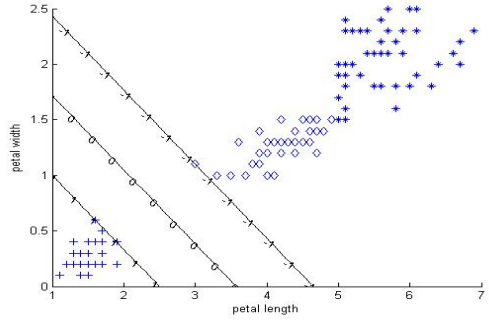
6. EXPERIMENTAL RESULTS

The SVM training algorithm presented in this paper was tested on the Iris data set. A data set with 150 random samples of flowers from the iris species *setosa*, *versicolor*, and *virginica* collected by Anderson. From each species there are 50 observations for sepal length, sepal width, petal length, and petal width in cm.

For training a SVM on the Iris dataset, the examples in *setosa* were used to represent the positive examples while the remaining examples defined the negative examples. Each example is the vector of 4-dimention.

The following parameters are defined as follows: The KKT conditions needed to be satisfied within an error threshold of 0.02 from the right hand side of equations. Optimization of the working set terminated when the KKT conditions on the working set were met within an error of 0.001.Let γ =10, a total of 20 initial support vectors were chosen to start the algorithm. The swarm size in each experiment was 20, and the iteration is 500. While the inertia weight *w* was set to change from 0.7 to 0.3, w=0.7**(Maxiter-t)/Maxiter*+0.3. The acceleration coefficients c_1 and c_2 were both set to 1.4.

The upper bound *C* was kept at 100.0.



Fig, 1. Separate Iris data sets to two classes by clearly identifying the petal length and petal width with PSO

The PSO training algorithm was written in matlab, and does not make use of caching and shrinking methods to optimize its speed. In order to show the possibility of training SVM with PSO, first show the condition of separating Iris data sets to two classes by clearly identifying the petal length and petal width with PSO, Fig. 1.. Do with a polynomial kernel. The working size is 4. We can see PSO algorithm can training SVM well. Experimental results show successful and accurate training on the Iris training.

Table 1. Influence of different working size on Iris dataset

Working set size	Time	SVs
4	104.41s	27
6	104.09s	29
8	105.82s	32
10	107.11s	24

ors of PSO is fewer.

From Table 1. , the time of PSO running is long. So the drawback from the experiment of PSO algorithm is its slow performance times. But with the development of optimization, there will be a good method to train SVM.

7. CONCLUSION

Experimental results show that particles group algorithms (PSO) have an excellent performance. PSO can be used for training support vector machines effectively, some PSO algorithms to optimize performance by QP problem reflects good performance. PSO algorithm is simple to operate, and does not require any background of numerical methods. Although speed of the algorithm is a practical drawback, suggest applying more convergent algorithm to training SVM So applications can be very effective in large-scale sample secondary planning, and provide a new way for the training of large samples.

8. REFERENCES

[1] U. Paquet, A.P. Engelbrecht, “Training Support Vector Machines with Particle Swarms”[C], 2003 IEEE, pp. 1593-1598.

[2] P. Laskov, “Feasible direction decomposition algorithms for training support vector machines”, in Machine Learning, Volume 46, N.Cristianini, C.Campbell, and Chris Burges, editors, 2002, pp. 315-349.

[3] J. Kennedy, R.C. Eberhart, “Particle swarm optimization”, in Proceedings of the IEEE International Conference on Neural Networks, IV, 1995, pp. 1942-1948.

[4] E. Osuna, R. Freund et al, “Support vector machines: Training and applications”, A.I.Memo AIM-1602, MIT A.I. Lab.1996.

[5] E. Osuna, R. Freund et al, “An improved training algorithm for support vector machines”, in Neural Networks for Signal Processing VII-Proceedings of the 1997 IEEE Workshop, IEEE, New York, 1997,pp.276-285.

[6] N.Y. Deng, Y.J. Tian, “ New Method In Data Mining--Support Vector Machine”, in Science Press, 2004.

[7] J.C. Zeng, Q. Jie, et al, “Particle Swarm Optimization Algorithm”, in Science Press, 2004.

[8] X.F. Song, D.Z. Chen et al, “Optimization Algorithm In Support Vector Machine”, in Computer Science, 2003Vol.30 No.3, pp.1029-1035.

[9] P. Wang, J.Q. Mao, “Training Algorithm And Application In Support Vector Machine,” in Information and Electronic Engineering, Vo1.3，No.4 Dec.2005, pp.309-314.

[10] V. Vapnik, “Estimation of Dependences Based on Empirical Data *[*in Russian*]*”, Nauka, Moscow, 1979. (English translation: Springer Verlag, New York, 1982.

Then training is done with a Radial kernel:

$k(x, x_i) = \exp(-\frac{\| x - x_i \|^2}{1.0^2})$. Table 2. gives the influence of different working set sizes. A working set of size q=6 gives the fastest convergence time and A working set of size q=4 has fewest support vectors. But with the size of working set growing, the time becomes longer. So selection of the working size also burdens the speed of the algorithm.

Training SVMs is to solve the QP problem, and make all Lagrange multipliers meet KKT conditions. PSO is easy to implement. The parameters are simple. The experimental results indicate that the convergence of PSO is fast. And comparing with previous studies the number of support vect_

Application of Fuzzy control in Internet congestion control system*

Hongyong Yang[1], Fuzeng Zhang[1], Fusheng Wang[2]
School of Computer Science and Technology[1], Information Technology Office of Library[2], Ludong University
Yantai, Shandong 264025, P.R. China
Shu-yan Bai
YanTai Vocational College
Yantai, Shandong 264025, P.R. China
E-mail: hyyang@yeah.net

ABSTRACT

Internet congestion control algorithm decides the quality of service (QoS) of network. Based on the mechanism of TCP windows control, a fuzzy logical controller (FLC) is presented as a substitute for active queue management (AQM) algorithm. The FLC is applied to realize the switch between a sliding-mode controller (SMC) and a state feedback controller (SFC). Finally, the stability of the new algorithm is proved.

Keywords: Quality of service; congestion control; Fuzzy Logical Controller; Active Queue Management; algorithm; switch.

1. INTRODUCTION

The development of new active queue management (AQM) routers will play a key role in meeting future increasing demand for performance in Internet applications. Nowadays, for ensuring the quality of service (QoS) and the capacity of Internet, the sources of Internet apply TCP congestion control algorithms to avoid network congestion, such as TCP Reno [1] (and its variants), and the link nodes use AQM schemes to improve the serving capacity of Internet, such as DropTail [1], RED [2]. However, the existing congestion control algorithm which are based on ”trial-and-error” methods employed on small test beds may be ill-suited for future networks [3,4] where both communication delay and network capacity can be large, which has been proved by the fact of two times to revise the parameters of the RED algorithm [5,6]. This has motivated the research on theoretical understanding of TCP congestion control and the search for protocols that scale properly so as to maintain stability in the presence of these variations. Many researches are motivated by these unstable phenomena, and lots of ameliorated algorithms are presented with modifications such as ARED [6], SRED [7] and REM [8] so on.

Based on the views of the network’s optimization, Kelly et al [11, 12] have developed a network framework with an interpretation of various congestion control mechanisms. They proposed a prime algorithm for TCP rate control and a dual algorithm for AQM scheme, which generalize the Additive Increase/Multiplicative Decrease (AIMD) congestion avoidance strategy [1] to large-scale networks. The advances in mathematical model of Kelly’s primal algorithm have stimulated the research on the analysis of the behavior such as stability, robustness from the views of the control theory [13, 14]. The reader is also directed the recent paper [9, 10, 15] which provide additional control-theoretic perspective on congestion TCP networks.

In this paper, our research is to design an AQM schemes that maintains the queue sizes stability. A switch controller based on the fuzzy logical controller is presented as Internet congestion control algorithm. When the system states are far off the equilibrium, a slide mode controller (SMC) is applied to control the states returning the neighbors of the equilibrium in a limit time. For decreasing the vibration of SMC, a state feedback controller (SFC) is gradually used to replace SMC in a small region of the origin. A fuzzy logical controller is introduced to realize the switch between SMC and SFC to adjust the congestion probability and advance the QoS of the network.

The remainder of the paper is organized as follows. The network congestion control model is introduced in section 2. The design and the stability of the switch controller is presented in section 3 and section 4. Finally, a conclusion is given in Section 5.

2. MODEL OF INTERNET CONGESTION CONTROL SYSTEM

In [11], a nonlinear dynamic model for TCP congestion control was derived, where the network topology was assumed to be a single bottleneck link with *M* homogeneous TCP sources that share the bottleneck link and have toughly the same RTTS, but don’t necessarily transverse the same path. For TCP with ECN, the AIMD behavior in congestion avoidance phase can be modeled as follows:

$$\frac{1}{w_i(t)}(1-p(t))-\frac{w_i(t)}{2}p(t)$$

where $p(t)$ is the marking probability, $w_i(t)$ is the windows sizes. The dynamical model of *M* TCP connections through a congested AQM router.

$$\begin{aligned}\dot{r}(t) &= \frac{M}{D^2(t)}-(\frac{M}{D^2(t)}+\frac{r^2(t)}{2M})p(t) \\ \dot{q}(t) &= r(t)-C\end{aligned} \tag{1}$$

and

$$D(t)=\frac{q(t)}{C}+T_p,$$

where $r(t)$ is the incoming traffic rate per unite time, $q(t)$ is the queue length on the router at time t, M is the

* This work was supported by the National Natural Science Foundation of China (under the grant 60574007), the Natural Science Foundation of Shandong Province (under the grant Y2006G01), the Science Foundation of LuDong University (under the grant 22320301) and the Postdoctoral Science Foundation of Northeastern University (under the grant 20060110), P R China.

number of TCP sessions, C is the link capacity, $D(t)$ is the round-trip time and T_p is the propagation delay. Let $e_r = r(t) - r_0$, $e_p = q(t) - q_0$, and $D_0 = q_0 / C + T_p$, where (r_0, q_0) is the desired point, the linearizing dynamical model is described as

$$\begin{aligned} \dot{e}_r(t) &= -a_1 e_q(t) - a_2 e_r(t) - bp(t) \\ \dot{e}_q(t) &= e_r(t) \end{aligned} \tag{2}$$

where $a_1 = \frac{2M}{D_0(2M + C^2 D_0^2)}$, $a_2 = \frac{2MC}{2M + C^2 D_0^2}$, and $b = \frac{M}{D_0^2} + \frac{C^2}{2M}$. Denoting $x_1 = e_p$, $x_2 = \dot{e}_p$, $u = p$, the system (2) can be rewrite as

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -a_1 x_1 - a_2 x_2 - bu \end{aligned} \tag{3}$$

We have the following form

$$\dot{x} = Ax + Bu \tag{4}$$

where $x = [x_1 \quad x_2]^T$, $A = \begin{bmatrix} 0 & 1 \\ -a_1 & -a_2 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ -b \end{bmatrix}$.

3. AQM SCHEME BASED ON FLC

In this section, we discuss the design of new AQM scheme for Internet congestion control system. Firstly, we create a SMC and a SFC to adjust the congestion probability. When the system states are far from the origin of the operating point, the SMC will take a major part of control to give a fast transient response. When the states are approaching the equilibrium values, the SFC will gradually replace the SMC, in order to avoid chattering. Secondly, a FLC is presented to realize the switch between the SMC and SFC.

3.1. The design of SMC

Define a sliding plane $\sigma = x_1 + cx_2 = 0$, where c is a constant. Obviously, the sliding plane is stable. Let $s = [1 \quad c]$, we can choose the equivalent control

$$u_{eq} = -(sB)^{-1} sAx$$

The final control is realized as

$$u = u_{eq} + u_d$$

where $u_d = -(sB)^{-1} k_m \operatorname{sgn}(\sigma)$, $k_m > 0$, and

$$\operatorname{sgn}(\sigma) = \begin{cases} 1, & \sigma > 0 \\ 0, & \sigma = 0 \\ -1, & \sigma < 0 \end{cases}$$

Defining a Lyapunov function $V = \frac{1}{2}\sigma^2$, we have

$$\begin{aligned} \dot{V} &= \sigma\dot{\sigma} \\ &= \sigma[s\dot{x}] \\ &= -\sigma k_m \operatorname{sgn}(\sigma) \\ &= -k_m |\sigma| \\ &\le 0. \end{aligned}$$

3.2. The design of SFC

Design a state feedback controller $u = k_f x$, where the control gains $k_f = [\frac{1/c - a_1}{b} \quad \frac{1.2/c - a_2}{b}]$. We apply the same Lyapunov function V, we have

$$\begin{aligned} \dot{V} &= \sigma\dot{\sigma} \\ &= x^T s^T s\dot{x} \\ &= x^T \begin{bmatrix} 1 & c \\ c & c^2 \end{bmatrix} (A + Bk_f)x \\ &= -x^T \begin{bmatrix} 1 & 0.2 \\ c & 0.2c \end{bmatrix} x. \end{aligned}$$

Let $c = 0.2$, we obtain

$$\dot{V} = -(x_1 + 0.2x_2)^2 \le 0.$$

3.3. Switch controller based on FLC

Based on the SMC and the SFC presented in front of subsections, we design a FLC to achieve the gradual switch between SMC and SFC. The rules of the FLC is defined as follows:

Rule 1: If σ is PE then $u = -(sB)^{-1}(sAx + k_m)$;

Rule 2: If σ is ZE then $u = k_f x$;

Rule 3: If σ is NE then $u = -(sB)^{-1}(sAx - k_m)$.

Where PE, ZE and NE are fuzzy level of σ, and their membership functions are shown in Fig. 1.. By applying the weighted sum defuzzification method, the output of the FLC is given by

$$u(x) = \frac{\sum_{i=1}^{k} \mu_i(x) u_i(x)}{\sum_{i=1}^{k} \mu_i(x)}$$

where $u_i(x)$ is control input, and $\mu_i(x)$ is membership degree.

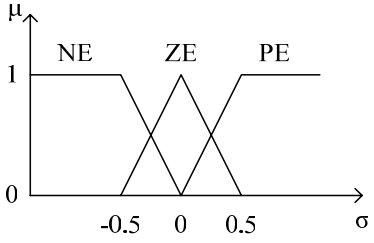


Fig. 1. Fuzzy levels of σ

4. STABILITY OF SWITCH CONTROLLER

Suppose the Lyapunov function is same as the last section, we prove the stability of the Internet congestion control system with the switch controller. Known from the Fig. 1., there is $\sum_{i=1}^{k} \mu_i(x) = 1$ and k = 2. The output of the FLC is

$$u(x) = \sum_{i=1}^{2} \mu_i u_i / \sum_{i=1}^{2} \mu_i = \sum_{i=1}^{2} \mu_i u_i$$

where control input $u_i(x)$ is an SMC or an SFC. Since $\mu_1 + \mu_2 = 1$, we have

$$\begin{aligned} \dot{x} &= Ax + Bu \\ &= Ax + B(\mu_1 u_1 + \mu_2 u_2) \\ &= \mu_1(Ax + Bu_1) + \mu_2(Ax + Bu_2) \end{aligned}$$

Suppose u_1 is an SMC, u_2 is an SFC, then we obtain

$$\begin{aligned} \dot{V} &= \sigma\dot{\sigma} \\ &= \sigma(s\dot{x}) \\ &= \mu_1\sigma s(Ax + Bu_1) + \mu_2\sigma s(Ax + Bu_2) \\ &= -\mu_1 k_m |\sigma| - \mu_2(x_1 + 0.2x_2)^2 \\ &= -\mu_1 k_m |\sigma| - \mu_2\sigma^2 \\ &\le 0. \end{aligned}$$

Since $\dot{V} = 0$, if only if $\sigma = 0$, only the sliding-model plane is satisfy for $\dot{V} = 0$. By LaSalle's Invariance principle [16], the system is asymptotically stable on the sliding-model plane.

5. CONCLUSIONS

In this paper, a new AQM algorithm is presented to adjust Internet congestion probability. The new AQM algorithm is a switch controller based on the Fuzzy logical controller. When the system states are far off the equilibrium, a SMC is applied to control the states returning the neighbor of the equilibrium in a limit time. For decreasing the vibration of SMC, a SFC is gradually used to replace SMC in a small region of the origin. A FLC is introduced to realize the switch between SMC and SFC.

6. REFERENCES

[1] V. Jacobson, "Congestion avoidance and control", Proceedings of ACM SIGCOMM'88, Stanford, CA, 1988, pp.314-329.

[2] S. Floyd, V. Jacobson, "Random early detection gateways for congestion avoidance", IEEE/ACM Transactions on networking, 1(4), 1996, pp.397-413.

[3] S. Low, F. Paganini, et al, "Dynamics of TCP/RED and a scalable control", IEEE INFOCOM, 2002, pp.239-248.

[4] S. Low, F. Paganini, et al, "Internet congestion control", IEEE control systems magazine, No.4, 2002, pp.28-43.

[5] S. Floyd, M. Handley, et al, "Equation-based congestion control for unicast applications", Proceedings SIGCOMM Symposium on Communications Architecture and Protocol, August 2000, pp.43-56.

[6] S. Floyd, R. Gummadi, et al, "Adaptive RED: An Algorithm for Increasing the Robustness of RED", Technical Report, 2001.

[7] T. Ott, T. Lakshman, et al, "SRED: stabilized RED", Proc. of The INFOCOM'99, New York, 1999, pp.1346-1355.

[8] S. Athuraliya, S. Low, et al, "REM: active queue management", IEEE network, 15(3), 2001, pp.48-53.

[9] S. Low, R. Srikant, "A mathematical framework for designing a lowloss, low-delay Internet", Networks and Spatial Economics, No.4, 2004, pp.75-101.

[10] F. Kelly, "Fairness and stability of end-to-end congestion control", European Journal of Control, No.9, 2003, pp.159-176.

[11] F. Kelly, Mathematical modeling of the Internet, Mathematics Unlimited-2001 and Beyond, Springer-Verlag, Berlin, 2001.

[12] F. Kelly, A. Maullo, et al, "Rate control for communication networks: shadow prices, proportional fairness, and stability", Journal of the Operational Research Society, 49(3), 1998, pp.237-252.

[13] L. Massoulie, "Stability of distributed congestion control with heterogeneous feedback delays", IEEE Transactions on Automatic Control, 47(6), 2002, pp.895-902.

[14] Y.P. Tian, H.Y. Yang, "Stability of the Internet congestion control with diverse delays," Automatica, 40(9), 2004, pp.1533-1541.

[15] R. Srikant, The mathematics of Internet congestion control, Birkhauser, 2004.

[16] H. Khalil, Nonlinear systems, 2nd edition, Prentice Hall, 1996.

Predicting the runtime of tasks based on neural networks on grids

Jingbo Yuan [1, 2], Shunli Ding [1, 2], Jiubin Ju [1], Liang Hu [1]
[1]College of Computer Science and Technology, Jilin University, Changchun, Jilin, China
[2]Department of Computer Engineering, Northeast University at Qinhuangdao, Qinhuangdao,Hebei,China
Email: jingboyuan@hotmail.com , dingsl@163.com

ABSTRACT

Application run-time information is a fundamental component in application and job scheduling. Predicting the runtime of a task, an important component of the resource management, plays an important role in the task scheduling and the resource using in computational grid. Such techniques can improve the performance of scheduling algorithms. However, the runtime of a task is a variable affected by many factors, accurate predictions of runtimes are difficult to achieve for parallel applications running in shared environments where resource capacities can change dynamically over time. This paper presents a predicting model for task's runtime based on BP neural networks considering several factors. The method has many advantages including small network structure, quick learning and use conveniently etc. The result of prediction indicates that the method is effective and has higher accuracy.

Keywords: computational grid, predicting the runtime of tasks, neural networks, and BP algorithm.

1. INTRODUCTION

The fundamental function of the computational grid is organizing many complex distributed computational resources to buildup a huge distributed loosely coupled computing system using the grid technologies. The main duty of such computing system is to complete those complex super-large-scale simulation tasks and computing tasks. The major problem of the computational grid is to organize, access and manage computing tasks though there are data in system, because the science computation issues are generally computation-intensive that have very large CPU overhead. Furthermore, grid resources will dynamically increase or decrease over time, which require that resource management must adequately consider and solve the problem well, so the grid computing system should be able to adapt such change of scale or structure, and resource managements or applications must adapt these dynamic actions in order to gain higher effective services. Therefore, in such grid environment, the scheduler need consider the dynamic changes of resource performance, and carefully choose how and where to run their tasks.

Resource performance prediction is a key to reach an optimal scheduling. Our goal is to provide high level predictive services in shared, unreserved distributing computing environments that are useful to applications make different kinds of scheduling and adaptation decisions, that is, to support adaptation frameworks by providing performance predictions on which valid scheduling decisions can be made. Our work is to research one form of such application-level performance predictions – predicting the runtime of task.

The next section summarizes our approach to predicting application run times and the approaches of other researchers. Section 3 describes the BP networks and BP Algorithm we adopt. Section 4 presents a predicting model for tasks runtime and relevant BP neural network structure. Section 5 presents the experiment results and analyzes the performance of using our prediction technique. Section 6 presents our conclusions and future work.

2. RELATED WORKS

Performance monitoring is capacity that the computing grid must have, which has been comprehensive recognized [1, 2]. The monitoring information is mainly used to optimize application execution. Starting in the late 90s, research began on how to build scalable systems for measuring the dynamic properties of distributed environments, leading to such well known systems as the Network Weather Service [3] (NWS), RPS [4] and so on. NWS focus mostly on network performance monitoring. It measures network resources and gives the future performance prediction of these resources with statistical technology. The application-level scheduler is able to use the prediction to get better executing performance. RPS is based on explicit resource-oriented prediction using the techniques of linear time series analysis to predict available CPU time and the running time of compute-bound task. The resource signal is host load. RPS and NWS group have independently demonstrated that host load prediction is feasible.

In [5] presents a holistic approach to estimation that uses *rough sets* theory to determine a similarity template and then compute a runtime estimate using identified similar applications in a history.

Both NWS and RPS adopts time series method, so the system is basically linear model. In practice, the runtime of a task is affected by many factors; so linear model cannot describe system features well. The neural networks have well non-linear mapping capability, fast parallel processing ability, powerfully self-learning and self-organizing ability and so on. BP network have simple structure, powerful simulation ability and implement conveniently etc. BP networks have been widely used for evaluation and prediction, expert system, fault diagnosis and so on. With a view to several factors like task scale and host load, we present a dynamic predicting system to predict the runtime of task on hosts using improved BP algorithm. It can fleetly provide reliable information for the task scheduling and the resource management in computing grid environment.

3. BP ALGORITHM

3.1 BP networks and BP algorithm

BP (Back-Propagation) method is a learning procedure for multilayered, feed-forward neural networks. Typically, BP networks structure includes three-layers: input layer, hidden layer (or middle layer) and output layer. Hidden layer may be a layer or multilayer. The neurons (or units) on same layer are not coterminous each other, neurons between the

previous layer and in the back layer are connected by the weight of the connection. An error-minimization method called gradient descent is used to reduce the output error by adjusting the network weights.

BP algorithm includes the forward-propagation and the back-propagation. On forward-propagation, first the samples are transmitted to the hidden layer from input layer and calculated and then transmitted to the output layer, that is, to verify that the output is correct. If it is not correct, the algorithm goes back and modifies the weights of the inputs to the output layer, such that a correct response would have been received. These inputs are, in turn, outputs of a hidden layer, and the weighted inputs of the middle layer must now also change to reflect the new output. This process continues successively until input weights in each of the layers have changed and error signal tend to minimum. There are essentially two passes through a neural network: the forward learning pass and the backward re-weighting pass. This process continues until the network responds for each input unit with an output unit that is sufficiently close to the desired one or the error signal arrives at given minimum.

Suppose that in BP networks the input layer has n units, $X = (x_1, x_2, \cdots, x_n)$ represents the input sample, the hidden layer has h units and the output is $U = (u_1, u_2, \cdots, u_h)$, and the output layer has m units and the output is $Y = (y_1, y_2, \cdots y_m)$.

1) Forward-propagation calculating procedure

Calculating the outputs of the hidden layer units and output layer units, using the formula (1) and (2)

$$u_j(t) = f(\sum_{i=1}^{n} w_{ij} \cdot x_i(t) + \theta_j), \ j = 1,2,...h \quad (1)$$

$$y_k(t) = f(\sum_{j=1}^{h} w_{jk} \cdot u_j(t) + \gamma_k), k = 1,2,...m \quad (2)$$

The sigmoid function is as the activation function defined by the formula (3).

$$f(x) = \frac{1}{1+e^{-x}}$$

$$\text{and its derivative } \frac{\partial f}{\partial x} = f(x)(1 - f(x)) \quad (3)$$

The network computes the squared error E and the system total error E_r according to (4).

$$E = \frac{1}{2}\sum_{k=1}^{m}(d_k - y_k(t))^2$$

$$E_T = \sum E \quad (4)$$

Here w_{ij} is the weight of the connection between the *i*th and *j*th units, θ_j is the threshold value of *j*th units of the hidden layer, w_{jk} is the weight of the connection between the *j*th and *k*th units, γ_k is the threshold value of the *k*th units of the output layer, t is epochs (learning times), d_k is the teacher signal of the *k*th units of the output layer, that is, the desired output.

2) Back-propagation calculating procedure

Correcting the weight using the formula (5)

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}}$$

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij} = w_{ij}(t) - \eta \frac{\partial E}{\partial w_{ij}} \quad (5)$$

Here η is the step length of learning, called the learning rate, generally, η=0.01~1.

3.2 The improved BP algorithm

For accelerating the convergence, we adopt the improved BP algorithm with a momentum term [6]. The momentum term can enlarge the oscillation of the BP network, consequently speedup the convergence when the network keeps away from convergence point or traps in the local minimum point. Here, adjusting and correcting the weight and the threshold of the output layer using the following formula (6).

$$w_{jk}(t+1) = w_{jk}(t) + \eta p(t) + \alpha \Delta w_{jk}(t)$$

$$\gamma_k(t+1) = \gamma_k(t) + \eta p(t) + \alpha \Delta \gamma_k \quad (6)$$

here

$$p(t) = \frac{\partial E}{\partial w_{jk}} = y_k(1-y_k)(y_k - d_k)u_j(t)$$

$$j = 1,2,...h; \quad k = 1,2,...m$$

Adjusting and correcting the weight and the threshold of the hidden layer using (7).

$$w_{ij}(t+1) = w_{ij}(t) + \eta q(t) + \alpha \Delta w_{ij}(t)$$

$$\theta_j(t+1) = \theta_j(t) + \eta q(t) + \alpha \Delta \theta_j \quad (7)$$

here

$$q(t) = \frac{\partial E}{\partial w_{jk}} = u_j(1-u_j)(\sum_{k=1}^{m} p(t) w_{jk}) x_i(t)$$

$$i = 1,2,...n; \quad j = 1,2,...h$$

The momentum controls how much the weights are changed during a weight update by factoring in previous weight updates. It acts as a smoothing parameter that reduces oscillation and helps attain convergence. This must be a real value between 0.0 and 1.0, and a typical value for momentum is 0.9.

To overcome the local minimum, adopt the error distributing function to adjust the momentum coefficient α and the learning rate η [7]. The formula (8) defines the average error function of the output units and (9) is the error distributing function.

$$E_A = \sum_{k=1}^{m} [(d_k - x_k)^2 / d_k]^{1/2} / m \quad (8)$$

$$P(\Delta E_A, t) = e^{-\Delta E_A (C_0 + \beta t)}$$

$$\Delta E_A = \Delta E_A(t) = E_A(t) - E_A(t-1) \qquad (9)$$

Here C_0 and β are constants and C_0=10, 0<β<1. P denotes the adjusting probability of α and η. P will decrease with increasing of the iteration times t, thereby the learning process converge steadily.

Define:
η'=η (1+η_1×m)
When ΔE_A >0 and η'>C_1 (η+α). Adjusting α and η according to (10), or else not adjusting.

$$\begin{cases} \eta = \eta' \\ \alpha = 0 \end{cases} \qquad (10)$$

Here η_1=η_1 (t) =η_1 (t-1) ×R, η_1 (0)>1, 5≤C_1<1, R<1. η_1 (0) is the initial value of η_1.

4. THE PREDICTING MODEL

4.1 The factors of influence the runtime

Many factors have impact on the runtime of a task on hosts, of which include the character of tasks like the type of a task, the issue scale, the length of program, the memory, the number of computing nodes and so on. Some performances of a host have huge influence too, as the processor speed (CPU capacity), CPU availability, host load, the storage volume and the type of operating system etc. In this paper, we choose three major factors, host load and processor speed (or CPU speed) and computing scale of a task, as parameters of the predicting model considering to following reasons.

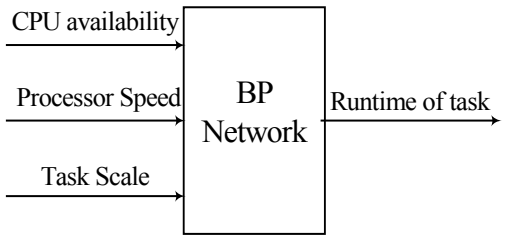


Fig. 1 Predicting Model

- Can objectively reflect the influence on the runtime
- Should be independent each other in measurement and control
- Be easy to measure or estimate, and low overhead
- Propriety in amount and convenient for calculation

Fig.1 presents the predicting model based on BP network. The inputs of the model are CPU availability, processor speed and task's computing scale. The output is the runtime of task on hosts.

4.2 BP neural network structure and parameters design

To a multilayer neural network, it is necessary to determine how many layers the hidden layers have. By Kolmogrov theorem, a three-layer forward neural network can implement a mapping from any n dimension to m dimension [8]. Therefore, this system uses simple three layers feed-forward BP network: an input layer, a hidden layer and an output layer.

According to analysis for the influence factors on the runtime of task in 4.1, we design the BP network structure to predict the runtime of task as following fig. 2. There, the input layer has 3 units and the output layer has 1 unit. Selecting of the number of the hidden layer units is a complex problem. If the hidden layer has fewer units, it will cause that the network training hasn't results, and can't recognize the samples that were not learned previously, thereby fault tolerance is worse. Contrariwise, if the hidden

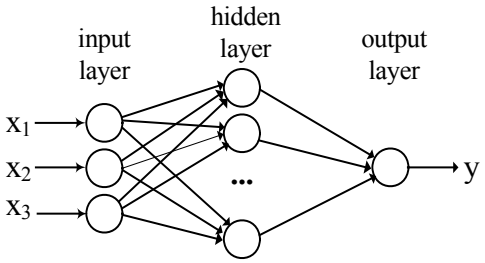


Fig. 2 BP Network Structure for Predicting

layer has more units, it may cause that the network training is longer (over-learning), and can't ensure the error is optimal. In the light of the Hecht-Nielsen theory [9], we set the number of the hidden layer units h=2n+1=7.
Define:
n: the number of the input layer unit, n=3
X=(x_1, x_2, x_3): the input vector, represent respectively processor speed, host load at present and the computing scale of a task.
m: the number of the output layer unit, m=1
Y=(y): the output vector, represent the runtime of a task on a host
h: the number of the hidden layer unit. In the light of the Hecht-Nielsen theory, let h=2n+1=7
α: the momentum coefficient, initial α=0.9
η: the learning rate, initial η=0.85
ε: the error tolerance, the total error upper limit, that is, the iteration stops when E<ε
All initial connection weights are generated by random function.

To avoid overflow and affect convergence during training process, it requires that data have same unitary proportion, so the system uses the formula (11) to normalize samples, which will ensure the scope of data is in [0,1].

$$Xn = \frac{x - \min x}{\max x - \min x} \qquad (11)$$

5. EXPERIMENTS

It had presented the grid architecture based on performance prediction in [10]. In the system, the Resource Selector chooses advisable host resources to build up a candidate hosts list. The predicting system will forecast the runtime of a submitted task on these candidate hosts. The Resource Information Management collects these resources information like CPU availability and processor speed of candidate hosts at present. If a host load is heavy (for example, CPU availability > 60%), in view of load balance, the scheduler shouldn't assign a new task to the host, so such host isn't regarded as a candidate host in order to reduce computing and enhance predicting speed. Users submit a task with the task's size to the Task Management. Considering the calculation of different kinds in grid, the system uses the Time Complexity to characterize the task's size. Due to possible great difference among resource availability and task size and machine speed, the difference between the runtime of tasks is possible larger. Therefore, for the sake of reducing the difference among the input samples and the training error, we design two kinds of the BP network; one deal with small-scale calculation and another deal with general scale calculation.

Furthermore, after many tests, we found that about 70% relative error is less than 1% if the learning samples are as to the test data. Therefore, the actual data of a completed task will be recorded in files as history information, like current CPU availability, the processor speed, the task scale, actual runtime and so on. The training of the network may be updated by the history information to acquire higher precision.

The system is tested in Linux with C program. We designed two cases to test the model.

Case1: The network deals with small-scale calculation, and is trained by 12 samples and tested by 30 data. Here the error tolerance ε=0.001.

Case2: The network deals with general scale calculation, and is trained by 12 samples and tested by 30 data. Here the error tolerance ε=0.001.

Fig. 3 and Fig. 4 represent the relative error between the predicting results and outputs. Table 1 represents average relative error and epochs of two cases. The relative error is calculated as following formula (12).

$$error = \left| \frac{t_{act} - y}{t_{act}} \right| \times 100\% \qquad (12)$$

Here t_{act} denotes actual runtime of a task, and y denotes predicting value.

The result shows the most relative errors is under 5% and the total average error is 3.795%. Aiming at the goal of the predicting runtime and scheduling, such error scope should be feasible. If the error tolerance is much smaller it will cause increasing iterations and excess training (over learning), consequently result in local optimum and influence convergence. The average iterations is about 23000, which means that the predicting system is faster and the computing process won't have an impact on the real execution time of tasks. Thus it can be seen that this predicting model of task's runtime is effective and the predicting results are satisfactory.

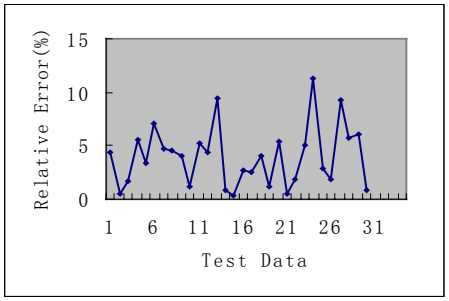


Fig. 3. Test Date vs Relative Error of Case 1

6. CONCLUSIONS AND FUTURE WORK

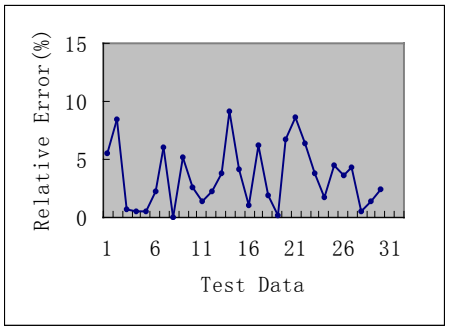


Fig. 4. Test Date vs Relative Error of Case 2

This paper presents a dynamic forecasting method based on BP neural network and improved BP algorithm to predict the runtime of tasks in grid. The system achieves a high level of accuracy for exemplar applications. The method is a useful and effective for resource performance prediction with the following characteristics:

Table 1 Average Error and Epochs

	Case 1	Case 2	Average
Average Relative Error(%)	4.06	3.53	3.795
Average Epochs	15000	31500	23250

The prediction model is real-time and on-line learning. It dynamically adapts in time to changes in the environment by accounting for changes in the workload and task scale.

The method may deal with many factors that have influence on the runtime of tasks, which make the model can better represent the relation between the actual characteristic of grid resource and task, consequently the calculating result is close to the real value.

Predictions can be computed with low overhead costs - the network can directly use to predict after a train - making it computing possible to have less impact on the scheduling and actual task execution.

The predicting network has some advantages including simple structure, fewer hidden units, faster learning, higher accuracy, on-line predicting, convenient use and so on. Using improved BP algorithm is good for enhancing

convergence and reducing computing.

In the future, we will integrate the method with the scheduling algorithm in [10] to assign tasks in grid. And we hope also to be able to design new scheduling algorithm based on the method. In extending the scope of this work to new scenarios, we also intend to look at the possibility of on-demand refinement of the model to include additional factors, such as disk or memory behavior, server load and communication network performance (for example bandwidth), in order to do a better job of assigning between tasks and hosts, and hence enhance application performance in computing grid environment. The estimation techniques should also have potential application for estimating quality-of-service levels in grid environments. We will investigate these areas of application.

8. REFERENCES

[1] B. Tierney, R. Aydt, et al, “A Grid Monitoring Architecture.Grid forum working group document”, Grid Forum, Feburary 2001,http://www.gridforum.org.

[2] F. Berman, A. Chien, et al, “The GrADS Project: Software Support for High-Level Grid Application Development”, International Journal of High Performance Computing Applications, 2001,15(4): 327-344.

[3] R. Wolski, N. Spring, et al. “The network weather service: A distributed resource performance forecasting service for metacomputing [J]”, Journal of Future Generation Computing Systems, 1999, 15(5-6): 757-768.

[4] P. A. Dinda, D. R. O’Hallaron, “An extensible toolkit for resource prediction in distributed systems”, Technical Report CMU-CS-99-138, School of Computer Science, Carnegie Mellon University, July 1999

[5] Shonali Krishnaswamy, Seng Wai Loke, et al, “Estimating Computation Times of Data-Intensive Applications”, IEEE DISTRIBUTED SYSTEMS ONLINE 1541-4922 © 2004 Published by the IEEE Computer Society, 5(4), April 2004.

[6] R.P. Lippman, “An introduction to computing with neural net”, IEEE ASSP Magazine, 1987, 4(2): 4-22.

[7] K.H. Fun, S.B. Wang, “Force modeling and forcasting in creep feed grinding using improved BP neural network”, Int. J. Mach. Tools Manufact, 1997, 37(8):I 167-I 178.

[8] Niclscn R H.Kolmogrov's, “Mapping Neutral Network Existence Theorem[R]”, Proc.IEEE First International Conference on Neural Networks, San Diego, 1987,Vol. 3:1114.

[9] R. Hecht-Nielsen, “Theory of Backpropagation Neural Networks”, Proceedings of the International Joint Conference on Neural Networks,Washington, 1989,Vol. 1: 593-605.

[10] J.B. Yuan, S.L. Ding, et al, Task Scheduling Using Performance-Driven. Proceedings of the fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 2005 Vol.6, pp.3899-3904.

The Modeling Mechanism of Cascade-forward Back Propagation Neural Network Based on Self-correlation and Its Application

Wenzhan Dai

College of Information and Electronic Engineering, Zhejiang Gongshang University

Hangzhou, China

Email: dwzhan@zist.edu.cn

And

Junfeng Li[1,2](Corresponding author)

[1.] Department of Automatic control, Zhejiang Sci-Tech University

Hangzhou, China

[2.] College of Mechanical Engineering, Donghua University

Shanghai, China

Email: ljf2003zz@163.com

ABSTRACT

In this paper, based on self-correlation, a new method of modeling mechanism of cascade-forward back propagation neural network is put forward. First, the original sequence is divided into sub-sequences according to the prominence of self- correlative coefficients. Then, the framework of neural network is structured and input and output are defined reasonably; and then ameliorating BP algorithm with momentum factor is studied. Finally the method presented in this paper is applied in building the model of total residence number and the results show the model has very high precision.

Keywords: Self-correlation Theory, Feed-forward Neural Network, Modeling and Forecasting.

1. INTRODUCTION

Neural networks (NN) are composed of simple elements, which are inspired by biological nervous systems. It is because they have ability of self-organization, self-adaptation and self-learning that they can provide powerful means for solving many problems especially in nonlinear, real-time, adaptive, and blind signal processing [1,2,3,4,5]. The popularity of neural network methodology is rapidly growing in a wide variety of areas from basic research to date mining applications, business forecasting, engineering and others. Neural network can be trained to perform any particular function by adjusting weights among neurons It has already been proved that any nonlinear function can be approached by backward propagation (BP) network which consists of at least one S-type hidden layer and one linear output layer.

On the other side, Self-correlation is one of important theory to analyze sequence. The self-correlative coefficient shows mutual relation between sequence and its different time shift [6,7,8,9]. It is proved that using self-correlative coefficients can help us understand the inherent rule of sequence. So it is obvious that a new modeling approach that combines the strongpoint of the self-correlation with neural network can greatly improve model precision.

In this paper, based on self-correlation theory and cascade-forward back propagation neural network, a new method of modeling is put forward. First, the original sequence is divided into sub-sequences according to its prominence of self- correlative coefficients. Then, the framework of neural network is structured and input and output are defined reasonably; and then ameliorating BP algorithm with momentum factor is studied. Finally the method presented in this paper is applied in building the model of total residence number and the results show the model has very high precision.

This paper is supported by the Natural Science Fundation of Zhejiang Province, P.R .China(No： 602016) and Returnee Foundation of Ministry of Personal, P.R.C.(No:2003236)

2. THE MODELING MECHANISM OF CASCADE-FORWARD BACK PROPAGATION NEURAL NETWORK BASED ON SELF-CORRELATION DATA PROCESSING

Denote the non-negative time sequence by

$$x_t(l),\ l = 1,2,...n\,. \tag{1}$$

In order to improve the precision of model, we can process the sequence by means of self-correlation theory [9]. The sequence (1) is analyzed by means of self-correlation theory and its self-correlative coefficients $r_1, r_2, ..., r_{n-1}$ can be obtained. The r_k is named as self-correlative coefficient with k time-shift which shows the correlative degree between x_t and x_{t+k}. r_k is defined as follow：

$$r_k = \frac{\sum_{i=1}^{n-k}(x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

It is well known that it is difficult to judge whether the mutual relation is prominent when r_k is about 0.5. So the prominence of self-correlative coefficients must be verified.

A statistical variable t_k is introduced as follow:

$$t_k = \frac{|r_k|}{\sqrt{(1+r_k^{\,2})/(n-2)}} \qquad (3)$$

Where n is the number of data points and (n-2) is free degree. Comparing t_k with "t numerical value table", we can consider self-correlative coefficients to be prominent under the condition of given confidence if t_k is larger than or equal to the corresponding numerical value of "t numerical value table". Generally, the confidence is more than 95%.

According to above theory, the sequence (1) is divided into sub-sequences as follow:

$$\begin{aligned} X_1 &= [x_1(1), x_1(2), ..., x_1(n-m)] \\ &= [x(m), x(m+1), ..., x(n-1)] \\ X_2 &= [x_2(1), x_2(2), ..., x_2(n-m)] \\ &= [x(m-1), x(m), ..., x(n-2)] \qquad (4) \\ &\cdots\cdots \\ X_m &= [x_m(1), x_m(2), ..., x_m(n-m)] \\ &= [x(1), x(2), ..., x(n-m)] \end{aligned}$$

Where m is the number of self-correlation coefficients $r_i (i = 1,2,...,m)$, which are prominent.

The network structure of the cascade-forward back propagation neural network is as Fig. 1.:

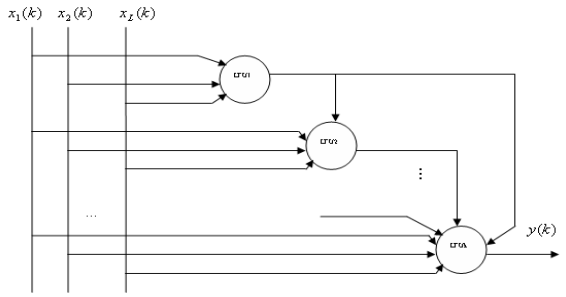


Fig. 1. The framework of the cascade-forward backprop neural network

Let inputs of cascade-forward back propagation neural network be $x_1(k), x_2(k) \ldots x_m(k)$(here L=m) and output be $y(k)=x(m+k)$, k=1,2,…n-m

Train the cascade-forward back propagation neural network by means of BP algorithm with momentum value until the objective function reaches anticipant

After training is finished the trained-network can be used to predict output of next time because of neural network's general ability. From the structure of network it can be regarded as prediction of next time.

3. BP ALGORITHM WITH MOMENTUM WEIGHT

BP is a specific learning algorithm. Nevertheless, the momentum term is often utilized The BP algorithm is used for updating the weights of each layer based on the error present at the network output.

3.1 Forward propagation process

Suppose that w_{ji} is weight between the j^{th} neuron of hidden layer and the i^{th} neuron of input layer and v_{kj} is weight between the k^{th} neuron of output layer and the j^{th} neuron of hidden layer.

Input layer: the output value o_i of the i^{th} neuron equals to input value x_i;

Hidden layer: the output net_j^1 of the j^{th} neuron equals to the sum of all the weighted output o_i of its forward layer and the bias.

$$net_j^1 = \sum_i w_{ji} o_i + \theta_j^1 \qquad (5)$$

output value a_j:

$$a_j = \log sig(net_j^1) \qquad (6)$$

Out layer: the output value net_j^2 of the k^{th} neuron equals to the sum of all the weighted output value a_j of its forward layer and the bias.

$$net_k^2 = \sum_j v_{kj} a_j + \theta_k^2 \qquad (7)$$

Output value

$$y_k = \log sig(net_k^2) \qquad (8)$$

3.2 Back propagation process

Define the following error function E_P:

$$E_P = \frac{1}{2}\sum_k [y(k) - y_k]^2 \qquad (9)$$

Adjust weight between input layer and hidden layer to be as follow:

$$\Delta w_{ji} = -\eta \frac{\partial E_P}{\partial w_{ji}} \qquad (10)$$

Adjust weight between hidden layer and output layer to be as follow:

$$\Delta v_{kj} = \eta \delta_k a_j \qquad (11)$$

where $0 < \eta < 1$. η is named as learning rate.

The emendatory values of weights between input layer and hidden layer can be gotten in the same way

$$\Delta w_{ji} = \eta \delta_j o_i \qquad (12)$$

where $\delta_j = a_j(1-a_j)\sum_k \delta_k v_{kj}$.

3.3 The ameliorating BP algorithm with momentum factor

As we all know, the choice of learning rate η is very important. In general way, the larger η is, the quicker constringency of the object function is. However, larger η cause object function to run into local minimum easy.

Luckily, momentum allows the network to ignore small features in the error surface. a network with momentum can slide through such minimum instead of getting stuck So the ameliorating BP algorithm with momentum item is adopted :

$$\begin{aligned} & w_{ji}(n+1) = w_{ji}(n) + \eta\delta_j(n)o_i(n) \\ & +a\Delta w_{ji}(n)v_{kj}(n+1) = v_{kj}(n) + \\ & \eta\delta_k(n)a_j(n) + a\Delta v_{kj}(n)\theta_j(n+1) \\ & = \theta_j(n) + \eta\delta_j(n) + a\Delta\theta_j(n) \end{aligned} \quad (13)$$

where α is momentum factor.

4. EXAMPLE

In order to show the effectiveness of the approach presented in this paper, a lot of simulation was made. Just pick the model of total residence family number as example. The total residence family number of shanghai city from 1979 to 2004 is as table 1(《Statistic almanac of sahnghai-2005》).

According to formula (2), the self-correlative coefficient $r_1 - r_{16}$ is gotten as Table 2.

Table 1. Residence number of Shanghai City
《Statistic almanac of sahnghai-2005》

Unit:ten-thousand

Year	Residence Number	Year	Residence Number
1979	296.71	1992	431.67
1980	303.87	1993	438.69
1981	314.56	1994	444.38
1982	321.71	1995	450.76
1983	330.60	1996	457.49
1984	340.78	1997	461.40
1985	351.72	1998	465.72
1986	364.92	1999	470.11
1987	380.19	2000	475.73
1988	394.95	2001	478.92
1989	406.82	2002	481.77
1990	415.28	2003	486.06
1991	425.84	2004	490.58

Table 2. Self-correlation parameters

r_1	r_2	r_3	r_4	r_5	r_6
0.865	0.718	0.571	0.418	0.264	0.114
r_7	r_8	r_9	r_{10}	r_{11}	r_{12}
0.031	0.159	0.266	0.347	0.399	0.425
r_{13}	r_{14}	r_{15}	r_{16}		
0.423	0.402	0.359	0.299		

From the table2, we can see r_1, r_2, r_3 is prominent, so m is taken as 3. The following sequence is defined:

X_1 = [314.56, 321.71, 330.60, 340.78, 351.72, 364.92, 380.19, 394.95, 406.82, 415.28, 425.28, 431.67, 438.69, 444.38, 450.76, 457.49, 461.40, 465.72, 470.11, 475.73, 478.92, 481.77, 486.06]

X_2 = [303.87, 314.56, 321.71, 330.60, 340.78, 351.72, 364.92, 380.19, 394.95, 406.82, 415.28, 425.28, 431.67, 438.69, 444.38, 450.76, 457.49, 461.40, 465.72, 470.11, 475.73, 478.92, 481.77]

X_3 = [296.71, 303.87, 314.56, 321.71, 330.60, 340.78, 351.72, 364.92, 380.19, 394.95, 406.82, 415.28, 425.28, 431.67, 438.69, 444.38, 450.76, 457.49, 461.40, 465.72, 470.11, 475.73, 478.92,]

And Define the following target sequence:

Y = [321.71, 330.60, 340.78, 351.72, 364.92, 380.19, 394.95, 406.82, 415.28, 425.28, 431.67, 438.69, 444.38, 450.76, 457.49, 461.40, 465.72, 470.11, 475.73, 478.92, 481.77, 486.06, 490.58]

Let $x_1(k), x_2(k), x_3(k)$ be input and y(k)be target. The cascade-forward back propagation neural network is trained by means of the ameliorating BP algorithm with momentum item. The structure map of the neural network is same as Fig. 1. The neural network has four layers. Input layer has three neurons and output layer has one neuron. Hidden layer has two layers and every layer has ten neurons. The transfer function of hidden layer neuron is log sigmoid transfer function and others are linear transfer function. Learning rate η is 0.01 and momentum factor α is 0.9. The objective function gets to anticipant value 2.7. Fig. 2. is the training map.

Table 3. is model value and predicting value. Fig. 3. is its fitting curve.

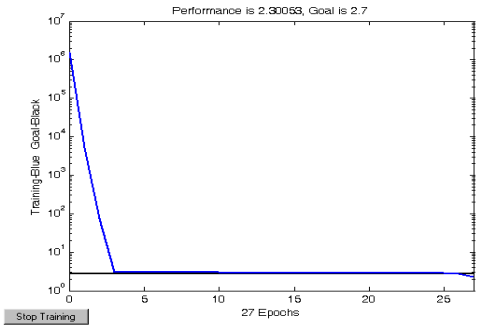


Fig. 2. The training map

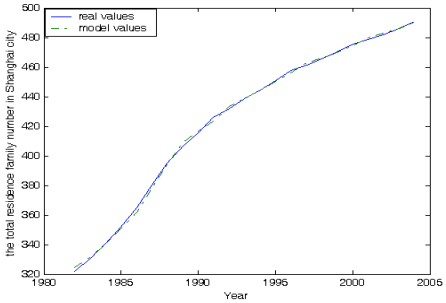


Fig. 3. The fitting curve

Table 3. Output values of the model

year	real value	model value	Relative error（%）
1982	321.71	324.72	-0.93
1983	330.60	331.90	-0.39
1984	340.78	340.00	-0.23
1985	351.72	350.57	0.33
1986	364.92	361.68	0.90
1987	380.19	378.40	0.47
1988	394.95	394.20	0.19
1989	406.82	408.98	-0.53
1990	415.28	416.60	-0.32
1991	425.84	423.50	0.55
1992	431.67	433.32	-0.38
1993	438.69	438.65	0.01
1994	444.38	444.53	-0.03
1995	450.76	450.12	0.14
1996	457.49	456.13	0.30
1997	461.40	462.91	-0.33
1998	465.72	466.29	-0.12
1999	470.11	469.87	0.05
2000	475.73	474.26	0.31
2001	478.92	479.97	-0.22
2002	481.77	482.93	-0.24
2003	486.06	486.02	0.01
2004	490.58	490.38	0.04
2005		496.62	
2006		500.40	
2007		504.1	
2008		508.06	
2009		511.94	
2010		515.72	

From Table 3.it is easy to see that worst relative error is 0.93%. The error inspection of post-sample method can be used to check quality of the model. The average relative error is 0.3%. The post-sample error is $c = S_1 / S_0 = 0.0285$ (where S_1^2 is variation value of the error and S_0^2 is variation value of the original sequence). The small error probability is $p = P\{| e^{(0)}(i) - e^{-(0)} | < 0.6745\, S_0\} = 1$. The above c and p are checked against the index JPA of Table 4. , it is found that the cascade-forward back propagation neural network model proposed by this paper has excellent precision.

Table 4. Overall model precision assessment and concrete index JPA

Model precision degree(JPA)	the small error probability	The post-sample error
Rank 1：good	0.95<=p	c<=0.35
Rank 1：qualified	0.80=<p<0.95	0.35<c<=0.50
Rank 1：just	0.70=<p<0.80	0.50<c<=0.65
Rank 1：unqualified	P<0.70	0.65<c

5. CONCLUSIONS

In this paper, based on self-correlation, a new method of modeling mechanism of cascade-forward back propagation neural network is put forward. First, the original sequence is divided into sub-sequences according to prominence of the self- correlative coefficients. Then, the structure framework of the neural network is presented and input and output are defined reasonably; and then the ameliorating BP algorithm with momentum item is studied. Finally the method is applied in building model, which has four layers. Input layer has three neurons and output layer has one neuron. Hidden layer has two layers and every layer has ten neurons. The transfer function of hidden layer neuron is log sigmoid transfer function and others are linear transfer function. Learning rate η is 0.01 and momentum factor α is 0.9. The results show the model's precision is above 99%.

6. REFERENCES

[1] F. Lin, Neuron Networks and Neuron computers, Zu-Lin Pub.Co.Taipei, Taiwan, China, 1992.

[2] D. S. Chen, R. C. Jain, "A Robust Back Propagation Learning Algorithm for Function Approximation," IEEE Trans. On NN, June 1994，5（3）: 467-479.

[3] D.E. Rumelhart, G.E. Hinton, et al, "Learning representation by back-propagating errors," Nature, vol.323, 1986, 533-536.

[4] R.A. Jacobs, "Increased rates of convergence through learning rate adaptation," Neural Networks, January 1988，1（4）: 295-308.

[5] A.A. Miniani, R.D. Wililiams, "Acceleration of back-propagation through learning rate and momentum adaptation," proceedings of International Joint Conference on Neural Networks, San Diego, CA, 1, 676-1,679.

[6] J. Sun, "State in Control for Autocorrelated Process," Journal of Industrial Engineering Management, February 2003，17（2）: 75-80.

[7] C.Q. Xu, H.L. Jin, "Spectrum Characters for Almost Perfect Autocorrelation Binary Sequences Pairs," Journal of Yanshan University, January 2003，27（1）: 13-16.

[8] X.L. Xu, J. Luo, "A Kind of New Combination Forecasting Model of Grey GM (1,1) and Neural Network," Journal of Xiamen University (Natural Science), February 2002，41（2）: 164-167.

[9] Y.L. Wang, "Forecasting Count Method", Beijing：Science and Technology Press, 1986.

A New RBF Neural Network Training Algorithm Based on PSO*

Dingxue Zhang, Xinzhi Liu, Zhihong Guan
Control science and technology Department, Huazhong University of Science & Technology
Wuhan, Hubei, 430074, China
Email: zdx7773@163.com

ABSTRACT

Based on the study of Radial Basis Function (RBF) neural network training algorithm and Particle Swarm Optimization (PSO) algorithm, a new RBF neural network training algorithm with modified PSO algorithm is formulated, in which a control gene is introduced into basis PSO algorithm. The algorithm can determine network structure and parameters, such as centers and widths of hidden units by combining with least square method. The new training algorithm is applied to the nonlinear system identification problem, comparing with hierachical genetic algorithm and orthogonal least squares algorithm (OLS), the simulation results illustrate its efficiency.

Keywords: RBF neural network, PSO algorithm, least square method, nonlinear system identification.

1. INTRODUCTION

Radial Basis Function (RBF) neural network is a type of feed-forward artificial neural network, its performances are better than other types of networks, such as simpler network structures, better approximation capabilities, faster learning algorithm, and so on [1,2]. It has been used in pattern recognition, function approximation, prediction of nonlinear times series and nonlinear system identification.

The performance of RBF neural network depends on the number and centers of hidden units [3]. On the choice of hidden units, there are some traditional training methods, for example, K-means clustering, orthogonal least squares algorithm (OLS) [4], and so on, but they require to fix the number of hidden units and then proceed with the calculation of model parameters, or select from training data as centers of hidden units. For improving generalization performance, genetic algorithm (GA) is used in optimizing RBF neural network [5]. In the method, a chromosome consists of connection weights, centers and widths of hidden units and parameters of network, and net parameters can be optimized by iterative processing. In particularly, hierachical genetic algorithm (HGA) can optimize not only network parameters but also network structure [6]. But GA have some defects such as more predefined parameters, more intensive programming burden etc.

Particle Swarm Optimizer (PSO) algorithm is a recently proposed optimization algorithm, motivated by social behavior of bird flocking [7,8]. It has exhibited good performance not only on global search but also on local search by adjusting parameters. With no cross and mutation operation, it is easy to adjust parameters, the

*This work was supported by the National Natural Science Foundation of China under Grant 60573005.

method is fitted to program. It is demonstrated that PSO gets better results in a faster, cheaper way than GA, and can avoid degeneration in absolute optimization. In the paper, we propose a new method to train RBF neural network based on novel PSO. In the method, each individual particle makes up of many parameter units and employ control unit to determine states of parameter units. It can optimize centers and widths of hidden units and structure of RBF network, and good effects can be obtained with least square method.

This paper is organized as follows. Description of RBF neural network is given in section 2. Review of standard PSO is provided in section 3. In section 4, a training algorithm of RBF neural network based on Novel PSO is formulated. Next, experimental setting and results are shown in section 5. Finally, conclusions are given the paper in section 6.

2. RBF NEURAL NETWORK

RBF neural network consists of input layer, hidden layer and output layer, it can complete a nonlinear mapping, as follows

$$f_n(X) = w_0 + \sum_{i=1}^{n} w_i \phi(\|X - c_i\|)$$

Where, $X \in R_n$ and is input vector, $\phi(\cdot): R^+ \to R$ is a nonlinear function, the Gaussian function is used in our work.

$$\phi(\|X - c_i\|) = \exp(-\frac{\|X - c_i\|^2}{\sigma_i^2})$$

Where, w_i is connection weights, c_i is centers of hidden unit, σ_i is widths of hidden unit and n is number of hidden unit.

The pending parameters of RBF neural network have two types, one is center, widths and number of hidden unit, and the other is connection weight. If centers, widths and number of hidden unit are determined, least square method can be used to compute the weights that connect the hidden layer with output layer. So it is the key to determine centers, widths and number of hidden unit.

3. STANDARD PARTICLE SWARM OPTIMIZATION ALGORITHM

PSO algorithm was first introduced by James Kennedy and Russel C. Eberrhart in 1995, and it was discovered through simulation of a simplified social model. It can search the optimization result through cooperation and competition among the individuals in a population of particles. Suppose that the search space is N-dimensional

and a particle swarm consists of n particles, then the i-th particle of the swarm can be represented by a N-dimensional vector, $x_i = (x_{i1}, x_{i2}, \cdots, x_{iN})$ (i=1,2, ⋯ ,n). The velocity of this particle can be represented by another N-dimensional vector, $v_i = (v_{i1}, v_{i2}, \cdots, v_{iN})$ (i=1,2, ⋯,n). Every particle has a fitness that is determined by the objective function of optimization problem, and knows the best previously visited position ($pbest$) and present position (x_i) of itself. Every particle has knowing the position of the best individual of the whole swarm ($gbest$). Then the velocity of particle and its new position will be assigned according to the following equation (1) and (2).

$$v_{id}^{k+1} = w \times v_{id}^{k} + \eta_1 \times rand() \times (p_{id} - x_{id}^{k}) + \eta_2 \times rand() \times (p_{gd} - x_{id}^{k}) \quad (1)$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \quad (2)$$

Where v_{id}^{k} represents the velocity in the d-th dimension of the i-th particle at time k. x_{id}^{k} represents the position in the d-th dimension of the i-th particle at time k. p_{gd} is the best global position in the d-th dimension of the whole swarm. p_{id} is the best individual position in the d-th dimension of the i-th particle. $rand()$ is random number uniformly from the interval [0,1]. η_1 and η_2 are two positive constants called acceleration coefficients. w is called inertia weight and can be determined by equation (3).

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{Num_{\max}} \times Num \quad (3)$$

Where $w_{\max}$ and $w_{\min}$ are separately maximum and minimum of w. $Num_{\max}$ is maximum iteration time. Num is present iterative time.

4. RBF NEURAL NETWORK TRAINING ALGORITHM BASED ON PSO

4.1 Novel of standard PSO

For RBF neural networks implementation, a specially designed PSO algorithm is used to the optimum number, centers and widths of hidden units. In standard PSO algorithm, every particle is regard as many parameter units, each parameter unit represents hidden unit parameter of RBF neural network, and control unit is introduced to determine status of every parameter unit, the structure of particle is depicted in Fig 1.

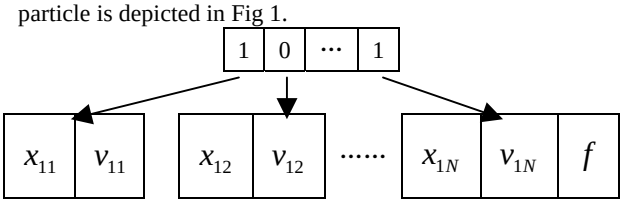


Fig.1. Structure of particle

Control unit is coded with binary, "1" represents activation status of parameter unit that is valid. Oppositely, "0" represents dormancy status that is invalid. Then the velocity of particle and its new position will be assigned according to the following equation (4) and (5).

$$v_{id}^{k+1} = C_{gd}(w \times v_{id}^{k} + \eta_1 \times rand() \times (p_{id} - x_{id}^{k}) + \eta_2 \times rand() \times (p_{gd} - x_{id}^{k})) \quad (4)$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \quad (5)$$

Where C_{gd} is the d-th code of the best individual of the whole swarm. Mutation can be used to change control unit, randomly select an individual by a probability and change the value in a random position.

4.2 Encoding and fitness function

In PSO algorithm, a particle is a feasible result, so parameter unit of particle must include centers and widths of RBF neural network hidden unit, velocity and fitness value. The object of neural network training is to optimize structure and network parameters, and to minimize the mean square error (MSE) between the desired outputs and the network predictions, so the reciprocal of mean square error can be regard as fitness function. The fitness value of the i-th individual is depicted as following.

$$f_i = 1/R_i \quad R_i = \frac{1}{M}\sum_{k=1}^{M}(y_k - \hat{y}_k)^2$$

Where y_k is the desired output, $\hat{y}_k$ is the network prediction and M is the number of training sample.

4.3 Steps of the algorithm

The algorithm can be summarized in the following steps.
1). Select training sample.
2). Initialize parameters of PSO.
3). Initialize the state of each particle. Store the best individual position of each particle. Evaluate and store best individual fitness of each particle, the best global position and best global fitness of the whole swarm.
4). Update the velocity of particle according to equation Update the position of particle according to equation (5).
5). If necessary, update and store the individual's best position and best fitness of each particle, and the best global position and fitness of whole swarm.
6). If the stopping condition is not satisfied, go to step 4. Otherwise, stop iterating and obtain the result from the best global position and fitness of the whole swarm.
7). Decode and train RBF neural network.

5. EXPERIMENT

In the section, simulation clearly demonstrates the ability of the RBF neural network trained by PSO to learn the dynamics of the unknown system. The system to be identified is described by equation (6).

$$y(k+1) = \frac{y(k)}{1+y^2(k)} + [u(k)+1]u(k)[u(k)-1] \quad (6)$$

if a series-parallel identification model is used for identifying the nonlinear system, the model can be described by the equation (7).

$$\hat{y}(k+1) = f(y(k), u(k)) \quad (7)$$

Where f is an RBF neural network trained by PSO with two inputs and one output.

For comparison purpose, RBF neural network training algorithm based on HGA is used to the nonlinear system identification problem. The simulation produced 100 data points by exciting the system with random signals uniformly distributed between –1 and 1.

In PSO algorithm, $\eta_1 = \eta_2 = 2.05$, $w_{\max} = 0.905$, $w_{\min} = 0.4$, population size is 20. In HGA, population size is 20, and crossover probability is 0.8, and mutation probability is 0.1. To the two algorithms, the maximum hidden unit number of RBF neural network is 30, and the maximum generation is 20.Table 1 illustrates the performance of RBF neural network based on two methods. The result shows that the MSE of training using OLS is the biggest. The evolving process of hidden unit number to the best network structure at each epoch is shown in Fig.2.

Table 1. Comparison of mean squared errors (MSE) and hidden unit number of RBF neural network constructed by PSO, HGA and OLS training algorithm

Training algorithm	Hidden units number	MSE of training
PSO	25	4.0×10^{-4}
HGA	23	7.813×10^{-4}
OLS	32	6.198×10^{-3}

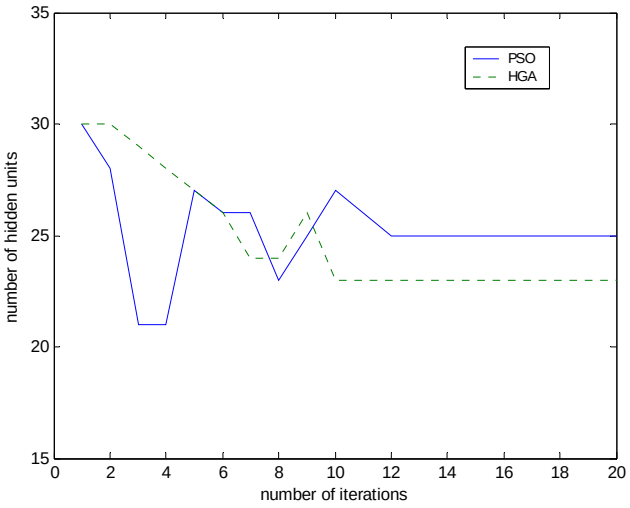


Fig.2. Evolving process of hidden unit number to the best network structure at each epoch

To test the result of identification, the input signal $u(k) = 0.25[\sin(2k\pi/25) + \sin(2k\pi/10)]$ has been applied to both the system and the model and generated 100 data. Fig. 3 shows the result of identification obtained from the RBF network based on PSO and HGA. MSE of testing is 3.897×10^{-5} based on PSO algorithm and is 1.25×10^{-4} based on HGA. The error of identification result obtained from the RBF network based PSO and HGA is shown in Fig.4. The interval of error based on PSO is from –0.0181 to 0.0119, however, the interval of error based on HGA is from –0.0277 to 0.0256. The mean relative error based PSO is 1.00%, and on HGA-1.38%. The result based on PSO is obviously better than based on HGA. The RBF neural network based on PSO algorithm has not only generated the most parsimonious structure but also provided better accurate outcome than HGA. As far as training time, based on PSO is only a third of based on HGA.

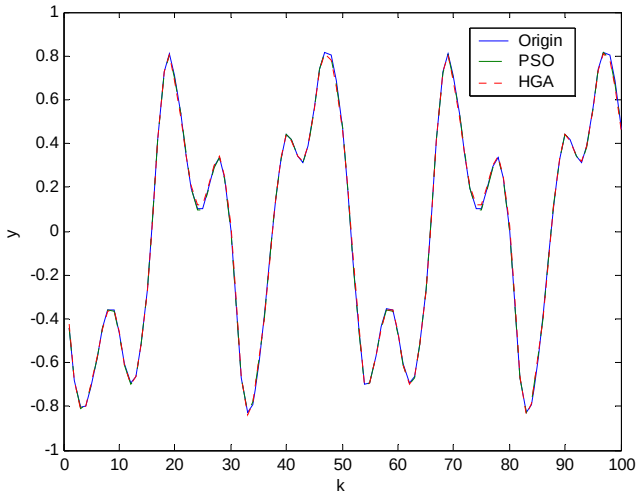


Fig.3. The result of identification obtained from the RBF neural network based on PSO and HGA

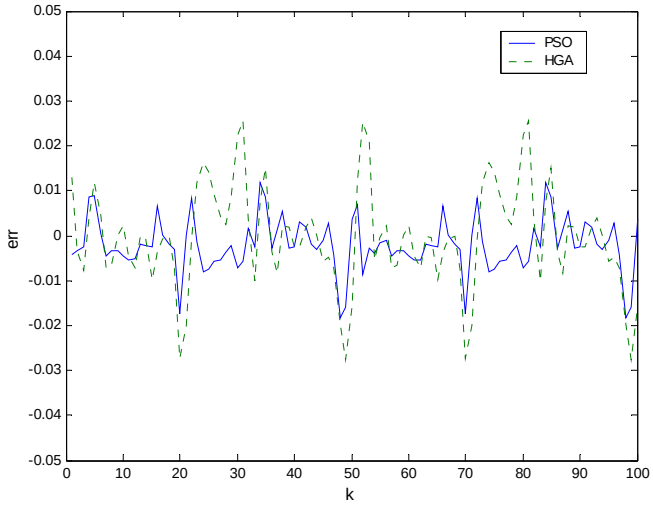


Fig. 4. Error of identification result obtained from the RBF network based on PSO and HGA

6. CONCLUSIONS

In the paper, a new RBF neural network training algorithm with novel PSO algorithm is introduced, based on the study of RBF neural network training algorithm and standard PSO algorithm. In the method, control gene is introduced into PSO algorithm, the algorithm can determine network structure and parameters of hidden units, combining with least square method, can training RBF neural network. The new method is used to nonlinear system identification, comparing with RBF network training algorithm based on HGA, the simulation results illustrate that the method has not only generated the most parsimonious structure but also provided the accurate outcome better than HGA. Training time based on PSO is only a third of based on HGA.

7. REFERENCES

[1] M. Binchini, P. Frasconi, et al, Learning without local minima in Radial Basis Function networks. IEEE. Transactions on Neural Net-works, 1995, 6(3): 749-755.

[2] J. Chai, Q. Y. Jiang, et al, Function approximation capabilit and algorithms of RBF neural networks. Pattern Recognition & Artificial Intelligence, 2002, 15(3); 310-316.

[3] J.Y. Cao, G.P. Lu, et al, Research on the generalization of radial basic function neural network and its application [J]. Aerospace Electronic Information Engineering and Control. 2006,28(1): 72-74.

[4] W.J. Liu, J. Guo, Study on the OLS algorithm of RBF neural network [J]. Journal of Tianjin Polytechnic University. 2002,21(2): 71-73.

[5] B.A. Whitehead, Genetic evolution of radial basis function coverage using orthogonal niches. IEEE Trans on Neural Networks. 1996. 7(6): 869-880.

[6] P.E. Zheng, Y.H. Ma, A New Hierarchical Genetic Algorithm for Training of RBF Neural Networks [J] Control and Decision.2000, 15(2): 165-168.

[7] J. Kennedy, R. Eberhart, Particle swarm optimization, IEEE Int. Conf. Neural Networks, 1995:1942-1948.

[8] Y. Shi, R. Eberhart, modified particle swarm optimizer[C]:IEEE World Congress on Computational Intelligence, 1998,69-73.

A Novel Fuzzy Cognitive Maps Learning Algorithm

ZHAO JING[1,2] , XU WENBO[1], SUN JUN[1], LI MING[3]
[1] School of Information Technology Southern Yangtze University
Wuxi, Jiangsu, P. R. China, 214122
[2]Center of Modern Education Technology, Shandong Institute of Light Industry
Jinan, Shandong, P. R. China, 250100
[3]Institute of Technology, Shandong University of Traditional Chinese Medicine
Jinan, Shandong, P. R. China, 250014
Email: lmzjmail@163.com , xwb@sytu.edu.cn , sunjun_wx@hotmail.com

ABSTRACT

A technique for Fuzzy Cognitive Maps learning, which is based on the Quantum-behaved Particle Swarm Optimization algorithm, is introduced. The proposed approach is used for updating the nonzero weight values that lead the Fuzzy Cognitive Map to desired steady states. The workings of the approach are applied to an industrial control problem. The results support the claim that the proposed technique is a promising methodology for Fuzzy Cognitive Maps learning, and the methodology is effective and efficient.

Keywords: PSO algorithm, QPSO algorithm, fuzzy cognitive maps, weight matrices, objective function.

1. INTRODUCTION

The representation of causal relationships and reasoning are important research fields of AI (artificial intelligence). Political scientist Robert Axelrod originally proposed Cognitive Maps in 1976 [1]. Cognitive Maps (CM)[1] are a useful model to represent concepts or variables in a given domain and their causal-effect relations. The concepts are represented as *nodes*, and the causal relationships between these concepts are represented as *edges*. Kosko [2] enhanced the power of cognitive maps considering fuzzy values for concepts of the cognitive map and fuzzy degrees of interrelationships between concepts. Then Fuzzy Cognitive Maps (FCM) were introduced in 1986 as signed directed graphs for representing causal reasoning and computational inference processing.

FCMs have been applied in various fields, including modeling of complex [3], decision analysis [4], management science, operations research and organizational behavior [5].

FCMs have the character of abstraction, flexibility and fuzzy reasoning, and FCMs promote the research on new concepts in complex systems. However, the existing learning methods need further enhancement, stronger mathematical justification, and further testing on systems of higher complexity. Moreover, the elimination of deficiencies, such as the abstract estimation of the initial weight matrix and the dependence on the subjective reasoning of experts' knowledge, will significantly improve the performance of FCMs. In this context, the improvement of learning algorithms is a primary research topic.

Up to date, a few learning algorithms have been proposed [6,7]. The most task of learning procedure is to find a set of the FCM's weights, that leads the FCM to a desired steady state, this can be achieved through the minimization of a properly defined objective function, they are mostly based on ideas coming from the field of neural network training. The existing algorithms are mostly dependent on the initial weight matrix approximation, which is provided by the experts.

In this paper, an approach for FCMs learning, based on QPSO (Quantum-behaved Particle Swarm Optimization) algorithm, is presented. QPSO is applied to update the weight values of the FCM, and determined the proper weight matrices for the system, so that leads the FCM to a desired steady state.

2. PARTICLE SWARM OPTIMIZATION ALGORITHM AND QUANTUM-BEHAVED PARTICLE SWARM OPTIMIZATION ALGORITHM

2.1 PARTICLE SWARM OPTIMIZATION (PSO)

Particle swarm optimization (PSO) is an evolutionary computation technique developed by Dr. Eberhart and Dr. Kennedy in 1995[8], inspired by social behavior of bird flocking or fish schooling.

PSO is similar to the other evolutionary algorithms in that the system is initialized with a population of random solutions. However, each potential solution, call particles, flies in the D-dimensional problem space with a velocity which is dynamically adjusted according to the flying experiences of its own and its colleagues. The location of the i-th particle is represented as $X_i=(x_{i1}, x_{i2}, \ldots,x_{iD})^T$. The best previous position of the i-th particle is denoted as $P_i=(pi1, pi2 \ldots piD)^T$, which is also called pbest. The index of the best particle among all the particles in the population is represented by the symbol g_i. The location Pg_i is also called gbest. The velocity for the i-th particle is represented as $V_i=(v_{i1}, v_{i2}, \ldots, v_{iD})^T$. Then the swarm is manipulated by the equation [9]:

$$V_i(t+1)=\omega V_i(t)+c_1 r_1\left(P_i(t)-X_i(t)\right)+c_2 r_2\left(P_{g_i}(t)-X_i(t)\right) \quad (3)$$

$$X_i(t+1)=X_i(t)+V_i(t+1) \quad (4)$$

where ω is inertia weight [9], c1 and c2 are acceleration constants [10], r1 and r2 are a random function in the range [0, 1]. Generally, the default values $c1 = c2 = 2$ have been proposed, and the parameter ω decreases linearly from 0.9 to 0.4[11].

2.2 QUANTUM-BEHAVED PARTICLE SWARM OPTIMIZATION ALGORITHM（QPSO）

In PSO algorithm, the i-th particle is depicted by its position vector X_i and velocity vector V_i, which determine the trajectory of the particle. The particle moves along a determined trajectory in classical mechanics, but this is not the case in quantum mechanics. In quantum world, the term *trajectory* is meaningless, because X_i and V_i of a particle cannot be determined simultaneously according to *uncertainty principle*. Therefore, if individual particles in a PSO system have quantum behavior, the PSO algorithm is bound to work in a different fashion. In [12,13], Jun Sun *et al.* introduce a Quantum-behaved PSO (QPSO) algorithm. The experiment results indicate that the QPSO works better than standard PSO on several benchmark functions.

In QPSO algorithm, only position vector *p* is needed to depict a particle, and there is only one parameter β. The equation is as follows:

$$p = \frac{\varphi_1 * p_i + \varphi_2 * p_g}{\varphi_1 + \varphi_2} \tag{5}$$

$$mbest = \frac{\sum_{i=1}^{M} P_i}{M} = \left(\frac{\sum_{i=1}^{M} P_{i1}}{M}, \frac{\sum_{i=1}^{M} P_{i2}}{M}, \cdots, \frac{\sum_{i=1}^{M} P_{id}}{M} \right) \tag{6}$$

$$x(t+1) = p \pm \beta * \left| mbest - x(t) \right| * \ln\left(\frac{1}{u}\right) \tag{7}$$

where φ_1 and φ_2 are a random function in the range [0, 1], p_i is the best position of particle i, p_g is the position of the best particle among all the particles in the population. The global point, denoted as *mbest*, is defined as the mean of pbest positions of all particles. β is called Creativity Coefficient, M is the population size and u is a random function in the range [0,1].In the process of iteration, ± is decided by the random number, when it is bigger than 0.5, minus sign (－) is proposed, others plus sign (＋) is proposed.

3. FUZZY COGNITIVE MAPS

3.1 FUZZY COGNITIVE MAPS

Fuzzy Cognitive Maps (FCM) are soft computing tools. Concepts are pictured different aspects of the system and their behavior, and the dynamics of the system are represented by the interaction of concepts [2]. An FCM models consists of nodes–concepts, C_i (i=1,2…N), where *N* is the total number of concepts. Each node–concept represents one primary factor of the system and it is depicted by a value $A_i \in [0,1]$(i=1,2…N). The weight, W_{ij}, indicates whether the relation between the two concepts is positive or negative. The direction of causality demonstrates whether the concept C_i causes the concept C_j or reverse. Thus, the values of weights are in continuum [-1, 1].

The value A_i, of a concept C_i, is obtained by the transformation of the fuzzy values to numerical values. After the assignation of values to the concepts and the weights, the FCM is let to converge to a steady state through the interaction subsequently described.

At each step, the value A_i of a concept C_i is influenced by the values of concepts–nodes connected to it, and is updated according to the equation [6]:

$$A_i(k+1) = f\left(A_i(k) + \sum_{j=1}^{n} W_{ji} A_j(k) \right) \tag{1}$$

where *k* stands for the iteration counter; and W_{ji} is the weight of the arc connecting concept C_j to concept C_i. The function *f* is the sigmoid function:

$$f(x) = \frac{1}{1 + e^{-\lambda x}} \tag{2}$$

Where λ > 0. In the present study the value of was set to 1. This function can restrict the values A_i of the concepts within [0,1]. The interaction of the FCM results after a few iterations in a steady state, i.e. the values of the concepts are not modified further.

After the determination of FCM's structure, and using the initial concept values A_i and the matrix initial $W_{initial}$, which are provided by the experts, then the FCM is let to converge to a steady state through the application of Eq. (1).

The heavy dependence on the experts' opinion regarding the FCM's design and the convergence to undesired steady states starting from the experts' recommendations are the two most significant weaknesses of FCMs. However, we apply the new technique to update the weight matrix of FCM so as to avoid convergence to undesired steady states.

3.2 THE NEW LEARNING APPROACH

The purpose of the new approach is to search a proper weight matrix W=[W_{ij}](i,j=1,2,…,N), and then the FCM is let to converge to a steady state.

Set C_1, …, C_N, be the concepts of an FCM, and set C_{out_1}, …, C_{out_m} (1≤m≤N) be the output concepts, while the remaining concepts are considered input or interior concepts. The output concepts keep in strict bounds:

$$A_{out_i}^{\min} \le A_{out_i} \le A_{out_i}^{\max}, \quad i=1,2,\ldots,m$$

Thus, the objective function is considered [14]:

$$F(W) = \sum_{i=1}^{m} H\left(A_{out_i}^{\min} - A_{out_i}\right)\left|A_{out_i}^{\min} - A_{out_i}\right| + \sum_{i=1}^{m} H\left(A_{out_i} - A_{out_i}^{\max}\right)\left|A_{out_i}^{\max} - A_{out_i}\right| \tag{5}$$

Where H is the famous Heaviside function

$$H(x) = \begin{cases} 0, x < 0, \\ 1, x \ge 0, \end{cases}$$

A_{out_i} (i=1,2,…,m) are the steady state values of the output concepts, which are obtained through the application of Eq. (1). Obviously, the global minimization of the objective function *F* is weight matrices that lead the FCM to a desired steady state.

The application of QPSO for the minimization of the objective function *F*, starts with an initialization state, where a swarm of M particles is generated randomly, and it is evaluated using *F*. Then, Eq. (5), Eq. (6) and Eq. (7) are used to evolve the swarm. When a weight configuration that globally minimizes *F* is reached, the algorithm stops.

4. AN INDUSTRIAL PROCESS CONTROL PROBLEM

The proposed learning algorithm previously described, is applied on a complex industrial process control problem. This problem consists of three tanks, five valves, one heating element, one sensor and two thermometers, as illustrated in Fig. 1. Valve 1 and Valve 2 pour two different liquids into the tank 1. During the mixing of the two liquids, a chemical reaction takes place in the tank, and a new liquid is produced. When the new liquid produced reaches a specific level, valve 3 opens, and the liquid comes into tank 2 from tank 1. A sensor is placed inside the tank 1 to measure the specific gravity of the produced liquid. The temperature of the liquid in tank 2 is regulated through a heating element, the temperature of the liquid in tank 3 is measured through a thermometer, when the temperature of the liquid 3 decreases, this cause the valve 4 to open, so hot liquid comes into tank 3.

The objective of the control system are firstly to keep the height of liquid, in all tanks, between some limits, an upper limit H_{max} and a low limit H_{min} ; secondly the temperature of the liquid in tank 2 and tank 3 must be kept between a maximum value T_{max} and a minimum value T_{min} ; thirdly the specific gravity of the liquid in tank 1 should be kept an upper limit G_{max} and a low limit G_{min}.

The control objective is to keep values of these variables in the following range of values:

$$G_{min} \leqslant G \leqslant G_{max}$$
$$H^1_{min} \leqslant H^1 \leqslant H^1_{max}$$
$$H^2_{min} \leqslant H^2 \leqslant H^2_{max}$$
$$H^3_{min} \leqslant H^3 \leqslant H^3_{max}$$
$$T^1_{min} \leqslant T^1 \leqslant T^1_{max}$$
$$T^2_{min} \leqslant T^2 \leqslant T^2_{max}$$

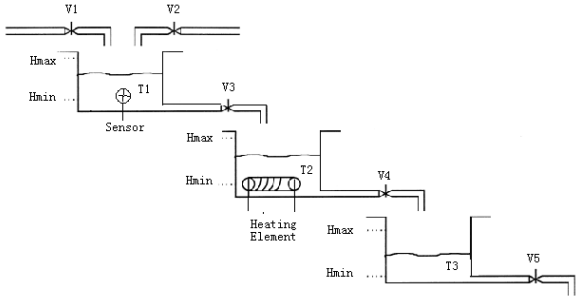


Fig. 1. The illustration of a process problem

Fuzzy cognitive map that models and controls this system is depicted on Fig. 2. It consists of twelve concepts that are defined as:

1)Concept 1——the amount of the liquid in tank 1.It depends on the operational state of Valve 1, 2, and 3;

2)Concept 2——the amount of the liquid in tank 2.It depends on the operational state of Valve 3 and 4;

3)Concept 3——the amount of the liquid in tank 3.It depends on the operational state of Valve 4 and 5;

4)Concept 4——the state of Valve 1(closed, open or partially opened);

5)Concept 5——the state of Valve 2(closed, open or partially opened);

6)Concept 6——the state of Valve 3(closed, open or partially opened);

7)Concept 7——the state of Valve 4(closed, open or partially opened);

8)Concept 8——the state of Valve 5(closed, open or partially opened);

9)Concept 9——the specific gravity of the liquid in the tank 1;

10)Concept 10——the temperature of the liquid in tank 2;

11)Concept 11——the temperature of the liquid in tank 3;

12)Concept 12——the operation of the heating element in tank 2.

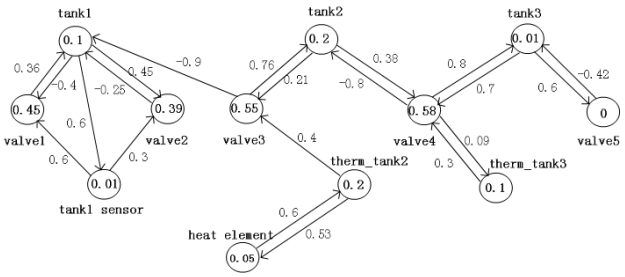


Fig. 2. The FCM that corresponds to the problem of Fig. 1.

Experts determine the direction of the arcs and the weight value among the concepts. The ranges of the weights implied by the fuzzy regions are:

$-0.5 \leqslant W_{14} \leqslant -0.3$	$-0.4 \leqslant W_{15} \leqslant -0.2$	$0.4 \leqslant W_{19} \leqslant 0.7$
$0 \leqslant W_{26} \leqslant 0.5$	$0 \leqslant W_{27} \leqslant 0.5$	$0.5 \leqslant W_{37} \leqslant 0.75$
$0.25 \leqslant W_{38} \leqslant 0.75$	$0.3 \leqslant W_{41} \leqslant 0.4$	$0.4 \leqslant W_{51} \leqslant 0.5$
$-1.0 \leqslant W_{61} \leqslant -0.8$	$0.5 \leqslant W_{62} \leqslant 1.0$	$-1.0 \leqslant W_{72} \leqslant -0.5$
$0 \leqslant W_{73} \leqslant 0.75$	$0 \leqslant W_{7,11} \leqslant 0.25$	$-0.75 \leqslant W_{83} \leqslant 0$
$0.5 \leqslant W_{94} \leqslant 0.7$	$0 \leqslant W_{95} \leqslant 0.8$	$0 \leqslant W_{10,6} \leqslant 0.5$
$0.25 \leqslant W_{10,12} \leqslant 0.7$	$0 \leqslant W_{11,7} \leqslant 0.5$	$0.25 \leqslant W_{12,10} \leqslant 0.7$

the nonzero weight values of initial weight matrix is:

$W^{initial}$= [-0.4, -0.25, 0.6, 0.21, 0.38, 0.7, 0.6, 0.36, 0.45, -0.9, 0.76, -0.8, 0.8, 0.09, -0.42, 0.6, 0.3, 0.4, 0.53, 0.3, 0.6]

QPSO is applied to update the twenty-one nonzero weight values of the FCMs. The bounds [-1,0] or [0,1], implied by the directions of the corresponding arcs of the FCM, are imposed on each weight.

The output concepts for this problem are the concepts C_1 , C_2 , C_3 , C_9 , C_{10} , C_{11}. The experts have defined the desired regions for these concepts:

$$0.68 \leqslant C_1 \leqslant 0.7 \qquad 0.55 \leqslant C_2 \leqslant 0.75$$
$$0.75 \leqslant C_3 \leqslant 0.8 \qquad 0.78 \leqslant C_9 \leqslant 0.85$$
$$0.65 \leqslant C_{10} \leqslant 0.7 \qquad 0.65 \leqslant C_{11} \leqslant 0.7$$

5. SIMULATION RESULTS

In our experiments, a total of 100 independent experiments have been performed using QPSO algorithm, and the obtained results compare with the results of using PSO algorithm.

The parameters of PSO algorithm are defined as follows: the swarm size is set to 30, the default values $c1 = c2 = 2$ have been proposed, the parameter ω decreases linearly from 0.9 to 0.4, the number of iterations required is 40. The parameters of QPSO algorithm are defined as follows: the swarm size is set to 30, the parameter β decreases linearly from 1.0 to 0.5, the number of iterations required

is 40.The accuracy for the determination of the minimized objective function has been equal to 10^{-12}.
One of the obtained sub-optimal matrices using PSO algorithm is the following:
W=[-0.3, -0.26, 0.7, 0, 0.5, 0.54, 0.75, 0.4, 0.5, -0.8, 0.5, -0.5, 0.51, 0.06, 0, 0.7, 0.8, 0.09, 0.7, 0.5, 0.25]
Which leads the FCM to the desired steady state:
C_1=0.6988, C_2=0.6297, C_3=0.7713, C_4=0.7473,
C_5=0. 7707, C_6=0.6737, C_7=0.8743, C_8=0.7985,
C_9=0.7807, C_{10}=0.7127, C_{11}=0.6738, C_{12}=0.7872.
However, one of the obtained sub-optimal matrices using QPSO algorithm is the following:
W=[-0.44, -0.4, 0.71, 0, 0, 0.54, 0.73, 0.4, 0.49, -0.8, 0.81, -0.57, 0.74, 0.11, -0.1, 0.71, 0.61, 0.01, 0.25, 0.5, 0.25]
Which leads the FCM to the desired steady state:
C_1=0.6909, C_2=0.6749, C_3=0.7921, C_4=0.7275,
C_5=0. 7152, C_6=0.6614, C_7=0.8443, C_8=0.7986,
C_9=0.7820, C_{10}=0.7485, C_{11}=0.6855, C_{12}=0.7105.

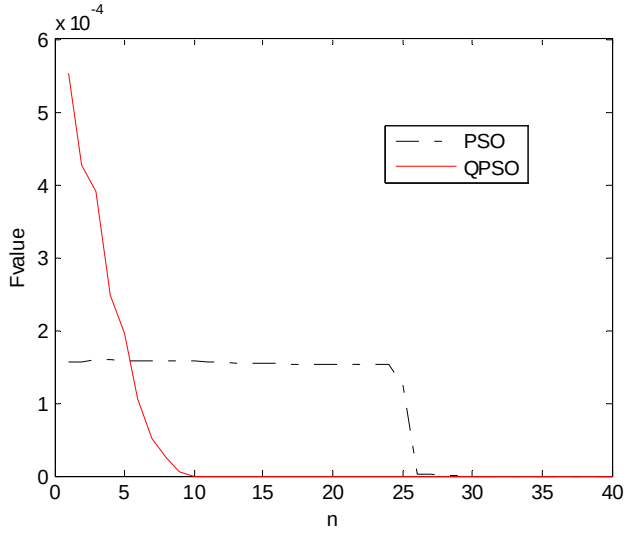


Fig. 3. The convergence graph of objective function F

The convergent process of objective function F is illustrated in Fig. 3. We can obtained the conclusion from the result that the convergent velocity of QPSO algorithm is much more faster than that of PSO algorithm. The obtained weight matrixes lie in experts' bounding regions. It is clear that the learning algorithm is capable of providing proper weight matrices for the FCM, and alleviating deficiencies caused by deviation in the experts' suggestions.

6. CONCLUSION

QPSO algorithm is the expanding of the standard PSO algorithm in quantum space, and has the virtues of briefness and less parameter. The experiment results indicate that the QPSO works better than PSO on an industrial process control problem and QPSO is a promising algorithm. Future work is directed towards more investigation the new approach on the applications of different scenarios of the employed industrial problem as well as applications on systems of higher complexity.

7. REFERENCES

[1] D.B. Chaib, J. desharnais, A relational model of cognitive maps [DB/OL]. http://citeseer.nj.nec.com/. 2001,12.
[2] B. Kosko, Fuzzy cognitive maps. Int. J. Man–Machine Studies, 24:65–75, 1986.
[3] J.S. Jang, C.T. Sun, et al, Neuro–Fuzzy and Soft Computing. Prentice Hall, Upper Saddle River, NJ, 1997.
[4] S. Khan, A. Chong, et al, Methodology for developing adaptive fuzzy cognitive maps for decision support. In Proc. 3rd Australia–Japan Evolutionary Systems, pp. 93–100, Canberra, 1999.
[5] P. Craiger, M.D. Coovert, Modelling dynamic social and psychological processes with fuzzy cognitive maps. In Proc. 3rd IEEE Conf. Fuzzy Systems, 1994.
[6] B. Kosko, Fuzzy Engineering. Prentice Hall, New York, 1997.
[7] E.I. Papageorgiou, C.D. Stylios, et al, "Activation hebbian learning rule for fuzzy cognitive maps". In Proc. 15th IFAC World Congress, Barcelona, Spain, 2002.
[8] J. Kenned, R.C. Eberhart, "Particle swarm optimization. In Proceedings IEEE International Conference on Neural Networks", volume IV, pages 1942–1948, Piscataway, NJ, 1995. IEEE Service Center.
[9] Y. Shi, R.C. Eberhart, "A modified particle swarm optimizer. In Proceedings IEEE Conference on Evolutionary Computation", Anchorage, AK, 1998. IEEE Service Center.
[10] R. Eberhart, Y. Shi, "Particle swarm optimization: developments, applications and resources," Proc. IEEE Int. Conf. on Evolutionary Computation, pp. 81-86, 2001.
[11] K.E. Parsopoulos, M.N. Vrahatis, Recent approaches to global optimization problems through particle swarm optimization. Natural Computing, 1(2–3): 235–306, 2002.
[12] J. Sun, B. Feng, et al. Particle Swarm Optimization with Particles Having Quantum Behavior[C].In:IEEE Proc of Congress on Evolutionary Computation,2004
[13] J. Sun, W.B. Xu, "Global Search Strategy of Quantum-behaved Particle Swarm Optimization". Proceedings of IEEE conference on Cybernetics and Intelligent Systems. (2004) 111 – 116.
[14] K.E. Parsopoulos, E.I. Papageorgiou, Groumpos, P.P., Vrahatis, M.N.: A first study of fuzzy cognitive maps learning using particle swarm optimization. In:Proc. IEEE 2003 Cong. Evol. Comput., (2003) 1440-1447.

Aggregating Linguistic Information by Case Learning

Mingrong Deng
School of Management, Zhejiang University
Hangzhou, Zhejiang, 310058, China
Email: dengmr@yahoo.com.cn

ABSTRACT

For linguistic information aggregation, both the inputs and outputs are expressed by linguistic variables. A linguistic term implies less information than a precise number does, but can tell more information than a nominal or ordinal number can. Using some aggregation operators, the mechanism aggregating linguistic information can be represented as a set of weights. Other than the methodologies of using weights given in advance, the approach proposed in this paper attempts to estimating the weights from a set of cases with known aggregation results. These weights are different from those in an artificial neural network. They have explainable meaning, and hence more like the explicit knowledge.

Keywords: Knowledge Discovery, Linguistic Information, Weight Estimation, Aggregation, and Decision Support System.

1. INTRODUCTION

The knowledge for aggregating criteria functions to form overall decision function is important. In the simplest cases such as both the criterion functions and overall function present as numerical scores, the knowledge can consist in determining the criterion weights, by which one can get overall grade through weighted averaging of all criteria grades. In the more complex situation, the approach of artificial neural network (ANN) provides an effective way, since it has been shown that there exists a three-layers network can approximate a continuous function for any required toleration. An artificial neural network can be trained to learn the processing by a set of instances with known results. The knowledge is embodied in the linkage weights between neurons [15]. However, one problem of a neural network roots in its explanation capability. The knowledge cannot only be used internally for reasoning and learning, but also externally for the explanation of results to a user. The full potential of trained artificial neural network may not be realized due to its limitation of explanation capability [1]. Some complementary techniques such as extracting rules from ANN have been researched, for instance, in [1, 10].

Uncertainties in inputs and outputs bring more complexity. When the uncertainty is due to ignorance, aggregating mechanisms can be constituted based on the Evidence theory of Dempster and Shafer [18]. Yager has developed a general framework for information fusing where multiple sources give their certain information, which maybe conflict each other [24]. Deng et al. proposed an evidential combination formula and developed an attribute weight estimating method in the circumstances each source information can be uncertain [5].

The best representation of uncertainty is the one that is able to handle all relevant information available, and does not require unavailable information [17].

When the inputs and outputs are expressed by linguistic variables, another kind of uncertainty is involved. A linguistic term (e.g., when evaluating the “comfort” or “design” of a car, terms like “good”, “medium”, “bad” can be used) implies less information than a precise number does, but can tell more information than a nominal or ordinal number can. Information will be lost if the terms are replaced by nominal or ordinal numbers.

Fuzzy set theory, proposed and developed by Zadeh [25], is often used to deal with the situation where the assessment information is stated in linguistic terms and some membership functions can be defined to denote the degree of the truthfulness of the proposition. The aggregation of fuzzy opinions can be realized by weighting averaging [2,6,7].

There are some difficulties lying in the application of fuzzy set theory. One is the agreement of all individuals on the same membership function associated to linguistic terms [12], and another is the complexity of calculation on the fuzzy numbers. Improper methods of calculation may introduce additional uncertainties not present in the original problem Linguistic approximation is another problem for aggregating information using membership functions associated with the linguistic terms. The linguistic terms often correspond to regular membership functions, such as triangular, trapezoidal forms. After operations on fuzzy sets time and again, the shape of final membership function does not take normal form when the linguistic variables are interactive. If a definite result is desired, approximation is needed to finding a linguistic term whose meaning is the same or closest to the meaning of the fuzzy set generated by the linguistic aggregation model [3,19].

Because of the difficulties in defining consensus membership function and result approximation of the ultimate irregular fuzzy set, a direct operation method on linguistic variables has been proposed in [11,12,13]. The method is based on the concepts of ordered weighted averaging (OWA) operators, due to Yager [21,22,23], and convex combination of linguistic labels defined by [4]. However, using linguistic ordered weighted averaging (LOWA) operator means that all inputs are equally important for the output, and the weights, which determine the degree of “anding” and “oring” implicit in the aggregation, cannot be predefined easily.

In this paper, we extend the LOWA operation by allowing the inputs may be not equally important. The knowledge about the aggregation mechanism and importance of each input variable are expected to be learnable from a set of cases with some known outcome information.

The paper is structured as follows. In Section 2, the problem is briefly reviewed. The LOWA operator and the weights in different means are discussed in Section 3. In Section 4, we propose an operator allowing for two kinds of weights, and present how it can be used to aggregating linguistic information based on mining known behavior cases. An example is illustrated in Section 5. Finally, some discussions and conclusions are given in section 6.

2. PROBLEM DESCRIPTION

The problem is equivalent to a multi-criteria problem, where linguistic assessments with respect to individual criteria are to be aggregated to form an overall assessment.

Let $A_1, A_2, ..., A_n$ be *n* criteria in a multi-criteria problem. X is the set of alternatives. For a proposed alternative $x \in X$, $A_j(x)$ indicates the degree to which $x \in X$ satisfies the criterion A_j. Our problem is to find out the mechanism to get the overall assessment of x from the set of $\{A_j(x) \mid j = 1, \ldots, n\}$.

Many theories and techniques are devoted to the problem when the assessment information is expressed quantitatively. When uncertainty involved, fuzzy sets and fuzzy numbers are often employed. However, some exact numerical values are still needed in advance, for example, the values of *a, b, c* in a triangular fuzzy number [*a, b, c*]. Their determination is not indisputable in many cases.

A linguistic variable differs from a numerical one in that its values are not numbers, but words or sentences in a natural or artificial language. Since words, in general, are less precise than numbers, the concept of a linguistic variable serves the purpose of providing a means of approximated characterization of phenomena, which are too complex, or too ill-defined, to be amenable to their description in conventional quantitative terms [12].

Although linguistic terms can be viewed as fuzzy numbers defined in the [0,1] intervals, we are interested in how the information in qualitative aspect can be interpreted and utilized to the maximal extent, and the least unavailable information is required.

What we can make use of is a number of alternatives with known aggregated result. They are similar to a series of cases in case-base reasoning.

Noted that linguistic terms imply less information than precise numbers do, but can tell more information than nominal and ordinal numbers can, many techniques such as decision tree, rough set, or artificial neural network cannot directly apply [14].

3. LOWA OPERATOR AND WEIGHTS

We begin with the situation where $A_1, A_2, ..., A_n$ are equally important.

An aggregation operator OWA (ordered weighted averaging) F can be defined [21], if the domain of each A_j is [0,1], $j = 1,2,...n$, as follows.

Given $NW_i \in [0,1], \sum_{i=1}^{n} NW_i = 1$, and

$$(A_1(x), A_2(x), ..., A_n(x)) = (a_1, a_2, ..., a_n),$$

$$F(a_1, a_2, ..., a_n) = NW_1 b_1 + NW_2 b_2 + ... + NW_n b_n,$$

where b_i is the i th largest element in the collection $\{a_1, a_2, ..., a_n\}$.

The vector NW determines some aggregation structure. When $NW = (1,0,...,0)$, the aggregation is the least rigorous one, permitting the alternative is good as long as one of the criteria is good. F behaves as a pure "or" operator. At the other extreme $NW = (0,...,0,1)$, the alternative is desired to satisfy all the criteria. F becomes a pure "and" operator. An ordinary aggregation acts like a combination of "or" and "and". Note that one of the OWA operator's advantages is that it makes possible that deficiency of the alternative on a criterion cannot get compensated from the other criteria.

A measure of "orness"* associated the weight $W = (W_1, W_2, ..., W_n)$ can be defined as

$$orness(W) = (1/(n-1)) \sum_{i=1}^{n} ((n-i) W_i)$$

It is obvious that for $W = (1,0,...0)$, $orness(W) = 1$; for $W = (0,...0,1)$, $orness(W) = 0$; and for $W = (1/n, 1/n, ...1/n)$, we have $orness(W) = 1/2$.

Some weighting functions have the same degree of "orness", for instance, for $W = (0,...,0,1_{(n+1)/2},0,...0)$, where *n* is odd, *orness(W)* is 1/2. The difference is that the latter is more volatile and uses less of the input. Hence a measure of "dispersion" (entropy) associated with a weighting function can be employed:

$$dispersion(W) = -\sum_{i=1}^{n} W_i \ln W_i$$

In linguistic situation, $A_j(x)$ is not an accurate number in [0,1], but relates to an appropriate linguistic term set, $S = \{s_i \mid i \in \{0,...,T\}\}$, which is chosen and used to provide vague or imprecise performance value about alternatives according to the different criteria.

For example, S can be the set of $\{s_0 = None, s_1 = Very\_Low, s_2 = Low, s_3 = Medium, s_4 = High, s_5 = Very\_High, s_6 = Perfect\}$

If the term set has the following characteristics:

- ✧ The set is ordered: $s_i \geq s_j$ if $i \geq j$;
- ✧ There is the negation operator: $Neg(s_i) = s_j$ such that $j = T - i$

An aggregation operator, LOWA ϕ can be defined to compute directly on linguistic labels [12]. For the label set $\{a_1, a_2, ..., a_n\}$,

$$\phi(a_1, a_2, ..., a_n) = NW \cdot B^T = C^n\{NW_i, b_i, i = 1,..., n\}$$
$$= NW_1 \odot b_1 \oplus (1 - NW_1) \odot C^{n-1}\{\beta_h, b_h, h = 2,..., n\}$$

where, NW and $B = \{b_1, ..., b_n\}$ are as before, $\beta_h = NW_h / \sum_{k=2}^{n} NW_k, h = 2,..., n$, and C^m is the convex combination of *m* labels and if $m = 2$, C^2 is defined as

$$C^2\{NW_i, b_i, i = 1,2\}$$
$$= NW_1 \odot s_j \oplus (1 - NW_1) \odot s_i = s_k, s_j, s_i \in S, (j \geq i)$$

such that $k = \min\{T, i + round(NW_1 \cdot (j - i))\}$, where $b_1 = s_j, b_2 = s_i$.

If $NW_j = 1$ and $NW_i = 0$ with $i \neq j, \forall i$, then the

* The concept of orness and andness was first introduced by Dujmovic in 1974 (see [8,9]).

convex combination is defined
$C^m\{NW_i, b_i, i=1,...,m\} = b_j$
So, calculating the weighting vector of LOWA operator, NW, is a basic question to be solved.
On the other hand, generally, the criteria are not really of equal importance. Issues of weighted aggregation operators have been studied in [7,12,16,22]. How to identify the importance weight vector IW of criteria and integrate it into the aggregation process is another problem.

4. WEIGHT ESTIMATING

LOWA operator with determinate weights

Most of the information aggregation methodologies in literature presume the weights are given in advance. Since it is not easy for the decision-maker to give certain and consistent weight vector, the weights are often regarded as fuzzy numbers or linguistic variables (e.g. [20]).

In our methodology the decision maker need not provide the weights subjectively; weights are regarded as interior quantitative variables.

We are first at the position to construct a general aggregation mechanism allowing for the both kinds of weights.

If the importance weight vector of criteria is $IW : IW_i \in [0,1], \sum_{i=1}^{n} IW_i = 1$, then for some $x : (A_1(x), A_2(x),..., A_n(x)) = (a_1, a_2,..., a_n)$, an aggregation operator φ can be defined as:

$$\varphi(a_1, a_2,..., a_n) = C^n\{W_i, b_i, i=1,...,n\}$$

where $W_i = NW_i \cdot IW_{\sigma(i)} / \sum_{i=1}^{n} (NW_i \cdot IW_{\sigma(i)}), i = 1,..., n$, σ is a permutation over {1,2,…,n} such that $(b_1, b_2,..., b_n) = (a_{\sigma(1)}, a_{\sigma(2)},..., a_{\sigma(n)})$.

When $\sum_{i=1}^{n} (NW_i \cdot IW_{\sigma(i)}) = 0$, $\varphi(a_1, a_2,..., a_n)$ is defined to be

$$\min\{a_1, a_2,..., a_n\} = \min\{b_1, b_2,..., b_n\}.$$

For the operator φ (we can call it LOW^{2}A), we have simple properties as follows.

1) If $NW = (1/n, 1/n,..., 1/n)$, then $W_i = IW_{\sigma(i)}$,
$\varphi(a_1, a_2,..., a_n) =$
$C^n\{IW_{\sigma(i)}, a_{\sigma(i)}, i = 1,..., n\}$.
This means that the aggregation result is consistent with the weighted averaging using criteria grades and criteria importance.

2) If $IW = (1/n, 1/n,..., 1/n)$, then $\varphi = \phi$
This means that if the criteria are equally important, then the aggregation is identical with the ordinary LOWA operation.

3) If $IW_i = 1$ for some $i \in \{1, 2,..., n\}$ and $NW_i \neq 0$ for all i, then $\varphi(a_1, a_2,..., a_n) = a_i$
In other words, if all the criteria but the *i*th criterion are not important, then the aggregation result is determined by the assessment on the *i*th criterion.

Proof: Because $W_{\sigma^{-1}(i)} = 1$, and $W_j = 0$ with $j \neq \sigma^{-1}(i), \forall j$, we have $\varphi(a_1, a_2,..., a_n) = b_{\sigma^{-1}(i)} = a_i$.

4) If $IW_i = 0$ for some $i \in \{1, 2,..., n\}$, then $\varphi(a_1, a_2,..., a_n)$ is independent of a_i.

Proof: This is obvious from the fact of $W_{\sigma^{-1}(i)} = 0$ and the definition of the convex combination C^m.

The function of φ can be illustrated by a simple example.

Assume the aggregation problem involves four criteria. $NW = (0.25, 0.25, 0.25, 0.25)$, which means both the "orness" and the "andness" of aggregation are 0.5, and the "dispersion" is maximum.

For an alternative x, if its assessment about four criteria is (s_0, s_3, s_2, s_4), the aggregation result by ordinary LOWA operator will be $\phi = s_3$, which doesn't take account of the importance of the four criteria.

However, if the criteria weight vector is given, for instance, $IW = (0.2, 0.3, 0.4, 0.1)$, then one can calculate $W = (0.1, 0.3, 0.4, 0.2)$, and $\varphi = s_2$. This is justifiable since the assessment for x on the 3rd criterion is s_2 and the importance of the 3rd criterion increases.

Weight estimating model

Making use of the LOW^{2}A operator defined above, the aggregation mechanism is mapped to configuration of the weights NW and IW.

We are now at the position to build a model to learning the weights from the set of alternatives with known aggregation results.

Taking $NW_1,..., NW_n, IW_1,..., IW_n$ as variables, according to the cases and information given, a mathematical programming can be established as follows.

1) Each alternative case $(a_1, a_2,..., a_n)$ with known result s_k can be transformed into a constraint equation. $\varphi(a_1, a_2,..., a_n) = s_k$

2) The information about the comparison of criteria importance should be expressed by equation or inequation in weight variables.

3) The objective function can be maximizing or minimizing the dispersion of NW, or IW, or both.

The ultimate model turns out to be a nonlinear programming.

5. AN EXAMPLE

In this section, we employ an example to show how the approach can be used to evaluate the experts through their own evaluation history.

For a group of alternatives c_1, c_2, c_3, c_4, the assessment results according to four experts and their actual performance ranks are as follows (a_{ij} is the assessment of

the alternative c_j given by the *i*th expert).

1) For c_1: $a_{11} = s_1, a_{21} = s_5, a_{31} = s_4, a_{41} = s_6$,

its actual performance turns out to be s_5.

2) For c_2: $a_{12} = s_1, a_{22} = s_3, a_{32} = s_3, a_{42} = s_1$,

its actual performance turns out to be s_2.

3) For c_3: $a_{13} = s_4, a_{23} = s_5, a_{33} = s_3, a_{43} = s_2$,

its actual performance turns out to be s_4.

4) For c_4: $a_{14} = s_1, a_{24} = s_6, a_{34} = s_1, a_{44} = s_0$,

its actual performance turns out to be s_3.

In addition, it is known that the 3rd expert is more knowledgeable than the 2nd expert.

Then, the mathematical programming minimizing the sum of dispersion of NW and of IW can be described as:

$$\max \; -\sum_{i=1}^{4} NW_i \ln NW - \sum_{i=1}^{4} IW_i \ln IW$$

$$s.t. \varphi(s_1, s_5, s_4, s_6) = s_5$$

$$\varphi(s_1, s_3, s_3, s_1) = s_2$$

$$\varphi(s_4, s_5, s_3, s_2) = s_4$$

$$\varphi(s_1, s_6, s_1, s_0) = s_3$$

$$IW_3 \geq IW_2$$

$$NW_i \geq 0, IW_i \geq 0, (i = 1,2,3,4);$$

$$\sum_{i=1}^{4} NW_i = 1, \sum_{i=1}^{4} IW_i = 1$$

Solving this programming, one can get,

$$NW = (0.225, 0.246, 0.248, 0.28);$$
$$IW = (0.243, 0.244, 0.244, 0.269)$$

If the objective is to minimize the sum of dispersions, the result will be:

$$NW = (0., 0.244, 0.310, 0.446);$$
$$IW = (0., 0.308, 0.308, 0.383)$$

This means, the overall performance can be obtained without the evaluation information due to the first expert. It doesn't mean the first expert is the least knowledgeable, since there may be another sets of weight configuration can lead to the same conclusion. Nevertheless, the uncertainty will be reduced when more information is given or more cases are presented.

Using these weights, aggregation over a new alternative can be processed conveniently.

It can be seen that, for some alternatives, if their actual performance can't be identified definitely, the information in other form, such as $\varphi(a_{1j}, ...a_{nj}) \geq \varphi(a_{1k}, ...a_{nk})$ (i.e. the actual performance of c_j is better than c_k), can also be easily integrated into the model.

6. DISCUSSIONS AND CONCLUSIONS

In this paper, we present an aggregation approach based on regression from a set of cases. The weights, which play a key role in aggregation, are adjusted essentially by the case set. So the cases should be selected carefully. The case set may be expanded incrementally. For example, if a new alternative, assessed by the expert, has a result evidently different from the one obtained by this approach, then the alternative should be considered to add as a member of the case set.

An amendment of the model can be a goal programming with some percentage of the cases satisfying constraint equations.

Compared with the weights learned from an artificial neural network, the weights obtained here have explainable meaning.

7. REFERENCES

[1] R. Andrews, J. Diederich, et al, "Survey and critiques for extracting rules from trained artificial neural networks", Knowledge-Based Systems, 1995,8:373-389.

[2] W. Cholewa, "Aggregation of fuzzy opinions: An axiomatic approach", Fuzzy Sets and Systems, 1985,17:249-259.

[3] R. Degani, G. Bortolan, "The problem of linguistic approximation in clinical decision making", International Journal of Approximate Reason, 1988,2: 143-162.

[4] M. Delgado, J. L.Verdegay, et al, "On aggregation operations of linguistic labels", International Journal of Intelligence Systems, 1993,9:351-370.

[5] M. Deng, W. Xu et al, "Estimating attribute weights through evidential reasoning and mathematical programming", International Journal of Information Technology and Decision making, 2004,3: 419-428.

[6] D. Dubois, H. Prade, "Weighted minimum and maximum operations in fuzzy set theory", Information Science, 1986,39:205-210.

[7] D. Dubois, H. Prade, et al, "Weighted fuzzy pattern matching", Fuzzy Sets and Systems, 1988, 28:313-331.

[8] J. Dujmovic, "Weighted conjunctive and disjunctive means and their application in system evaluation", Journal of the University of Belgrade, EE Dept., Series Mathematics and Physics, 1974,483:147-158.

[9] J. J. Dujmovic, H. Nagashima, "LSP method and its use for evaluation of Java IDEs", International Journal of Approximate Reasoning, 2006,41:3-22.

[10] A.E. Flalfi, R. Haque, et al, "Extracting rules from trained neural network using GA for managing E-business", Applied Soft Computing, 2004,4:65-77.

[11] F. Herrera, E. Herrera-Viedma, et al, "Direct approach processes in group decision making using linguistic OWA operators", Fuzzy Sets and Systems, 1996,79:175-190.

[12] F. Herrera, E. Herrera-Viedma, "Aggregation operators for linguistic weighted information", IEEE Transactions on Systems Man Cybernetics, 1997,27: 646-656.

[13] F. Herrera, E. Herrera-Viedma, "Linguistic decision analysis: steps for solving decision problems under linguistic information", Fuzzy Sets and Systems, 2000,115:67-82.

[14] E. Hinloopen, P. Nijkamp, et al, "Integration of ordinal and cardinal information in multi-criteria ranking with imperfect compensation", European Journal of Operational Research, 2004,158:317-338.

[15] F. Rosenblatt, Principles of Neurodynamics, New York, Spartan Books, 1962

[16] E. Sanchez, "Importance in knowledge systems", Information System, 1989,14:455-464.

[17] H. Schjaer-Jacobsen, “Representation and calculation of economic uncertainties: Intervals, fuzzy numbers, and probabilities”, International Journal of Production Economics, 2002,78:91-98.
[18] G. Shafer, A mathematical theory of evidence, Princeton, N.J: Princeton University Press, 1976
[19] M. Tong, P. P. Bonissone, “A linguistic approach to decision making with fuzzy sets”, IEEE Transactions on Systems Man Cybernetics, 1980,10:716-723.
[20] Z. Xu, “The hybrid linguistic weighted averaging operator”, Information, 2005,8(3): 453-456.
[21] R.Y. Ronald, “On ordered weighted averaging aggregation operators in multicriteria decision making”, IEEE Transactions on Systems Man Cybernetics, 1988,18:183-190.
[22] R.Y. Ronald, “On weighted median aggregation”, International Journal of Uncertainty, Fuzziness, Knowledge-Based System, 1994,2:101-113.
[23] R.Y. Ronald, “Modeling uncertainty using partial information”, Information Science, 1999,121:271-294.
[24] R.Y. Ronald, “On the determination of strength of belief for decision support under uncertainty—Part II: fusing strengths of belief”, Fuzzy Sets and Systems, 2004,142:129-142.
[25] L.A. Zadeh, “Fuzzy sets”, Information and Control, 1965, 8:338-353.

Research of Artificial Neural Network Algorithm Based on Fuzzy Clustering of Rough Sets

Yan Chen
School of Economics and Management, Dalian Maritime University
Dalian, Liaoning, China
Email: Chenyan_dlmu@163.com
And
Taoying Li
School of Economics and Management, Dalian Maritime University
Dalian, Liaoning, China
Email: ytaoli@126.com

ABSTRACT

BPNN (back propagation neural network) has advantages of good learning and memory, but it hasn't the ability to train samples with qualitative attributes. On the contrary, fuzzy clustering of rough sets based on fuzzy equivalent matrix has the advantage. In the paper, Fuzzy clustering of rough sets is used as anterior processor of BPNN, which is to say that firstly data records are partitioned by fuzzy clustering of rough sets and then are inputted into BPNN to be learned and memorized. BPNN model based on fuzzy clustering of rough sets and model of Multivariate Linearity Regression are both applied in costume field. By comparing the results of BPNN model with that of Multivariate Linearity Regression model, the results show that the BPNN model based on fuzzy clustering of rough sets is more effective and practical.

Keywords: Artificial neural network, BPNN, Fuzzy clustering, Fuzzy Equivalent Matrix, Fuzzy Similar Matrix.

1. INTRODUCTION

There are many useful rules for production and life in practice, outside of which are known, however, the inside is undefined. Function simulation is adopted in order to solve problems of this kind. The excellent capability of function simulation of ANN (artificial neural network) makes it one of best choices to simulate unknown function. Since the MP model was advanced by McCulloch and Pitts in 1943, ANN develops more and more quickly and always is a hot topic except in the intervals of two low tides. ANN trains samples by means of corresponding relationship between input and output of samples, and then reaches to favorable simulating effect in tolerable range so that the network can be used to simulate response of things. However, ANN can't be used to solve qualitative attributes, which is primary difficulty.

Rough sets can be used to accomplish clustering of objects, the attributes of which can not only be quantitative but also qualitative. Therefore, fuzzy clustering of rough sets is used as anterior processor of BPNN in this paper. Samples firstly are clustered by fuzzy clustering of rough sets, and then are inputted into BPNN. Finally, BPNN model based on fuzzy clustering of rough sets and multivariate regression model are both applied to size classification of costume. By comparing the results of the two models, result shows that the BPNN model based on fuzzy clustering of rough sets is more perfect and effective.

In section 2, the fuzzy clustering and its corresponding theories are introduced, which can be used to prune redundant attributes. Theory, structure and training process of BPNN are presented in section 3. Finally, the practical instance and its solution are given in section 4.

2. FUZZY CLUSTERING THEORY

Clustering algorithm is a method in common use in data mining. Existing clustering algorithms in common use are composed of partitioning method, Analytic Hierarchy Process, density method, grid method, and so on. Aiming at rough sets decision-making table consists of qualitative attributes and quantitative attributes, objects of rough sets can be measured by a measurement of comparability in statistics and be translated into fuzzy similar matrix [1], which will be formed into fuzzy equivalence matrix by transferable close package. Consequently clustering without surveillance will be accomplished.

2.1 Theory of Rough Sets

Knowledge system can be described as a four-item group: S= (U, A, V, f), U is set of all objects of data set, A is set of all attributes of data set, $A=C \cup D$, C is set of conditional attributes, D is set of decision-making attributes, V is set of attribute value, and f: $U \times A \rightarrow V$ is info-function and it designates the attribute value of every object.

Table 1.Decision-making table

Object	a_1	a_2	a_3	…	a_m	d
x_1	x_{11}	x_{12}	x_{13}	…	x_{1m}	d_1
x_2	x_{21}	x_{22}	x_{23}	…	x_{2m}	d_2
x_3	x_{31}	x_{32}	x_{33}	…	x_{3m}	d_3
…	…	…	…	…	…	…
x_n	x_{n1}	x_{n2}	x_{n3}	…	x_{nm}	d_n

Table 1 is decision-making table, $\{x_1,x_2,x_3,\ldots,x_n\}$ is set of objects, $\{a_1,a_2,a_3,\ldots,a_m\}$ is set of conditional attributes and {d} is set of decision-making attribute. Decision-making attribute, which contains two attributes or more, can be easily split into only one attribute [2].

2.2 Fuzzy Clustering Algorithm

Because attributes of object may be composed of qualitative attributes and quantitative attributes, the measurement of comparability must be flexible.

Euclidean distance like formula (1) often is adopted for quantitative attributes and Hamming distance like formula (2) usually is used for qualitative attributes. In order to be short and sweet, d_i can be expressed as $x_{i,m+1}$ and qualitative attributes stay behind quantitative attributes.

Euclidean distance:

$$d_E(x_i, x_j) = [\sum_{k=1}^{r} | x_{ik} - x_{jk} |^2]^{\frac{1}{2}} \quad (1)$$

x_i, x_j are two random objects, and r is the total number of quantitative attributes.

Hamming distance:

$$d_H(x_i, x_j) = \sum_{k=r}^{m+1} | x_{ik} - x_{jk} | \quad (2)$$

The value of Hamming distance of two objects is 0 when their value of the same qualitative attribute is equivalent, otherwise the value is 1.

In order to describe the clustering clearly, now the paper gives some definitions and theorems.

Definition 1: Supposing $R=(r_{ij})_{m\times m}$ is a fuzzy matrix, if R satisfies conditions:① self-same: $r_{ii}=1$;②symmetry: $r_{ij}=r_{ji}$;③ transfer: $\max\{\min(r_{ik},r_{kj})|1\le k\le m\}\le r_{ij}$, then R is a fuzzy equivalent matrix.

Definition 2: Supposing $R=(r_{ij})_{m\times m}$ is a fuzzy matrix, if R satisfies conditions:① self-same: $r_{ii}=1$;②symmetry: $r_{ij}=r_{ji}$, then R is a fuzzy similar matrix.

Definition 3: Supposing R and S are fuzzy similar matrixes, then the multiplication of fuzzy similar matrixes is defined as following:

$$T=R\circledcirc S=(t_{ij})_{m\times m},\ t_{ij}=\max\{\min\{r_{ik},s_{kj}\}|1\le k\le m\} \quad (3)$$

If S=R, then R⊛S can be denoted as R^2=R⊛R.

To popularize the algorithm, R^k= R⊛R⊛…⊛ R⊛R.

Theorem 1: Supposing $R=(r_{ij})_{m\times m}$ is a fuzzy similar matrix, then there will be a natural number k(k≤m), which satisfies the conditions that R^k is a fuzzy equivalent matrix and $R^k=R^d$(d is a natural number and d≥k).

Correlation of two random objects is measured by following formula:

$$r(x_i, x_j) = \frac{r}{m+1}\left(1 - \frac{d_H(x_i, x_j)}{\max\limits_{1\le u,v\le n} d_H(x_u, x_v)}\right) + \frac{m+1-r}{m+1}\left(1 - \frac{d_H(x_i, x_j)}{\max\limits_{1\le u,v\le n} d_H(x_u, x_v)}\right) \quad (4)$$

Formula (4) denotes similar coefficient and similar degree of x_i and x_j. Denominators are maximal value of Euclidean distance of quantitative attributes and maximal value of Hamming distance of qualitative attributes. A fuzzy similar matrix can be obtained by computing objects like formula (4). According to theorem 1, the relevant fuzzy equivalent matrix can be obtained, then designating threshold α . Objects, whose similar coefficient in fuzzy equivalent matrix is larger than or equal to α , are partitioned into one part and other objects are partitioned into another part. As long as we want to keep partitioning, we can continue use the algorithm of fuzzy clustering until it satisfies needs.

3. BPNN ALGORITHM

Usually, there are plentiful data, and it is necessary to abstract data from large numbers of samples. The easiest method to abstract data is proportional selection, which is to select appropriate sample randomly. Then, after having removed the redundant attributes of objects, these samples are the input data of BPNN. Fuzzy clustering is used just as the anterior processor of BPNN.

The detailed algorithm of BPNN is presented in following sections.

3.1 BP Neural Network Theory

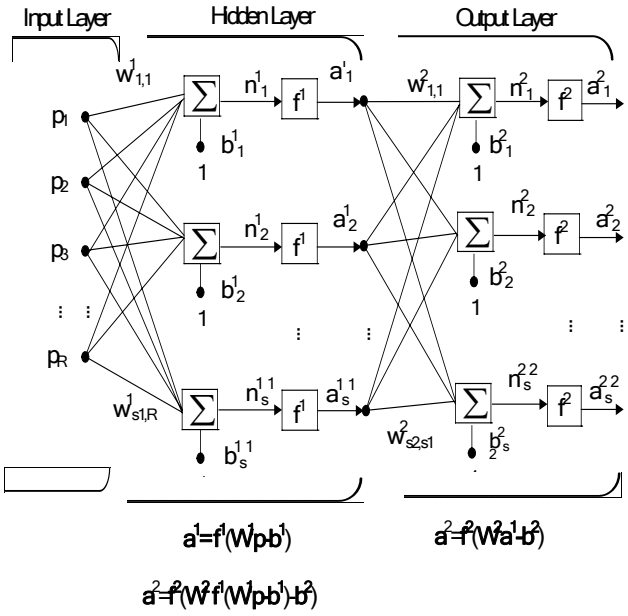


Fig. 1. Structure of BPNN

Fig.1 is the structure of BPNN with only one hidden layer. Superscript of letters denotes serial number of computing layers. In first computing layer between input layer and hidden layer, R is the total number of nerve cells in input layer. p_i (i=1,2,3,…,R) is i element of input sample. $\mathbf{W^1}$ is the weight matrix between hidden layer and input layer. $w^1_{j,i}$ is weight between nerve cell i in input layer and nerve cell j in hidden layer. b^1_i is the threshold of nerve cell i in hidden layer. n^1_i is the input of nerve cell i in hidden layer. $\mathbf{f^1}$ is transmission function. $\mathbf{a^1}$ is output vector in hidden layer and input vector in output layer. The second computing layer between output layer and hidden layer is just the same [3].

Usually, sigmoid function as Eq. (5) is used as the transmission function in hidden layer.

$$f^1 = \frac{1}{1+e^{-x}} \quad (5)$$

Transmission function in output layer generally adopts linearity function:

$$f^2 = ax + b \quad (6)$$

BPNN algorithm consists of four processes as follow:

1) Forwards propagation, which is the process that data transmits from input layer to output layer through hidden layer.

2) Error back propagation, which is the process that error between target output and actual output transmits from output layer to input layer through hidden layer.

3) Repeating memory training, which is the process of alternation between forwards propagation and error back propagation.

4) Network tending to convergence, which is the process of error verging minimum.

Essential thought of BPNN is LMS (Least Means Square), which adopts gradient-searching method and expects squared error sum between target outputs and actual output minimum [4].

3.2 Training of BPNN

Monolayer network is too simple to satisfy actual demand because of the complexity of practical problems, so multilayer network is often adopted in practice. But too many layers will lead to network structure complex and model hard to establish. BPNN with one hidden layer can simulate arbitrary function if only the number of nerve cells is enough large [4]. Therefore, it is widely used in practice.

The nerve cells in output layer can be limited into only one by partitioning the network.

Firstly, the training propagates forwards and produces the output of BPNN as following equations:

$$a^0 = p_q \quad (7)$$

$$a^{m+1} = f^{m+1}(w^{m+1}a^m + b^{m+1}), m = 0,1,..,M\text{-}1 \quad (8)$$

$$a^m = a^{m+1} \quad (9)$$

Squared error sum of each sample is able to obtain by computing as Eq. (10).

$$E_q = \sum_{j=1}^{S^M} (t_{j,q} - a^M{}_{j,q})^2 \quad (10)$$

In Eq. (10), E_q is squared error sum of sample q and M is total number of computing layers. The average of squared error sum can be computed as Eq. (11).

$$E = \frac{1}{Q}\sum_{q=1}^{Q} E_q \quad (11)$$

In Eq. (11), Q is the total number of samples. In this paper, there are two computing layers, so M=2. $t_{j,q}$ is element j of target output of sample q, $a^M_{j,q}$ is element j of actual output of sample q, and s^M denotes sensitivity of output layer.

Secondly, error propagates back.

$s^M_i = -2(t_j - a_j)\frac{\partial a_i}{\partial n^M_i}$, sensitivity propagates from last layer to first layer like: $s^M \rightarrow s^{M-1} \rightarrow \ldots \rightarrow s^2 \rightarrow s^1$.

$$s^m = F^m(n^m)(W^{m+1})^T s^{m+1} \quad (12)$$

Sensitivity of layer m can be obtained by computing as Eq. (12). $(W^{m+1})^T$ is transpose of weight matrix of layer m+1. $F^m(n^m)$ is diagonal matrix and the element of row i and column j is expressed as Eq. (13).

$$f^m(n^m{}_j) = \frac{\partial f^m(n^m{}_j)}{\partial n^m{}_j} \quad (13)$$

Finally, the weight can be modified as Eq. (14) and Eq. (15).

$$W^m(k+1) = W^m(k) - Lr\, s^m (a^{m-1})^T \quad (14)$$

$$b^m(k+1) = b^m(k) - Lr\, s^m \quad (15)$$

In Eq. (14), $W^m(k+1)$ is the new weight matrix of layer m for k+1 times, and it is equal to weight for k times plus a increment $-Lr\, s^m (a^{m-1})^T$. Lr is learning rate of BPNN and $(a^{m-1})^T$ is output transpose of layer m-1. In the same way, corresponding threshold is expressed as Eq. (15).

4. INSTANCE ANALYSIS

Firstly, collects some history data of male costume from corporation, tidies data up. More than 2000 records are collected, which contains man-made index data and corresponding size data. 800 records are randomly abstracted as samples, in which 600 records are training samples and the rest are test samples. SQL Server [5] is used as database. There are samples in Fig.2.

shg	xw	yw	tw	zjk	bc	xdc	byc	qyc	hbk	qxk	ywdg	ph	fzgg
175	102	94	105	48	56	75	47	47	45	44	105	feat	175/100A
170	90	78	96	43	54	73	39	42	39	38	102	feat	170/88A
178	93	78	101	44	57	76	41	42	39	39	106	small	175/92A
172	92	83	100	46	57	74	41	43	39	39	104	big	175/92A
170	102	92	105	45	53	68	41	43	41	41	98	small	165/104B
168	93	83	96	43	53	69	40	41	39	38	99	small	165/92A
170	98	89	103	44	53	72	41	43	40	40	102	feat	170/100B
178	92	83	97	44	56	77	43	45	39	38	106	small	175/92A
166	90	77	97	44	52	68	39	42	39	38	97	feat	165/88A
176	103	93	103	46	59	75	43	44	41	43	106	big	180/104B
174	97	85	98	46	57	74	41	43	39	42	103	feat	175/96A
174	91	78	99	44	55	75	40	43	39	37	104	feat	175/92A
173	102	92	104	45	53	72	42	44	42	41	102	small	170/100A
172	104	97	104	46	55	69	44	45	42	43	100	feat	170/104B
184	107	96	109	47	58	78	43	45	42	44	109	small	180/108B
170	99	87	99	45	52	70	40	43	41	40	100	small	165/100A

Fig. 2. Part of samples

Fig.2 is part of samples. Each record is costume information of each person. The numbers in Fig.2 are body data, and ph is preference of each person and fzgg is size standard of each person. After samples being pretreated, they should be clustered by fuzzy clustering algorithm based on rough sets and then be inputted into BPNN model to train. The run of BPNN is like Fig.3.

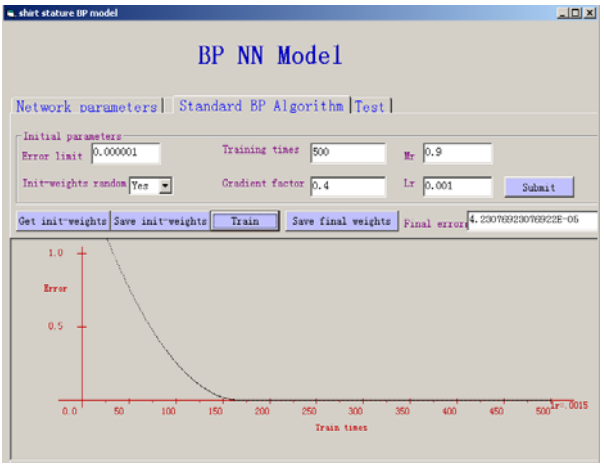


Fig. 3. Run of BPNN

Parameters in Fig.3 are parameters of BPNN. After all parameters are inputted, start the training. The change of squared error sum of training is like the curve in Fig.3. Save the weight matrix and threshold until average of squared error sum is in an acceptable range, and exit the training.

In order to verify the practicability of the model, Multivariate regression model is also established according to same samples. Finally, BPNN model can be used to test samples. The results of test samples are as Table.2.

Table 2. Results of test

Target Output	Output of BPNN	Output of Multivariate Regression
175/100B	175/100B	175/100B
175/96A	175/96A	175/96A
175/104B	175/104B	175/104B
175/100B	175/100B	175/100B
170/100B	170/100B	170/102B
180/104B	180/104B	180/104B
175/100B	175/100B	175/100B
…	…	…
Number of feat size	167	107
Rate of feat size	83.55%	53.5%

The results of test in Table.2 show that the feat rate of output of BPNN based on fuzzy clustering of rough sets is 83.55%, which outclasses that of Multivariate Regression. So BPNN based on fuzzy clustering of rough sets is practical and effective.

5. CONCLUSIONS

Because of the disadvantages of BPNN, it can't be used to train samples with quantitative attributes. However, fuzzy clustering of rough sets overcomes the flaws, and it is combined with BPNN in the paper. Fuzzy clustering of rough sets is used as anterior processor of BPNN. Then, the BPNN model based on fuzzy clustering of rough sets and the Multivariate regression model are both established in size classification of costume in order to show the effect of BPNN model. Finally, test samples are inputted into the two models, the results of test show that the accuracy of BPNN model based on fuzzy clustering of rough sets is outclasses that of Multivariate Regression, so it is more practical and effective.

6. REFERENCES

[1] P. Lingras, "Unsupervised rough set classification using gas", Journal of Intelligent Information System, Vol.16, No.3, 2001, pp: 215-228.

[2] P. Lingras, "Rough set clustering for web mining", Proceeding of IEEE International Conference on Fuzzy Systems, New York: IEEE CS, 2002, pp: 1039-1044.

[3] M.T. Hagan, H.B. Demuth, et al, Neural Network Design, Beijing: China Machine Press, 2003.

[4] P.F. Yan, C.S. Zhang, Artificial Neural Network and Simulant Evolutive Calculation, Beijing: Tsinghua University Press, 2001.

[5] S.X. Sa, S. Wang, An Introduction to Database System, Beijing: Higher Education Press, 2004.

Image Edge Detection Using Improved PCNN Model

Lan Di
School of Information Technology, Southern Yangtze University
Wuxi, 214122, China
Email: dilan126@163.com
Yi Lin
School of Information Technology, Southern Yangtze University
Wuxi, 214122, China
Email: liny01@vip.sina.com

ABSTRACT

A calculated method is proposed which uses the region areas to modulate the pulse of the corresponding neurons in PCNN. The edge detection for the blurred Image with noise can reduce the noise and obtain a precise edge value of the image. First, based on the analysis of the difference between the edge pixels and the noise pixels, it proposed a method, which suppresses the noise by the region areas. Second, it suggested that image segmentation uses the pulse of the corresponding neurons in PCNN. The experimental results show that this method is effective and feasible.

Key words: Region Area; PCNN; Edge Detection; Noise.

1. INTRODUCTION

Edge detection is one of the most commonly used operations in image analysis, and there are probably more algorithms in the literature for enhancing and detecting edges than any other single subject. The reason for this is that edges form the outline of an object. An edge is the boundary between an object and the background, and indicates the boundary between overlapping objects. This means that if the edges in an image can be identified accurately, all of the objects can be located and basic properties such as area, perimeter, and shape can be measured. Since computer vision involves the identification and classification of objects in an image, edge detection is an essential tool.

Faced with the blurred Image with noise, we should suppress the noise, and extrude the edge at the same time. Now image smoothing is the set of local methods whose predominant use is the suppression of image noises---- it uses average filter and median filter; and image sharpening is the set of usual methods whose predominant use is the detection of image edges ---- it uses Sobel operator, Kirsch operator, canny operator and Laplacian operator. These operators are represented by a collection of masks. They have some questions ----- a weak self-adaptability, a sensibility for the noise etc. It is our goal to find an edge detection method of a stronger self-adaptability and counteracting noise. So, based on the characteristic of noises and edge pixels, an edge detection algorithm is proposed which uses the region areas to modulate the pulse of the corresponding neurons in PCNN.

The Pulse-Coupled Neural Network, PCNN, is different from traditional artificial neural network. PCNN, according to pulse bursts phenomenon in the visual cortex of cat, monkey and so on, has biological background and was applied to image processing, image recognition, moving object recognition, communication, optimization [1,2,3,4,5,6], etc.

2. BASIC PCNN MODEL

According to the phenomena of synchronous pulse bursts in the cat visual cortex, Eckhorn brought forward the linking field network [7]. When experimental objects are monkeys, the same results of experiments are obtained [8]. When some modifications were introduced to the linking field network, it became PCNN-pulse coupled neural network. The Eckhorn model is shown in Fig. 1.

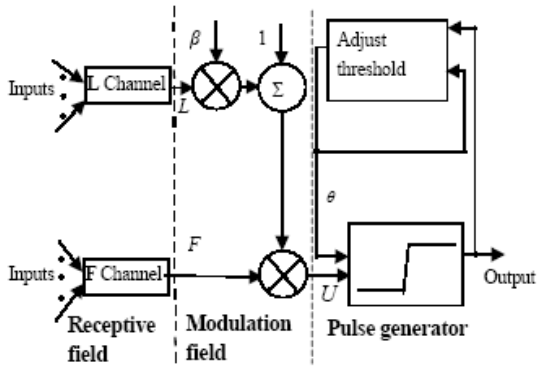


Fig. 1. A Pulse Coupled Neuron

PCN (Pulse Coupled Neuron) consists of three parts: the receptive field, the modulation field, and the pulse generator (See Fig. 1.). A PCN receives signals through the receptive field. There are two channels in a PCN. One channel called F channel is feeding input; the other called L channel is linking input. There are different synapses with different time constants in each channel. In general, the feeding connection has a slower characteristic response time constant than that of the linking connection. In modulation field, the linking input L is added a constant positive bias firstly. Then it is multiplied by the feeding input F and the bias is taken to be unity.

The equations from (1) to (5) describe this model [4]:

$$F_{ij}(n) = e^{-a_F} F_{ij}(n-1) + S_{ij} + V_F \Sigma M_{ijkl} Y_{kl}(n-1) \quad (1)$$

$$L_{ij}(n) = e^{-a_L} L_{ij}(n-1) + V_L \Sigma W_{ijkl} Y_{kl}(n-1) \quad (2)$$

$$U_{ij}(n) = F_{ij}(n)(1 + \beta L_{ij}(n)) \quad (3)$$

$$\theta_{ij}(n) = e^{-a_\theta} \theta_{ij}(n-1) + V_\theta Y_{ij}(n-1) \quad (4)$$

$$Y_{ij}(n) = step\left(U_{ij}(n) - \theta_{ij}(n)\right) \quad (5)$$

Where $S_{ij}, F_{ij}, L_{ij}, U_{ij}, \theta_{ij}$ is the external stimulation of neuron, feedback input, linking input, internal activity and dynamic threshold respectively? M And W are linking weight matrices (generally $W = M$), V_F, V_L, V_θ are amplitude constant, $\alpha_F, \alpha_L, \alpha_\theta$ are time decay constant, *n* is a cyclic iterative time, and Y_{ij} is binary value output.

A two dimensional image ($M * N$) can be thought as a PCNN network with $M * N$ neurons, and the gray level of pixels can be thought as S_{ij} , the input of the neuron. Obviously, M and W are the interior linking matrixes. When there are pixels whose gray levels are approximate in the neighborhood of M and W , one pixel's pulsating output can activate other corresponding pixels having the approximate gray level in the neighborhood and let them generate pulsating output sequence $Y(n)$. Obviously $Y(n)$ contains some information about this image such as regional information, edge, and texture features. Then the binary image constructed by $Y(n)$, the output of PCNN, is the segmented image. This is why the PCNN achieved the image segmentation.

3. IMPROVED PCNN ALGORITHM

Faced with the blurred Image with noise, we should suppress the noise, and protect the edge at the same time. We know that protecting edge and suppressing noise are incompatible. Protecting edge is to increase the brightness margins between the pixels, but suppressing noise is to decrease them. How to resolve this problem?

The noise means that an image is corrupted with noisy pixels whose brightness differs significantly from that of the neighborhood. The noise is usually described by its probabilistic characteristics, such as Gaussian noise, Salt-and-pepper noise, etc. In this paper, we consider only Salt-and-pepper noise. The Salt-and-pepper noise is used to describe saturated impulsive noise. It usually locates the brightest area or the darkest area of the pixels brightness. This noise's region areas are smaller. So we think that it is a noise if a region area is smaller than a special threshold.

In the case of noise density is not big, the function E_{ij} is:

$$E_{ij} = \begin{cases} 1, \text{ for } P(i, j) \text{ is not a noise} \\ 0, \text{ for } P(i, j) \text{ is a noise} \end{cases} \quad (6)$$

Where *P (i, j)* is a pixel of an image.

Following is the procedure to judge whether a pixel is a noise:

Step1.

Let pixel *P (i, j)* to be used as the center. If its 8-connected neighborhood satisfies a property of similarity with *P (i, j)*, merge it.

Step2.

Let new merging pixel to be used as the center. Repeat Step 1 until the region can't grow.

Step3.

Calculate the area of the region. Compare it with the special threshold (*T*). If area<*T*, *P (i, j)* is a noise; else it isn't.

Because the pixel of an image can either be the noise or non-noise, the threshold is confirmed by using the minimal threshold of the region area histograms. An area histogram is given in Fig. 2.. The horizontal axis is the area and the vertical axis is the number of area regions.

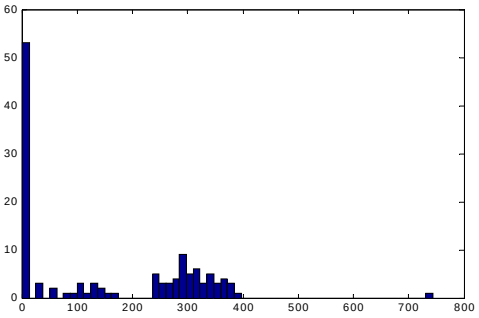


Fig. 2. The region area histogram

The region area is the total number of the pixels in the region. But it is believed that the neighborhood pixel is in the same region, its gray-levels difference is within certain limit in actual calculating.

This paper does not calculate the pixel edge value of an image by its gray-levels directly. It is to take the gray-levels of an image as the input of PCNN. Replace the gray-levels in the algorithm with the margin of an internal activity and a threshold $(U_{ij} - \theta_{ij})$, and use the function E_{ij} to regulate a pulse value of the neuron.

$$Y_{ij}(n) = step\left(U_{ij}(n) - \theta_{ij}(n)\right) * E_{ij} \quad (7)$$

Replace equation (5) with equation (7). The algorithm is called Improved PCNN. The Work principle of Improved PCNN is the same as the traditional PCNN.

4. EXPERIMENTAL RESULTS

Several experimental results are presented in this section to demonstrate the power of our Improved PCNN algorithm.

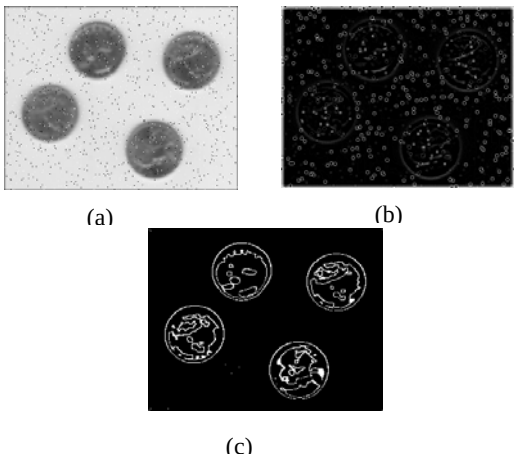


Fig. 3. Results for the testing image (a) original image (b) Laplacian algorithm (c) Improved PCNN algorithm

The blurred image, with salt and pepper noise, is employed and is shown in Fig. 3. (a). Edge Detection using our new

algorithm is compared with Laplacian algorithm. Fig. 3. (b) shows the result of using average filter and Laplacian operator filter. Fig. 3. (c) Shows the result of using our Improved PCNN algorithm. It can be seen that the difference between the original and the new filtered image is obvious, which implies that the ability of preserving edges and details of the new algorithm is quite well. All noisy images are generated by MATLAB 6.x, each channel of which is corrupted independently with impulsive salt and pepper noise.

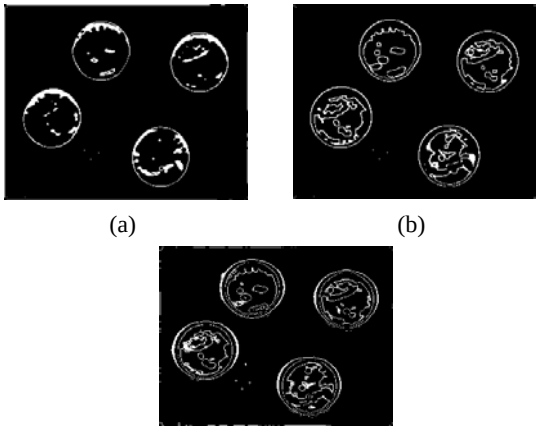
In our experiments, the area threshold: $T = 5$, the parameter of PCNN:

$V_\theta = 20, V_F = 0.5, V_L = 0.2, \alpha_\theta = 0.2, \alpha_F = 0.1, \alpha_L = 0.3,$ M And W masks are:

$$\begin{bmatrix} 0.707 & 1 & 0.707 \\ 1 & 1 & 1 \\ 0.707 & 1 & 0.707 \end{bmatrix}$$

It is found that the results of algorithm appear in periodic repetition with the iterative numbers and are shown in Fig. 4. .

So, the choice of the circulation iterative number *n* is related to the examination result of the image directly. It is said that *n* has a best value for particular image and the parameter of PCNN. In our experiments, choose *n* value is 26 by repeated comparison. The regulation of this approach remains to be further studied.



(a) (b)

(c)

Fig. 4. Results for the periodic repetition
(a) *n*=17 (b) *n*=26 (c) *n*=39

5. CONCLUSIONS

A new edge detection algorithm is proposed for the blurred Image with noise. The algorithm is divided into two groups. First, based on the analysis of the difference between the edge pixels and the noise pixels, it proposed a method, which suppresses the noise by the region areas. Second, it suggested that image segmentation uses the pulse of the corresponding neurons in PCNN. So this algorithm has a stronger self-adaptability and ability to counteract noise. The experiment shows that this method is effective and feasible.

6. REFERENCES

[1] R.P. Broussard, S.K. Rogers, “Physiologically Motivated Image Fusion for Object Detection using a Pulse Coupled Neural Network”, IEEE Trans. Neural Networks, Vol. 10(3), 1999, pp.554-563.

[2] X. Liu, D.L. Wang, “Rang image segmentation using a relaxation oscillator networks”, IEEE Trans. Neural Networks, Vol. 10(3), 1999, pp. 564-573.

[3] J.M. Kinser, “Foveation by a Pulse-Coupled Neural Network”, IEEE Trans. Neural Networks, Vol. 10(3), 1999, pp. 621-625.

[4] J.L. Johnson, M.L. Padgett, “PCNN Models and Applications”, IEEE Trans. Neural Networks, Vol. 10(3), 1999, pp. 480-498.

[5] H. Jcaufield, J.M. Kinser, “Finding shortest path in the shortest time using PCNN’s”, IEEE Trans. Neural Networks, Vol. 10(3), 1999, pp. 604-606.

[6] M.X. Shi, J.Y. Zong, “Image Binary Segmentation Based on Improved PCNN Model”, Computer Simulation, Vol. 19(4), 2002, pp. 42-46.

[7] R. Eckhorn, H.J. Reitboeck, et al, “Feature linking via synchronization among distributed assemblies”, Simulation of results from cat cortex. Neural Comput, Vol. 2(3), 1990, pp. 293-307.

[8] R. Eckhorn, A. Frien, et al, “High frequence oscillations in primary visual cortex of a wake monkey”, Neuron Rep., Vol. 4(3), 1993, pp. 243-246.

Research of Enhancing Smooth Degree of Sequence Based on Multiplex Transformation and Its Application

Junfeng Li[1,2]
1. Department of Automatic control, Zhejiang Sci-Tech University
Hangzhou, China
2. College of Mechanical Engineering, Donghua University
Shanghai, China
Email: ljf2003zz@163.com
And
Wenzhan Dai
College of Information and Electronic Engineering, Zhejiang Gongshang University
Hangzhou, China
Corresponding author
Email: dwzhan@zstu.edu.cn

ABSTRACT
In this paper, the method of improving GM (1,1) model is put forward based on multiplex transformation. First, the necessary and sufficient condition of smooth sequence is given. It is proved that smooth degree of sequence can be enhanced under some conditions. Secondly, the new concept of multiplex transformation is proposed because the enhancing smooth degree of sequence is limited by adopting unitary transformation. Thirdly, the mechanism of enhancing smooth degree of sequence based on multiplex transformation is studied and the steps of improving GM (1, 1) model based on multiplex transformation is presented. Finally, the method is applied in building the model of per-capita power quantity and the results show the method presented in this paper is effective.

Keywords: Grey Model, Multiplex Transformation, Smooth Degree.

1. INTRODUCTION

Grey system is such a system which includes incomplete information about elements, structure relation and behavior. Grey system theory was first proposed by professor Deng in 1982 and has already been widely used in many fields. Although lots of achievements have been made in application of grey system theory, there are some failures in building grey models. The reason of failure is that the smooth degree of sequence can not meet necessary requirement. It is proved that the smooth degree of sequence is one of key factors in building grey model. The precision of GM(1,1),GM(n, h), Verhulst Model and SGM(1,h) mainly dependents on the smooth degree of sequence[1,2,3,4,5,6,7,8,9]. The higher the smooth degree of sequence is, the better the precision of model is. Therefore, improving the smooth degree of sequence is significant for increasing grey model's precision. In order to enhance the smooth degree of sequence, some transformations such as logarithm transformation, exponent transformation and power function transformation [10, 11, 12, 13] were put forward respectively. But unitary transformation is often not perfect in some applications.

In this paper, the method of improving GM (1, 1) model is put forward based on multiplex transformation. First, the necessary and sufficient condition of smooth sequence is given. It is proved that smooth degree of sequence can be enhanced under some conditions. Secondly, the new concept of multiplex transfer is proposed because enhancing smooth degree of sequence is limited by just adopting unitary transformation. Thirdly, the mechanism of enhancing smooth degree of sequence based on multiplex transformation is studied and the steps of improving GM(1,1) model based on multiplex transformation is presented Finally ,The method is applied in building the model of per-capita power quantity and the results show the method. Presented in this paper is effective.

2. REVIEW OF ENHANCING SMOOTH DEGREE OF SEQUENCE

Definition[14] Let $\{x^{(0)}(k), k=1,2,...,n\}$ be any non-minus date sequence, for $\forall\ \varepsilon > 0$, if there exist k_0 ,when $k > k_0$ the following inequation $\dfrac{x^{(0)}(k)}{\sum_{i=1}^{k-1} x^{(0)}(i)} = \dfrac{x^{(0)}(k)}{x^{(1)}(k-1)} < \varepsilon$ is held ,then the sequence $\{x^{(0)}(k), k=1,2,...,n\}$ is defined as smooth sequence.

Theorem 1[14] The sufficient and necessary condition for $\{x^{(0)}(k), k=1,\ 2,...,n\}$ to be smooth sequence is that function $\dfrac{x^{(0)}(k)}{\sum_{i=1}^{k-1} x^{(0)}(i)}$ is decrement with k .

* This paper is supported by the Natural Science Foundation of Zhejiang Province, P.R .China(No：602016)and Returnee Foundation of Ministry of Personal, P.R.C.(No:2003236)

Theorem2[14] For any sequence $x(k) > 0$, if non-minus transformation $F_i(x(k))$ can be expressed as follow:

$$F_i(x(k)) = x(k) \cdot f_i(k)$$

where $f_i(k)$ is non-minus and monotone descending absolutely.

Then inequality $\dfrac{F_i(x(k))}{\sum_{i=1}^{k-1} F_i(x(i))} < \dfrac{x(k)}{\sum_{i=1}^{k-1} x(i)}$ is held

If $x(k)$ is monotone ascending absolutely, several transformation function which are accordant with this theorem are given as follows:

Ⅰ. $F(x(k)) = (x(k))^{-a}$, $x(1) > 1, \alpha > 0$

Ⅱ. $F(x(k)) = (x(k))^{1/p}$, $p > 1$.

Ⅲ. $F(x(k)) = x(k)/(k+C)^{\alpha}$, $\alpha > 0, k + C > 0$

Ⅳ. $F(x(k)) = \ln x(k)$, $x(k) > 1$

Ⅴ. $F(x(k)) = x(k)/\ln(\alpha \cdot k + C)$, $\alpha \cdot k + C > 0$

Ⅵ. $F(x(k)) = a^{x(k)}$, $0 < \alpha < 1$

3. MECHANISM OF ENHANCING SMOOTH DEGREE OF SEQUENCE BASED ON MULTIPLEX TRANSFORMATION

Let any non- minus incremental sequence be denoted as:

$$Y_{(0)}^{(0)} = \left\{ y_{(0)}^{(0)}(1), y_{(0)}^{(0)}(2), \cdots, y_{(0)}^{(0)}(n) \right\}$$

then the sequence $Y_{(0)}^{(0)}$ is transfered by a chosen non-minus transformation function $F_1(x(k))$, we can get sequence

$$Y_{(1)}^{(0)} = \left\{ y_{(1)}^{(0)}(1), y_{(1)}^{(0)}(2), \cdots, y_{(1)}^{(0)}(n) \right\}$$

where $y_{(1)}^{(0)}(k) = F_1(y_{(0)}^{(0)}(k))$

If $\dfrac{y_{(1)}^{(0)}(k)}{\sum_{i=1}^{k-1} y_{(1)}^{(0)}(i)}$ is not only smooth sequence but also its smooth degree is enough high then we build model directly based on the transformed sequence $Y_{(1)}^{(0)}$

If the smooth degree of transformed sequence $Y_{(1)}^{(0)}$ is not enough high, second transformation is necessary.

According to theorem2, sequence $Y_{(1)}^{(0)}$ need to be re-transferred by a chosen non-minus transformation function $F_2(x(k))$ into sequence

$$Y_{(2)}^{(0)} = \left\{ y_{(2)}^{(0)}(1), y_{(2)}^{(0)}(2), \cdots, y_{(2)}^{(0)}(n) \right\}$$

where $y_{(2)}^{(0)}(k) = F_2(y_{(1)}^{(0)}(k))$

If $\dfrac{y_{(2)}^{(0)}(k)}{\sum_{i=1}^{k-1} y_{(2)}^{(0)}(i)}$ is not only smooth sequence but also smooth degree is enough high then we build model directly based on sequence. $Y_{(2)}^{(0)}$

If not, according to theorem2, sequence $Y_{(2)}^{(0)}$ need to be re-transfered by a chosen non-minus transformation function $F_2(x(k))$ into other sequence, and so on.

The above transfer doesn't stop until smooth degree of transformed sequence is fit for modeling.

4 THE STEPS OF IMPROVING GM(1,1)BASED ON MULTIPLEX FUNCTION TRANSFORMATION

The modeling steps of improving GM(1,1) based on multiplex transformation is as follow:

Let original non-minus sequence be

$$Y_{(0)}^{(0)} = \left\{ y_{(0)}^{(0)}(1), y_{(0)}^{(0)}(2), \cdots, y_{(0)}^{(0)}(n) \right\}$$

Supposed that $Y_{(n)}^{(0)} = \left\{ y_{(n)}^{(0)}(1), y_{(n)}^{(0)}(2), \cdots, y_{(n)}^{(0)}(n) \right\}$ is n th transferred-sequence which is enough high smooth degree.

$$\text{Let } X^{(0)} = Y_{(n)}^{(0)} \tag{1}$$

and $x^{(0)}(k) = y_{(n)}^{(0)}(k)$.

The AGO (accumulated generation operation) of sequence $X^{(0)}$ is defined as:

$$X^{(1)} = \left\{ x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n) \right\} \quad (2)$$

where $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 1,2,...,n$.

The GM(1,1) model can be constructed by establishing first order differential equation for $x^{(1)}(t)$ as :

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = u \tag{3}$$

By using least square, we can get:

$$\hat{\Phi} = \begin{bmatrix} \hat{a} & \hat{u} \end{bmatrix}^T = (B^T B)^{-1} B^T Y_n$$

$$B = \begin{bmatrix} -\frac{1}{2}\left[x^{(1)}(1) + x^{(1)}(2)\right] & 1 \\ -\frac{1}{2}\left[x^{(1)}(2) + x^{(1)}(3)\right] & 1 \\ \vdots & \vdots \\ -\frac{1}{2}\left[x^{(1)}(k-1) + x^{(1)}(k)\right] & 1 \end{bmatrix}$$

$$Y = \left[x^{(0)}(2), x^{(0)}(3), ..., x^{(0)}(n) \right]^T \quad (4)$$

Discrete solution of equation (3) is:

$$\hat{x}^{(1)}(k+1) = [x^{(1)}(1) - \frac{\hat{u}}{\hat{a}}]e^{-ak} + \frac{\hat{u}}{\hat{a}} \quad (5)$$

where $k = 1,2,...,n$.

Applying the inverse AGO, we can get

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$$
$$= (1-e^{\hat{a}})[x^{(1)}(1) - \frac{\hat{u}}{\hat{a}}]e^{-\hat{a}k} \quad (6)$$

Finally, taking n times inverse transformation for above sequence, we can get:

$$\hat{Y}_{(0)}^{(0)} = \left\{\hat{y}_{(0)}^{(0)}(1), \hat{y}_{(0)}^{(0)}(2), \cdots, \hat{y}_{(0)}^{(0)}(n)\right\}$$

5. EXAMPLE

Take modeling of per-capita power quantity as an example. Per-capita power quantity in China from 1980 to 2001 is shown in《Chinese stat. annual-2005》.

The unitary transformation adopts logarithm transformation Ⅳ of second section. The modeling by only using unitary transformation is:

$$\hat{y}_{(0)}^{(0)}(k+1) = 292.95^{e^{0.01064k}}$$

The binary function transformation adopts logarithm transformation Ⅳ and exponent transformation Ⅵ of the second section and $a = 0.9$

The modeling by using binary function transformation is:

Table 1. Modeling and forecasting results

year	Shown in state. annual	Results of unitary transfer		Results of binary transfer	
		Model Values	Relative Error (%)	Model Values	Relative Error (%)
1980	306.35	306.35	0	306.35	0
1981	311.2	311.29	0	303.39	2.51
1982	324.9	330.99	-1.88	324.84	0.02
1983	343.4	352.18	-2.56	347.82	-1.29
1984	361.61	374.97	-3.70	372.41	-2.99
1985	390.76	399.51	-2.24	398.74	-2.049
1986	421.36	425.94	-1.09	426.94	-1.32
1987	458.75	454.43	0.949	457.13	-0.35
1988	494.9	485.16	1.97	489.46	1.10
1989	522.78	518.33	0.857	524.07	-0.25
1990	547.22	554.16	-1.27	561.13	-2.54
1991	588.7	592.89	-0.71	600.8	-2.06
1992	647.18	634.79	1.91	643.29	0.60
1993	712.34	680.14	4.52	688.78	3.31
1994	778.32	729.27	6.30	737.48	5.25
1995	835.31	782.54	6.32	789.63	5.47
1996	888.1	840.33	5.38	845.47	4.80
1997	923.16	903.07	2.18	905.26	1.94
1998	939.48	971.25	-3.38	969.27	-3.17
1999	988.60	1045.4	-5.74	1037.8	-4.98
2000	1073.6	1126.1	-4.89	1111.2	-3.50
2001	1164.3	1213.9	-4.26	1189.8	-2.19
2002*	1287.6	1309.7	-1.72	1273.9	1.06

***:** forecasting results

$$\hat{y}_{(0)}^{(0)}(k+1) = e^{(5.6466 - 0.1054e^{0.0072})}$$

The results of models of unitary and binary transformation are shown separately in table 1.

From table 1, it is easily to see that model precision by applying binary transformation is much better than model precision by using unitary transformation. According to theorem2, as long as appropriate transformation function can be chosen, the modeling precision is increased with multiplex transfer times. It can effectively improve the model precision

6. CONCLUSIONS

Different transformation functions have different characters. Utilizing multiplex transformation can effectively take advantages of their characters and is helpful to improve model's precision .One of difficult problems is how to choose transformation functions and how to combine them. There are still a lot of research works worthy to do.

7. REFERENCES

[1] C. S. Cheng, Y. S. Hsu et al, "Grey neural network", IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol.81, No.11, November 1998, pp.2433-2442.

[2] W. Z. Dai, J. F. Li, "Grey Verhulst Forward Neural Network Model and Its Application", IEEE Proceeding of The 5th World Congress of Intelligent Control and Automation., Vol.6, No.2, June 2004, pp.204-206.

[3] R. S. Liang, "Application of grey linear Programming to short-term hydro scheduling", Electric Power Systems Research, Vol.41, No.3, March 1997, pp.159-165.

[4] S. H. Huang, C. L. Huang, "Control of an inverted Pendulum using grey Prediction model", IEEE Transactions on Industry Applications, Vol.35, No.2, February 2000, pp.452-458.

[5] Y. P. Huang, T. M. Yu, "Hybrid grey-based models for temperature prediction", IEEE Transactions on Systems, Man, and Cybernetics, Part B：Cybernetics, Vol.27, No.2, February 1997, pp.284-292.

[6] H. Chenghsiung, "Grey neural network and its application to short term load forecasting problem", IEICE Transactions on Information and Systems, Vol.85, No.5, May 2002, pp.897-902.

[7] J. F. Li, W. Z. Dai, "Multi-stage Grey Modeling Method and Its Application in Chinese GDP Modeling Based on Stepwise ratio Data Processing", SYSTEMS ENGINEERING-THEORY & PRACTICE, Vol.24, No.9, September 2004, pp.98-101.

[8] J. F. Li, W. Z. Dai, "A New Approach of Background Value- Building and It's Application Based on Data Interpolation and Newton-Cores Formula", SYSTEMS ENGINEERING-THEORY & PRACTICE, Vol.24, No.10, October 2004, pp.122-126.

[9] J. F. Li, W. Z. Dai, "Research on the Ameliorating GM(1,1) Model and Its Application in the Power

Quantity Modeling of Shanghai City", SYSTEMS ENGINEERING-THEORY& PRACTICE, Vol.25, No.3, March 2005, pp.124-128.

[10] Q. Li, "The Further Generalization for Grey Forecasting Model", SYSTEMS ENGINEERING-THEORY & PRACTICE, Vol.13, No.1, January 1993, pp.64-66.

[11] B. He, Q. Meng, "Study on Generalization for Grey Forecasting Model", SYSTEMS ENGINEERING THEORY & PRACTICE, Vol.22, No.9, September 2002, pp.138-141.

[12] J. F. Li, W. Z. Dai, "GM(1,1) Model Based on Power Function x^{-a} Transformation and Its Application in Chinese Country Per-capita Housing Areas Modeling", SYSTEMS ENGINEERING- THEORY & PRACTICE, Vol.24, No.11, November 2004, pp.63-67.

[13] F. Y. Huang, "Transformation Methods of Grey System Modeling", SYSTEMS ENGINEERING-THEORY & PRACTICE, Vol.14, No.11, November 1994, pp.35-38.

[14] T. J. Chen, "A New Development of Grey Forecasting Model", SYSTEM ENGINEER, Vol.8, No.7, July 1990, pp.50-52.

A New Denoising Algorithm Based on Fuzzy Mathematical Morphological Operations with One Parameter

Qi-sheng YAN
School of Mathematics & Information Science, East China Institute of Technology
Fuzhou, Jiangxi 344000,China
Email:yanqs93@126.com

ABSTRACT

In order to enhance the robustness of morphological image processing as well as the performance in the anti-disturbance, a new denoising algorithm with one parameter based on fuzzy mathematical morphological basic operations has been presented, and then some initial analysis for the selection of parameter are given. The method is applied in denoising for binary image with noise. The experimental results demonstrate that the noise can be nearly removed by using the new method and the detail of original image can be kept clearly with the clear edge, thus its performance is better than the classical morphological filter. In addition, the method has the feature of flexibility and better practicability due to containing an adjustable restricted parameter.

Keywords: Fuzzy morphology, Mathematical morphology, Parameter, Filtering Algorithm

0 INTRODUCTION

Mathematical morphology, which is based on set-theoretic concept, provides an efficient tool to image processing, and analysis [1,2,3,4]. It is widely adopted in computer vision, signal processing, image analysis, pattern recognition, and it includes almost all the concerned contents of the image processing, such as nonlinear filtering, texture analysis, biological material analysis, edge detection, feature extraction, image compression, image segmentation and face recognition etc.

The classical filter based on the standard morphology has the better performance for denoising, at the same time it vagues the edge information of image, the robustness of morphological image processing as well as the performance in the anti-disturbance which is in the need of development. Since the efficiency of the morphological filter is mainly up to the following two factors: One is the definition of basic morphological operators (dilation and erosion) and the structuring element. The other is the selection of structuring element and the construction of composite operators. To the former, we can develop it in two ways: one is to divide the structure element into two parts: the core, the pixels participate with weights greater than one, and the soft boundary, the pixels participate with weights equal to one, then get the soft morphological operators [5]. The other is to define a class of morphological operations such as regulated morphological operations by extending the fitting interpretation of the ordinary morphological operations, which have a controllable strictness, and so they are less sensitive to noise and small intrusions or protrusions on the boundaries of shapes [6]. In addition, due to the complexity of image information and the strong relevance among them, there are probably some incompletion and inaccuracy under various situations during the process, so the better processing performance may be obtained by applying fuzzy set theory to the image processing and understanding in some situation. In this paper, a fuzzy morphological method with one restricted parameter (r-MM) is presented and the method is applied to denoising for binary image, experimental results demonstrate that it is possible to improve the results in denoising by using r-MM instead of the ordinary operations.

1 FUZZY MATHEMATICAL MORPHOLOGY FOR BINARY IMAGE

Fuzzy mathematical morphology (FMM) is an efficient method for image processing and analysis combined the mathematical morphology and fuzzy logic. Its initiate idea is to treat the image as a fuzzy set, and further to carry out the fuzzy operation on the fuzzy set, the according fuzzy morphology operation is fuzzy dilation, fuzzy erosion, and fuzzy open and fuzzy close. In the past several attempts have been made to apply fuzzy set theory to mathematical morphology. These attempts have resulted in different approaches and definitions. In this paper the approach described by Sinha and Dougherty [7] has been used. The operations of erosion and dilation of a fuzzy image by a fuzzy structuring element having a bounded support, are defined in terms of their membership functions as

$$\mu_{E_B(A)}(x) = \min_{y\in B}[\min[1,1+\mu_A(x+y)-\mu_B(y)]] = \min[1,\min_{y\in B}[1+\mu_A(x+y)-\mu_B(y)]] \quad (1)$$

$$\mu_{D_B(A)}(x) = \max_{y\in B}[\max[0,\mu_A(x-y)+\mu_B(y)-1]] = \max[0,\max_{y\in B}[\mu_A(x-y)+\mu_B(y)-1]] \quad (2)$$

where $x, y \in Z^2$ are the spatial co-ordinates and μ_A, μ_B are the membership functions of the image and the structuring element, respectively. it is obvious from Eq. (1) and Eq.(2) that the result of both fuzzy erosion and dilation have membership functions whose values are restricted to the interval [0,1].

2 RMM AND THE DEFINITION OF MEMBERSHIP FUNCTIONS

2.1 Basic operations of RMM

In the operations of classical binary morphology, the binary dilation collects shifts for which the kernel set intersects the object set without taking into account what is the size of the intersection, whereas the binary erosion collects shifts for which the kernel set is completely contained within the object set without considering shifts for which some kernel elements are not contained with the object set. As a result of these strict approaches, the ordinary morphological operations are sensitive to noise and small intrusions or protrusions on the boundary of shapes. By using the fitting

interpretation of the definition of ordinary dilation or erosion and introduce a restricted parameter r, regulated dilation and regulated erosion are defined by [6]:

$$E^r_{\overset{\vee}{B}}(X) = A \overset{r}{\Theta} B \equiv \left\{ x \middle| \#(A^c \cap B[x]) < r \right\},$$

$$D^r_B(X) = X \overset{r}{\oplus} B = \left\{ x \middle| \#(X \cap B[x]) \geq r \right\}, \quad r \in [1, \#B],$$

where A is a set of binary image object pixels and B is a set of binary structuring element pixels, $B[x]$ is denoted as the translation of B along the vector x, that is

$B[x] = \{ b + x \mid b \in B \}$, $\overset{\vee}{B}$ is the reflected or symmetric set of B given by $\overset{\vee}{B} = \{ -b \mid b \in B \}$,the symbol # denotes the cardinality of a set.

Obviously, when r is equal to 1,the regulated dilation or erosion results in the ordinary dilation or erosion:

$$E^1_{\overset{\vee}{B}}(X) = \{ x \in U \mid B[x] \subset X \}$$

$$D^1_B(X) = \{ x \in U \mid B[x] \cap X \neq \varnothing \}$$

where U is a set of binary image pixels.

Taken into account actual application, the following discussion is on the premise that set B is equal to $\overset{\vee}{B}$ (that is $B = \overset{\vee}{B}$).

2.2 Definition of membership functions

The sets are described by using membership functions in the fuzzy set theory. Let U be a nonempty set called a universe, any mapping $\mu_A : U \to [0,1]$ or $\mu \to \mu_A(u)$ determines a fuzzy subset A of U and μ_A is called as the membership degree of A. Given a fuzzy set A, for any $\lambda \in [0,1]$, The level cut set of A is defined as $A_\lambda = \{ x \in U \mid \mu_A(x) \geq \lambda \}$, it is an important concept in the transform between fuzzy sets and ordinary sets. Let degree of belonging of all the elements in U to set X as the membership degree of this element.

For any $u \in U$, we define the membership degree of u in X as $\dfrac{\#\{B[x] \cap X\}}{\#B[x]}$, (3)

Obviously, for any $u \in U$, we get that

$$0 \leq \frac{\#\{B[x] \cap X\}}{\#B[x]} \leq 1$$

Further we can conclude the fuzzy set

$$F^B_X = \left\{ (u, \mu_{F^B_X}(u)) \mid u \in U, \mu_{F^B_X}(u) = \frac{\#\{B[x] \cap X\}}{\#B[x]} \right\}$$

in U (4)

where $\mu_{F^B_X}(u)$ is the membership degree of u belonging F^B_X.

3 DISCUSSION ON RESTRICTED PARAMETER

Rewriting the expression of the regulated dilation to the following:

$$(A \overset{r}{\oplus} B)(x,y) = \begin{cases} 1, & if\ \#(A \cap \overset{\vee}{B}_{(x,y)}) \geq r \\ 0, & otherwise \end{cases} \quad (5)$$

where $r \in [1, \#B]$ (assume that #B<#A, otherwise let $r \in [1, \min(\#A, \#B)]$)

So we can get the equivalent definition of the regulated dilation taking into account the symmetry of B.

Definition 1. $(A \overset{r}{\oplus} B)(x,y) =$

$$\begin{cases} 1, & if\ \exists B1 \subseteq B\ st.\ \ \#(B1) = r \wedge (A \Theta B1)(x,y) = 1 \\ 0, & otherwise \end{cases}$$

Definition 1 simply states that some r-point subset of B must fit with A when B is centered at (x, y). Rewriting definition 1 using logical "or":

Definition 2. $(A \overset{r}{\oplus} B)(x,y) = \underset{\substack{B1 \subseteq B \\ \#(B1)=r}}{\vee} (A \Theta B1)(x,y)$

Applying above equation to all pixels in the image, we get:

Definition 3. $$A \overset{r}{\oplus} B = \underset{\substack{B1 \subseteq B \\ \#(B1)=r}}{\vee} A \Theta B1 \quad (6)$$

This equation expresses regulated dilation using only crisp erosions and logical "or". The number of crisp erosions needed on the right-hand side of the equation depends on the number of structuring elements B1, satisfying the condition: $B1 \subseteq B \wedge \#B1 = r$. That is

$$C_r^{\#B} = \binom{\#B}{r} = \frac{\#B!}{r!(\#B - r)!}$$

$C_r^{\#B}$ increase as $\#B$ grows and is maximized when $[\frac{\#B}{2}] = r$. Despite the combinatorial growth, the number of erosions is manageable when $\#B$ is relatively small or the value of r is near 1 or $\#B$. Actually, in the typical regulated dilation or erosion operations, r does not usually deviate greatly from 1 or #B.

Theorem 1. $A \overset{r}{\oplus} B =$

$$\begin{cases} \underset{\substack{B1 \subseteq B \\ \#B1 = r}}{\vee} A \Theta B1, & if\ r \leq \#B - r + 1 \\ \underset{\substack{B1 \subseteq B \\ \#B1 = \#B - r + 1}}{\wedge} A \oplus \overset{\vee}{B}1, & otherwise \end{cases} \quad (7)$$

Proof. （i）when r is smaller, theorem 1 can be derived directly from above analysis and the Eq.(6).

（ii）when r is bigger, the properties of the combination function yield:

$$\binom{\#B}{r} = \binom{\#B}{\#B - r} \Rightarrow C_r^{\#B} = C_{\#B-r}^{\#B}$$

Thus, thresholds of either r or $\#B$-r require the same number $C_r^{\#B}$ of erosions. However, for large values of r, the crisp erosions are more expensive since the structuring elements contain many points. That is, when r is large, eroding by $C_r^{\#B}$ structuring elements contain r points is more expensive than eroding by $C_{\#B-r}^{\#B}$ structuring elements containing only $\#B$-r points.

Since

$$(A \overset{r}{\oplus} B)(x,y) = \begin{cases} 1, & if\ \#(A \cap \overset{\vee}{B}_{(x,y)}) \geq r \\ 0, & otherwise \end{cases} \quad \Leftrightarrow$$

$$(A\overset{r}{\oplus}B)(x,y)=\begin{cases}0, & if\ \#(A\cap B_{(x,y)})<r\\ 1, & otherwise\end{cases}\quad\Leftrightarrow$$

$$(A\overset{r}{\oplus}B)(x,y)=\begin{cases}0, & if\ \#(A^c\cap B_{(x,y)})\geq \#B-r+1\\ 1, & otherwise\end{cases}$$

$$\Leftrightarrow (A\overset{r}{\oplus}B)(x,y)=$$

$$\begin{cases}0, & if\ \exists B1\subseteq B\ st.\ \#(B1)=\#B-r+1\wedge(A^c\Theta B1)(x,y)=1\\ 1, & otherwise\end{cases}$$ t

hen $A\overset{r}{\oplus}B=(\underset{\substack{B1\subseteq B\\ \#B1=\#B-r+1}}{\vee}(A^c\Theta B1))^c$

Using DeMorgan's law and the principle of operator duality, the above equation becomes:

$$A\overset{r}{\oplus}B=\underset{\substack{B1\subseteq B\\ \#B1=\#B-r+1}}{\wedge}(A^c\Theta B1)^c=\underset{\substack{B1\subseteq B\\ \#B1=\#B-r+1}}{\wedge}A\overset{\vee}{\oplus}B1$$

That is when $r>\#B-r+1$, the process is efficient using above equation.

4 FILTERING ALGORITHM BASED ON FUZZY MATHEMATICAL MORPHOLOGY WITH THE PARAMETER *r*

A restricted parameter r is introduced and $\mu^r_{F^B_X}(u)$ is defined by:

$$\mu^r_{F^B_X}(u)=\begin{cases}\dfrac{r}{\#(B[x]\cap X)}, & if\ u\in D^r_B(X)\\ 0, & if\ u\notin D^r_B(X)\end{cases}$$

where $r\in[1,\#B]$ (8)

Since $r\leq\#\{B[x]\cap X\}\leq\#B$, assume that $x\in D^r_B(X)$, we get that $\dfrac{r}{\#B}\leq\dfrac{r}{\#(B[x]\cap X)}\leq 1$,

Thus for given r, it follows that

If $x\in D^r_B(X)$,then

$\mu^r_{F^B_X}(u)\in[\dfrac{r}{\#B},1]\subset[0,1]$,otherwise $\mu^r_{F^B_X}(u)=0$

for the regulated erosion operation with restricted parameter,

since $X\overset{r}{\Theta}B=\{x\,|\,\#(X^c\cap B[x])<r\}$, $r\in[1,\#B]$,

or as $E^r_B(X)=\{x\in U\,|\,\#(B[x]\cap X)<r\}$

if the complement of r relative to the set B is defined by $\bar{r}_B\equiv\#B-r+1$, then we can get that

$$A\overset{r}{\Theta}B=\{x\,|\,\#(A\cap B[x])\geq\bar{r}_B\}$$

and $A\overset{r}{\oplus}B=A\overset{\bar{r}_B}{\Theta}\overset{\vee}{B}$, $A\overset{r}{\Theta}B=A\overset{\bar{r}_B}{\oplus}\overset{\vee}{B}$

and if $r=(\#B+1)/2$ then $A\overset{r}{\oplus}B=A\overset{r}{\Theta}B$.

So the definition is as following:

$$\mu^r_{F^B_X}(u)=\begin{cases}\dfrac{r}{\#(B[x]\cap X)}, & if\ u\in E^{\bar{r}_B}_B(X)\\ 0, & if\ u\notin E^{\bar{r}_B}_B(X)\end{cases},$$

where $r\in[1,\#B]$ (9)

Therefore, we can get the procedures of the calculation of the erosion and dilation for binary image using fuzzy morphology with the parameter r:

Step1. Chose a appropriate restricted parameter r.

Step2. Calculate the result of the original image's fuzzy set F^B_X and then the membership degree of every pixel from Eq. (8).

Step3. Calculate the fuzzy morphological erosion and dilation, and get the membership degree of every pixel.

Step4. Let threshold $\lambda=0.5r/(\#B)$ in view of level cut set, to turn the fuzzy set into ordinary set, when the membership degree is higher than $0.5r/(\#B)$, the .set. is viewed as object, when lower than $0.5r/(\#B)$, then as background.

5 EXPERIMENTAL RESULTS

In order to demonstrate the advantage of the new denoising algorithm based on fuzzy mathematical morphological operations with one parameter, we compared it with the ordinary mathematical morphology in noisy image processing. And we applied open-close composite operators to realize the bilateral filtering. Fig. 1. presents an example of morphological processing of a noisy image by using our method and the ordinary morphological operations respectively. The structuring element used in these operations is a 5-points diamond flat structuring element with the origin at its center, and their membership degree is 1.where r is selected as 3, $\lambda=0.5r/(\#B)$.

(a) original image X

(b) X corrupted with 10% pepper and salt noise

(c) filtering result using ordinary morphological method

(d) filtering result using our method

Fig. 1. Comparative effect of two filtering methods

From the above experiment we can observe that the noise can be nearly removed by using our method and the detail of original image can be kept clearly with the clear edge, thus its performance is better than the classical morphological filter. In addition, it contains an adjustable restricted parameter, thus it has the feature of flexibility and better practicability.

6 CONCLUSIONS

In this paper, we studied a new denoising algorithm based on fuzzy mathematical morphological operations with one parameter, and tried to apply the method in binary image filtering with great experimental success, it is to some degree development of the classical mathematical morphological method. But how to apply this method widely in the actual image like grey, colorful image etc and wider fields in image processing is our next step works. In addition, in the regulated morphological operations, besides the construction of composite operators, structuring element and restricted parameter is the most basic and important concept, in this paper the choice of size and shape of structuring element and parameter r and threshold λ will directly influence the result of image processing.

7 REFERENCES

[1] J. Serra, Image Analysis and Mathematical Morphology, New York: Academic Press, 1982.

[2] R.M. Haralick, Sternberg, et al, Image analysis using mathematical morphology, IEEE Trans.Pattern Anal.Mach.Intell. 1987, 9(7):532-550.

[3] F.Y. Shih, Y. T. Wu, Decomposition of binary morpho-logical structuring elements based on genetic algorithms, Computer Vision and Image Understanding. Vol.99, 2005, 291-302.

[4] F. Y. Shih, P. Puttagunta, Recursive soft morphological filters, IEEE Trans. Image Process. 1995, 4（7）:1027-1032.

[5] L. Koskinen. SoftMorphological Filter. Proc. Proceedings of SPIE-The International Society for Optical Engineering, 1991, 1568(7): 262-270.

[6] G. Agam, Its'hak Dinstein. Regulated morphological operations, Pattern Recognition, Vol.32, 1999, 947-971.

[7] D. Sinha, E. R. Dougherty, Fuzzy Mathematical Morphology.Journal of Visual Commuication and Image Representation, 1992, 3(7): 286-302.

Application of PSO Arithmetic in a Greenhouse Neural Network Expert System

Hong Liu
Computer and Information Engineering College, Zhejiang Gongshang University
HangZhou, 310035,China
Email: llh@mail.hzic.edu.cn

ABSTRACT

Aim at such drawbacks of neural network expert systems as slower convergence pace, longer learning time, etc. Based on an intelligent greenhouse neural networks expert system, PSO arithmetic is applied in its learning process and forecasting process. Simulation results demonstrate the efficiency and the effectiveness of the proposed method, the PSO arithmetic combined with neural networks, can improve the convergence pace and learning ability of neural network, and it has entirely stronger convergence ability.

Keywords: PSO Arithmetic, Expert System, and BP Neural Network.

1. INTRODUCTION

With the more and more research about the technology of intelligence information processing recently, the technique of combining artificial intelligence expert system with greenhouse system gets more and more attention. People pay more attention to how to regulate some environmental gene, such as temperature, humidity, ventilation, illumination in the greenhouse for the request of all kinds of crops growth. Literature [1,2,3,4] discussed application of expert system in the greenhouse crops production controlling atomization. If experts' experience and knowledge can be taken full advantage in the crops production, the crops production level and economy benefit can got improvement. In the literature mentioned before, working procedure and running rules of greenhouse controlling are organized by automation machine theory and production rule denotation form based on conventional artificial intelligence expert system theory, and greenhouse controlling sequences dynamically come into being, which on a certain extent improves greenhouse crops growth. However, as a symbolic transacting system of simulating artificial abstract thought, expert system has some drawbacks such as knowledge acquiring bottleneck, poor capability in association, reasoning and lower intelligent level, etc. so expert system is applied combined with the other technology, such as neural network. Aim at the obvious advantages of neural network as knowledge expressing, knowledge acquiring, self-adjustability learning, ideal reasoning, etc; integration of neural networks and expert system can redress the drawbacks of conventional expert system. But the BP neural network has some drawbacks such as slower convergence pace, longer learning time, etc.

In the paper, PSO arithmetic is introduced into the neural network expert system, with it to train the neural network of expert system. Aim at the obvious advantages of PSO arithmetic as stronger convergence ability; integration of neural networks and PSO in the expert system can redress the drawbacks of neural network. In the paper, based on an intelligent greenhouse neural networks expert system, PSO arithmetic is applied in its learning process and forecasting process. Simulation results demonstrate the efficiency and the effectiveness of the proposed method. The paper firstly sets forth system structure of the greenhouse expert system based on BP neural network, and then introduces the designing of BP neural networks system, the designing of PSO arithmetic, and in the end, discusses and simulates its learning process and forecasting process.

2. SYSTEM STRUCTURE

Intelligent greenhouse controlling means that based on climate gene such as temperature, humidity, ventilation, illumination outside environment, and enactment of the greenhouse expert system and user parameter, right greenhouse environment (which is confirmed by planning and optimization according to all kinds of crops request) for crops growth is created by accommodating temperature, humidity, ventilation, illumination in the greenhouse, that is to say, it makes a checking standard, which compares desired data with measured data got by real-time for checking greenhouse environment and modulates all kinds of controlling machines' state to fulfill environment gene require of sorts.

Based on previous greenhouse controlling require, the paper develops an intelligent greenhouse expert system based on back propagation neural networks (showed as Fig. 1).

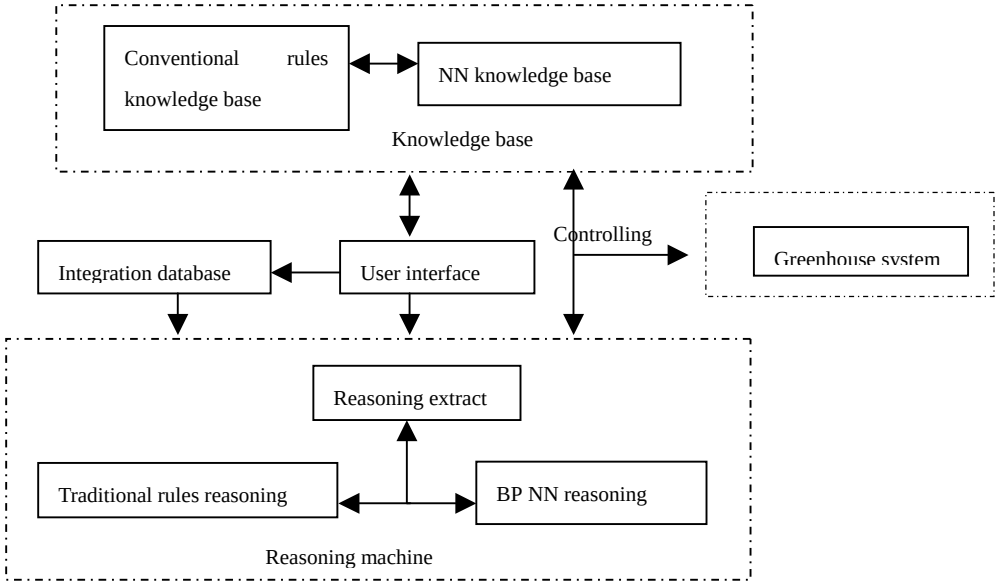


Fig. 1. System structure of an intelligent greenhouse expert system based on neural networks

In the Fig.1, the most important two function models are knowledge base and reasoning machine. Knowledge base describes specialty knowledge of greenhouse management, which is consisted of fact set, experience database. The fact set includes knowledge of controlled object, and self-adjusting controlling rules and parameter self-modulating rules. The experience database includes specialty knowledge extracted from expert experience in greenhouse management and work technique, and experience datum, such as parameter change scope and adjusting scope of controlled object, etc.

Reasoning machine includes two sub-reasoning machine: traditional reasoning machine and BP neural network reasoning machine. It extracts current state of controlling object from greenhouse state base by management task and object, compares such state with object state asked, and after controlling object is fixed on, produces relevant controlling order by exacting relevant machine name from integrate database based on known rules.

3. KEY TECHNOLOGIES

3.1 BP Neural Network Designing

BP (Back Propagation) network is a former direction feedback network adopting error counter-propagation as learning algorithmic in the neural network, it usually is consist of input layer, output layer and hidden layer, whose nerve between layer and layer is linked by full interlink age through relevant net power coefficient w, while never in every layer has no link. BP network can be looked as a kind of highly non-linear mapping F from input to output, in the mapping topology keeps invariability, which can be simply described as

$$y = F(x) \quad F : R^n \to R^M \qquad (1)$$

When applied in the forecasting, BP network needs a network learning process, which gets every nerve link power w and clique value by self-adapting and self-organizing input learning samples. By training, networks have abilities of memory and association for learning samples. The network learning process includes two repeatedly alternant processes: information former-propagation process and error reverse propagation process.

For depicting and stimulating domain expert knowledge more rightly and realistically, neural network uses subjection function to depict expert knowledge, the function can be showed as: $F_A(x) : x \to [0,1]$, whose value is a real number from 0 to 1, and shows degree of x being part of a set.

The BP learning algorithm in the paper introduces an improved form, which showed as:

$$f(x) = \frac{\sum_{l=1}^{M} y^{\bar{l}} \prod_{i=1}^{n} \exp[-[\frac{x_i - x_i^{\bar{l}}}{o_i^l}]]^2}{\sum_{l=1}^{M} \prod_{i=1}^{n} \exp[-[\frac{x_i - x_i^{\bar{l}}}{o_i^l}]]^2} \qquad (2)$$

Where $x = (x_1, x_2, \cdots, x_n)^T$ is unit input, $f(x)$ is unit output, $x_i^{\bar{l}}$ is a center point which is in the illegibility sets of the lth input and the ith variable, o_i^l is its' width, and $l = 1,2,\cdots,M$ is the number of rules. The other parameters in the BP neural network are shown as:

1) In the input layer, there are five variables as system input, such as environment temperature, sun radiation, wind speed, relative humidity, and greenhouse temperature.

2) In the output layer, there is a variable as system

output, which is greenhouse temperature.

3) Because of real datum got from greenhouse limited, some of the learning swatch in the BP neural network is from the real greenhouse datum, while the others are got by system simulation experiment g based on the real greenhouse datum.

4) System error uses square error function: $E = \frac{1}{2}\sum_{k=1}^{N}(y_k - \hat{\gamma}_k)^2$, where $\hat{\gamma}_k$ is network real output.

3.2 PSO ARITHMETIC DESIGNING

Particle swarm optimization (PSO) is an evolution computation technique developed by Kennedy and Eberhart, with roots in the preying of large number of birds or fish. The underlying rules of cooperation and competition within social swarms give it good capability to make global optimization with the help of memory rather than to simply random search in a certain area. So it has a better chance to fly into a better solution quickly than some previous optimizations have in addition to its better performance.

In the PSO arithmetic, the potential solutions is defined as particles, firstly, the particles are initialized with a population of random solutions. The arithmetic searches for the optimum by updating generations. The update of the particles is accomplished by the following equations Eq. (1) calculates a new velocity for each particle (potential solution) based on its previous velocity (v[]), the particle's location at which the best fitness so far has been achieved(pbest), and the population global location(gbest for global version) at which the best fitness so far has been achieved. Eq. (2) updates each particle's position in the solution hyperspace.

$$v[] = w*v[] + c_1*rand()(pbest[] - present[]) + c_2*rand()*(gbest[] - present[]) \quad (3)$$

$$present[] = present[] + v[] \quad (4)$$

In the Eq. (1) and Eq. (2), $v[]$ means the velocity for each particle, $present[]$ means the current location for each particles, $pbest[]$ and $gbest[]$ mentioned before, $rand()$ is got random data between 0 and 1; c_1 and c_2 are learning gene, generally, c_1 is equal to c_2, their value is 2. The use of the inertia weight w has provided improved performance in a number of applications. If the value of w is larger, it means the arithmetic have stronger convergence ability.

This paper adopts PSO algorithm to train the neural network. First of all, set all the linking weights of all the neural in the greenhouse system as individuals represented by strings of real numbers, so called particles. Suppose the network contains M optimized weights (including the threshold value), then every particle will be represented as a M-dimension vector composed by M weight parameters. Every component in the particle structure represents a weight value of the neural network. The domain of the weight is set as [-2,2], and the variance of the real output and expected output of the network is defined as the fitness function. The PSO is used to search the optimal particle positions, namely the optimal weights of the neural network, minimizing the variance.

1) Initializing particle swarm

For the scale of the particle swarm, a number of particles are randomly generated according to the particle structures to compose the population, where various particles represents a group of different weights of the neural network. At the same time, initialize Gbest and Lbese.

2) Network training (particle evaluation)

Every components of the particle swarm reflects the weight of the network, so constructs the entire neural network. For every particle, input the training samples. The optimization process of the network weights is an iterative process. To guarantee high generalization of the network in training, the given population space is divided into two parts: one is for training, called the training set; another is for testing, called the testing set. In the optimizing process, for each training set, the given samples should be categorized to make training with different samples. Compute the variance of every network with the training set, and then the variance would be the target function to construct the corresponding fitness function, in order to calculate particle fitness.

3) PSO model calculation

Evaluate every particles (every particle is considered as a flyable particle), and locate the optimal particle to determine whether Gbest and Lbest should be updated. After that, update the flying speed of different components of every particle according to the PSO model, and generate new particles.

4) The end of the algorithm

When the target function (the variance) is below the given value, then the algorithm stops.

4. SIMULATION RESULTS

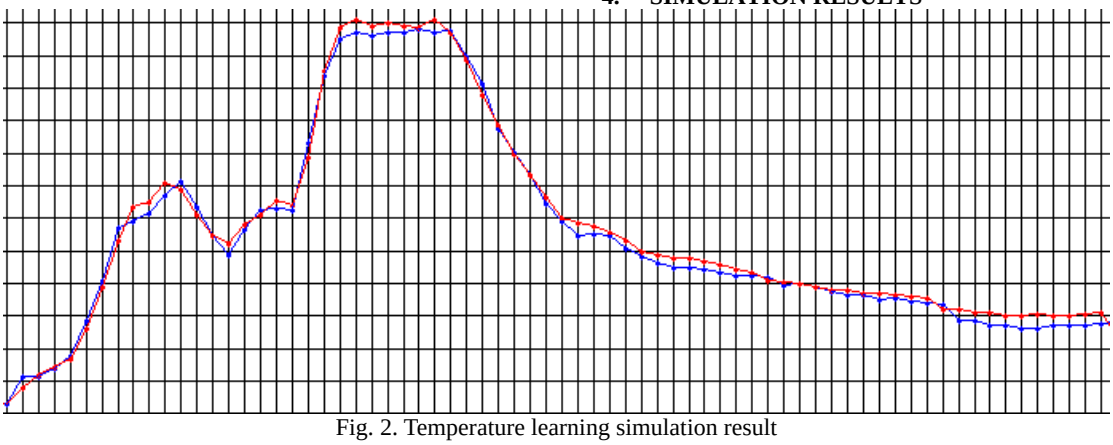


Fig. 2. Temperature learning simulation result

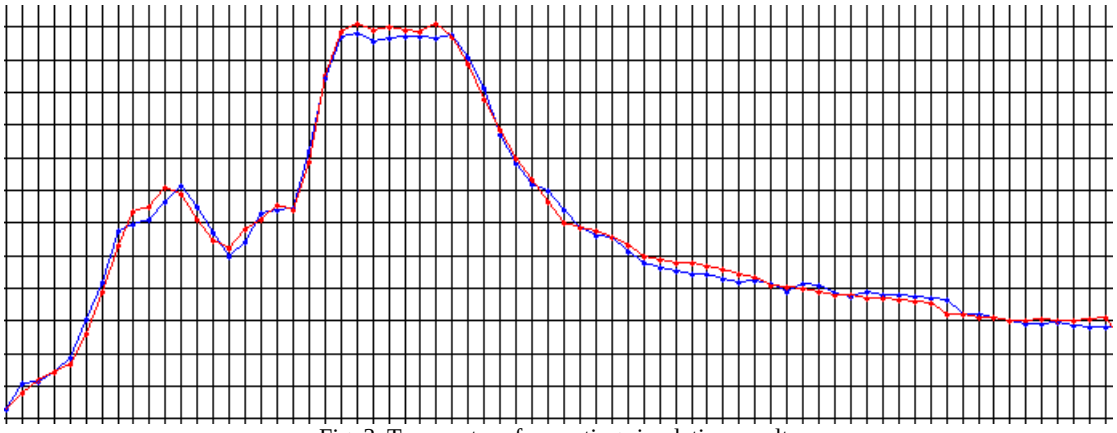


Fig. 3. Temperature forecasting simulation results

In the Fig.2 and Fig.3, red curve is scheduled by the system, and blue curve is simulation result. Fig.2 shows temperature learning stimulation results; it shows temperature-learning result good and curve simulation degree loftily. And Fig.3 shows temperature forecasting simulation results, it can be concluded that precision of the forecasting simulation results is high.

5. CONCLUSION

From the simulation results above, PSO arithmetic can be convergent, and its' convergence pace is repaid, its' capability is very good. Introduced into the greenhouse neural network expert system, it is demonstrated of the efficiency and the effectiveness of the proposed method, which can realize expert system self-learning and self-organizing which is efficient in overcoming knowledge acquiring bottleneck.

6. REFERENCES

[1] J.J. Li, Z.R. Shen, "Preliminary Study on Expert System for Tomato Long Season Cultivation in Greenhouse", Transactions of the CASE, 2003, 19(3): 267-268.

[2] X.L. Zheng, "Expert Control Systems for Modern Greenhouse Climate", Mechanical & Electrical Engineering Magazine, 2003,20(3): 42-45.

[3] T.F. Shen, B. Li, "An Expert Control Platform for Greenhouse", Journal of Shanghai University (Natural Science), Oct 2001, 7(5): 384-387.

[4] L.P. Long, K.Y. Zhang, "The Research on Real-Time Expert Control System for Greenhouse Heating System", Mechatronics, 2003, No.1: 27-29.

[5] Kassabalidis I N, Sharkawi M A, Marks R J, et al, "Dynamic Security Border Identification Using Enhanced Particle Swarm Optimization", IEEE Trans on Power Systems, 2002, 17(3): 723 -729.

[6] X.F. Li, J.P. Xu, "The Establishment of Self-adapting Algorithm of BP Neural Network and Its Application", Systems Engineering-Theory & Practice, 2004, No.5: 1-8.

[7] B.F. Vanden, A.P.A. Engelbrecht, "Cooperative Approach to Particle Swarm Optimization", IEEE Transon Evolutionary Computation, 2004, 8(3): 225-239.

[8] D.L. Chen, Y. Cheng, "Application Research of the ESS Based on Neural Networks and Rules", Computer Engineering & Science, 2004, 26(6): 70-72.

[9] K.E. Parsopoulos, M.N. Vrahatis, "On the Computation of All Global Minimizers through Particle Swarm Optimization", IEEE Trans on Evolutionary Computation, 2004,8(3): 211-224.

[10] C.Y. Zhou, W.S. Wang, "Fuzzy Adaptive Controller Based on BP Algorithm", Journal of Beijing University of Aeronautics and Astronautics, 1998, 24(3): 291-294.

Predicting PPII by Artificial Neural Network with a Modified Sequence Coding Method*

Kezhong Lu[1], Lanjuan Sun[3]
Department of Computer Science, Chizhou Teachers College, Chizhou 247000,China
E-mail: luke76@163.com
E-mail: sunlanjuan@yahoo.com.cn
Wenbo Xu[2]
School of Information Technology, Southern Yangtze University, Wuxi 214036, China
E-mail: xwb@sytu.edu.cn

ABSTRACT:

Polypro line type II (PPII) secondary structures are somewhat rare on proteins. Few works have been made to predict PPII with machine learning approaches. This paper predicted PPII secondary structure with BP-model neural network based on preprocessing protein sequences, and compared two numerical sequence coding methods in PPII, the traditional binary vector sequence coding method and the modified binary vector sequence coding method allowing for neighboring amino-acid residues' effect. Results show that the modified sequence coding method is better than the traditional sequence coding method in predicting PPII; the sensitivity can reach 82.5%.

Keywords: Artificial Neural Network, BP-model, Protein Secondary Structure, and Prediction of Polypro line Type II.

1. INTRODUCTION

The protein-folding problem is a key initiative in the fields of protein engineering and drug design. Traditionally, the protein structure has been predicted using methods such as x-ray crystallography, which involves manual and experimental procedures. However due to the slow velocity, it is becoming impossible for these approaches to keep up with the number of protein sequences needed to be mapped. As a result the existence of an artificially intelligent process of sequencing is becoming an imminent necessity. Currently artificial neural networks (ANN) have been used intensively in bioinformatics. Artificial neural networks have several unique characteristics and advantages as tools for the molecular sequence analysis problem [1]. A very important feature of these networks is their adaptive nature, where learning by example replaces conventional programming in solving problems. This feature makes such computational models very appealing in application domains where one has little or incomplete understanding of the problem to be solved, but where training data are readily available. Owing to the large number of interconnections between their basic processing units, artificial neural networks are error-tolerant, and can deal with noisy data. Artificial neural network architecture encodes information in a distributed fashion. This inherent parallelism makes it easy to optimize the network to deal with a large volume of data and to analyze numerous input parameters. Flexible encoding schemes can be used to combine heterogeneous sequence features for network input. Finally, a multiplayer network is capable of capturing and discovering high-order correlations and relationships in input data.

Polypro line type II (PPII) secondary structure is a kind of rare protein secondary structure. This structure plays a central role in numerous vital processes including signal transduction, transcription, cell motility, and the immune response [2, 3]. So far few works have been taken to predict PPII secondary structure computationally. Siermala M et al. developed a method on the basis of feed-forward multilayer perception neural networks with the back propagation (BP) learning algorithm to predict PPII secondary structure in 2000[4]. The method shows artificial neural network is suit to predict PPII secondary structure, but the result is not ideal even with uniformly distributed test set. In this paper, first, we briefly introduce the artificial neural network model. Then describe how the sequence data are chosen. A sliding windowing technique of sequences is tested to localize occurrences for the PPII class. Finally, results of our artificial neural network predictions are surveyed with a modified sequence coding method.

2. AN ARTIFICIAL NEURAL NETWORK [5]

The artificial neural network we use is a simple feed-forward network. It consists of layers of computational units as shown in Fig. 1. Connections between units are weighted so that each unit takes as its input the weighted outputs from the preceding layer. The weighted sum of all inputs to a given unit is termed "the activation." The activation a_j of unit j is a sum of the outputs from the previous layer multiplied by the connection weights between layers.

$$a_j = \sum_i w_{jk} x_k \tag{1}$$

Where w_{jk} is the weight of the connection from unit k to unit j, and x_k is the output from unit k.

The output from any unit depends on the activation function. A common activation function is the sigma function:

$$g(a_j) = 1/(1+e^{-aj}) \tag{2}$$

The network uses Eq. (1) and Eq. (2) at every unit to compute output patterns given any input pattern and a set of connection weights.

The weights of such a network can be optimized so that input patterns match to particular output patterns. This is achieved using an algorithm called "back propagation." Starting with randomly assigned weights, the network is trained with a set of corresponding pairs of input and output patterns. For every input pattern in this training set the actual output can be compared to the desired output pattern to produce a cost function:

* Research supported by the young teacher scientific research project of Anhui province (2006jql244)

$$E = \frac{1}{2}\sum_i \left|\xi_i - O_i\right|^2 \quad (3)$$

ξ_i is the target output for unit *i*, and O_i is the actual output for unit *i*.

The most common method for minimizing the cost function *E* is by steepest descent. In this case, for every pattern in the training set, weights are updated proportionally to the negative of the gradient of the cost in weight space. That is, according to the following rule,

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \quad (4)$$

Where η is a parameter defining the learning rate and w_{ij} is the weight being updated? The back propagation-learning algorithm is summarized as follows.

1) Initialize network weights to small random values.
2) Choose an input pattern and apply to input layer.
3) Propagate pattern through the network according to Eqs. (1) and (2) to calculate final outputs.
4) Use Eq. (4) to update every connection.
5) Return to step 2 and cycle.

Once a network has been "trained" by this procedure with a set of input patterns and corresponding output patterns, novel data can be fed through the network to test its ability to generalize.

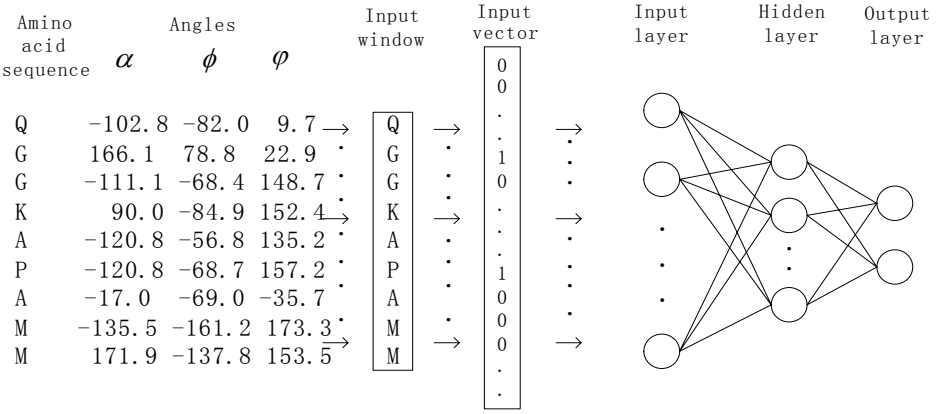


Fig. 1. A figure for a artificial neural network predicting

3. DATABASE

3.1 Selection of entries

The structural data were acquired from the PISCES server [6]. PISCES is a public server for culling sets of protein sequences from the Protein Data Bank (PDB) by sequence identity and structural quality criteria. In the paper, the conditions of culling cases are: the percentage identity cutoff is 30%, the resolution cutoff is 2.5 angstroms, and the R-factor cutoff is 0.2, Skipping non-X-ray and CA-only entries. We attained 1108 macromolecules in the end.

3.2 Definition and location of PPII

Three-dimension of proteins is a suitable format in the DSSP files. Information in DSSP files based on amino acids streochemical properties. There is sequence information, atom coordinates, torsion angles, and several secondary structure information for known protein secondary structure types. We use torsion angles as did literature [4,7] to locate PPII secondary structures from an amino acid sequence. By using these files we set the structural conditions for torsion angles that differentiated PPII and non-PPII cases.

The PPII structures often appear around the point a=-110°, Φ=-75°, ψ=145°(a is a virtual angle, a=180°+ψ_i+Φ_{i+1}+20°(sinΦ_i+sinψ_{i+1})), shows in Fig. 1. The regularity of a structure is an average distance formed with successive angles Φ and ψ. Regularity was computed with Φ and ψ angles using the subsequent equation, where n is the number of amino acids in a sequence:

$$\mathrm{D}=\frac{\sum_{k=1}^{n-1} d_{k,k+1}}{n} \quad (5)$$

Where

$$d_{k-1,k} = \sqrt{(\psi_{i-1} - \psi_i)^2 + (\phi_i - \phi_{i+1})^2}$$

The bound of PPII structures angles:
-145°<a<-70°;
-180°<ψ<-160° or 90°<ψ<180°.

For the non-PPII cases, the angles may be different from the preceding angles or as the preceding angles, but never such three or more successive angles, or as the preceding ones, but Φ and ψ are not sufficiently regular.

Structures were searched for with the following algorithm.

while there exist protein molecules do
while there exist amino acids do
if -145°<a<-70° and (-180°<ψ<-160° or 90°<ψ <180°) then
the next amino acid is taken into consideration
n=1
while -145°< a <-70° and (-180°<ψ<-160° or 90°< ψ<180°) do
the next amino acid is taken into consideration
n=n+1
end while
if n>3 then
regularity computing of the chain(D)
if structure is regular and there are more amino acids in the structure than more n then
the regularity of every part is checked
if regularity D of every part <50 then
the amino acids of the part are marked PPII class
end if
end if
end if
end if

the next amino acid is considered
end while
windowing of protein
the next protein is considered
end while

The algorithm was implemented in C++. A common windowing technique is tested to determine whether a part of a sequence had a PPII structure or not [8]. The purpose of the windowing is to choose sequences for artificial neural networks. The location of a PPII structure is considered in terms of the middle position of the window (Fig. 2.).In the figure, the window length is equal to 13; the grey positions indicate PPII structures.

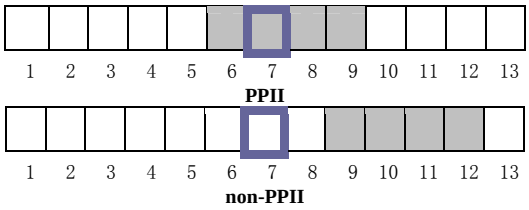


Fig. 2. Windowing technique as exemplified by the window length of 13 amino acids

The windowing technique produced more than 7000 valid sequences for artificial neural networks. The number of PPII structures with different window lengths is showed in table 1.. Naturally, the remaining windowed sequences (more than 300 000) belonged to the non-PPII class, and the distribution was markedly non-uniform. But artificial neural networks need that the distribution between classes in a learning set should be almost uniform. Therefore, we have few choices to equalize the sizes of classes. We chose the most secure means, in the sense of artificial neural network computation, to overcome this difficulty by decreasing the non-PPII class remarkably with random sampling. The distribution inside the non-PPII class is preserved in spite of decreased size. Finally, we attained a uniformly distributed test set.

Table 1. The number of PPII structures with the different window lengths

window length	5	9	13	17
PPII number	7152	7173	7118	7052

4. THE MODIFIED SEQUENCE CODING METHOD

As already demonstrated by many studies, a sequence encoding method that provides maximal feature extraction and information presentation will be the key to successful systems. So sequence-encoding problems are very important in the area of artificial neural networks. The standard data representation for amino acids used for artificial neural networks is to use an orthogonal representation, 20 digit binary number. Thus 10000000000000000000 will represent the amino acid Alanine, 01000000000000000000 will represent Cystine, and so on. In this paper, we used windowing technique to choose sequences for artificial neural networks. The structural type is determined by the amino-acid residue's type in the middle position of the window, that's to say, the structural type of the middle amino-acid residue fully affects the sequence's type. Therefore, we think whether the amino-acid residues beside the middle affect the sequence's structural type. Here, considering the neighboring amino-acid residues' effect, we appended neighboring residues' information in the end of the standard local coding method, and formed the following modified sequence coding method: allowing for neighboring amino-acid residues' effect, we supposed the residue in the middle position of window contribute to form a secondary structure most, and the affecting become less and less when residues are far away the middle position of widow gradually. So the appending unit in the end of standard coding is calculated as follows:

$$\text{Appending unit value}=a_0-a_1*k/K_0 \quad (k=0,1,\ldots, K_0) \qquad (6)$$

where, K_0=integer(l/2), l is current windowing length, k is the distance between residue and the middle of window, a_0 and a_1 are constants.

So the modified sequence code for an amino acid is 20 digits binary plus 1 decimal numbers.

5. MEASURE OF ACCURACY

Cross validation extensively used for artificial neural networks train on one set of data, judge how well it performs on an independent set of data. Here we used 7-fold cross validation. In the cross validation test, dataset was divided into seven parts. Six parts were used to train artificial neural network. Then the trained artificial neural network was used to make prediction for the remaining one part. Because the latter was unknown to the artificial neural network during training, the prediction was realistic.

We used sensitivity, specificity, and total accuracy to measure prediction accuracy. They are defined as follow[9]:

Sensitivity:

$$Se = TP/(TP+FN) \qquad (7)$$

Specificity:

$$Sp = TP/(TP+FP) \qquad (8)$$

Total accuracy:

$$Ta= (TP+TN)/(TP+TN+FP+FN) \qquad (9)$$

Where TP is true positive, FN is false negative, FP is false positive and TN is true negative PPII cases according to decisions made by the artificial neural network.

6. RESULTS

We trained three-layer perceptrons with the back propagation learning algorithm to predict PPII structures. The networks were implemented in the Matlab programming environment. The number of the input nodes of artificial neural network is 21l (l: window length) for modified coding and 20l for standard coding. The number of output nodes was 2. The numbers of hidden nodes N are 4,8,15, 30 or 40 respectively. The transfer function was sigmoid function. The learning rate η is equal to 0.6. The momentum factor is 0.9. The goal error is 0.01. The maximal epochs is 10 000. The a_0 and a_1 in the Eq. 6 are set to 2.0 and 1.9 respectively. Validation sets were applied to prevent over-learning and cross-validation was used. The test result is showed in Table 2.. We also experimented with a few successive artificial neural networks, but their overall results are not improved compared to one network.

The test results for different hidden nodes with the different window lengths are given in Table 2.. As we can see, the Sensitivity, Specificity, and Total accuracy attained with modified coding are better than those with standard coding almost. Especially, when the windowing length is 5 and the hidden node number is 8, the Sensitivity with modified coding can reach 82.5%. That's to say, it can predict 82.5% PPII helix structure from total PPII helix structures. The reason for the results is that the modified

coding allows for the neighboring residues and strengthens local coding information, so the modified coding can represent residues better, and can give more information to artificial neural network. The best total accuracy for modified coding doesn't locate in window length 5. It locates in window length 9 and reaches 73.5%. But the total accuracy little changes in window length 9 and 13, this result is consistent with other secondary structures prediction.

Table 2. The test results for different hidden nodes with the different window lengths

l	*N*	Results for standard coding			Results for modified coding		
		Se(%)	*Sp*(%)	*Ta*(%)	*Se*(%)	*Sp*(%)	*Ta*(%)
	4	72.0	67.0	70.5	82.2	67.1	71.0
	8	72.5	67.3	70.9	82.5	66.9	70.9
5	15	72.3	68.1	71.3	80.7	67.3	70.8
	30	71.7	67.1	70.3	80.3	66.8	70.2
	40	71.5	66.7	70.1	79.2	66.9	70.0
	4	73.8	68.1	71.9	76.6	72.0	73.4
	8	73.9	68.4	72.1	76.3	72.2	73.5
9	15	74.4	70.4	73.1	75.1	72.3	73.2
	30	73.6	67.6	71.7	76.8	71.4	73.0
	40	73.2	67.5	71.4	69.7	73.8	72.5
	4	73.4	69.7	72.2	75.3	72.6	73.4
	8	73.7	69.4	72.3	74.2	71.4	72.3
13	15	74.7	72.1	73.8	74.2	72.2	73.2
	30	72.3	71.7	71.9	72.7	72.7	72.7
	40	72.0	70.7	71.6	72.1	73.8	73.3
	4	72.7	69.1	71.6	73.9	71.1	72.0
	8	72.7	70.1	71.9	74.6	71.1	72.1
17	15	72.6	69.4	71.7	75.1	71.4	72.5
	30	72.2	69.0	71.2	74.6	70.5	71.7
	40	71.4	68.7	70.6	74.1	70.7	71.7

7. CONCLOSION

Artificial neural networks are efficient in solving the complicated prediction problems related to rare protein structures. Especially, neural networks with modified coding can solve the problem of PPII helix structure classification better than artificial neural networks with standard coding. When the windowing length is 5 and the hidden node number is 8, the Sensitivity with modified coding can reach 82.5%, greater than 74.7% attained by standard coding. In the paper, the dataset for ANN is uniform, how to predict non-uniform dataset is the future work.

8. REFERENCES

[1] C.H. Wu, "Artificial neural networks for molecular sequence analysis", Computers Chemistry, Vol.21, No.4, April 1997, 237-256.

[2] P. Vlasov, G. Kilosanidze, et al. "Left- handed Conformation of Polyproline II Type in Globular Proteins", Biophysics, Vol.46, No.3, September 2001,546-550.

[3] B. Stapley, T. Creamer, "A survey of left-handed polyproline II helices", Protein Science, Vol.8, No.3, March 1999,587-595.

[4] M. Siermala, M. Juhola, et al, "On preprocess- ing of protein sequences for neural network prediction of polyproline type II secondary structures", Computers in Biology and Medicine, Vol.31, No.5, September 2001, 385-398.

[5] T.W. Barlow, "Feed-forward neural networks for secondary structure prediction", Journal of Molecular Graphies, Vol.13, No.3, June 1995, 175-183.

[6] G.Wang, R.L.Dunbrack, "PISCES: a protein sequence culling server", Bioinformatics, Vol.19, No.12, August 2003, 1589-1591.

[7] A. Adzhubei, M. Sternberg, "Left handed polyproline II helices commonly occur in globular proteins", Journal of Molecular Biology, Vol.229, No.2, January 1993, 472-493.

[8] N. Qian, T.J. Sejnowski, "Predicting the secondary structure of globular proteins using neural network models", Journal of Molecular Biology, Vol.202, March 1988, 865-884.

[9] B. Pierre, B. Soren, Bioinformatics, Beijing: Zhong-Xing press, Pub., 2003.

A Novel Training Algorithm for BP Neural Network

Fang Bao[1,2]
[1]**School of Information Technology, Southern Yangtze University**
[1]**No. 1800, Lihudadao, Wuxi Jiangsu 214122, China**
[2]**Jiangyin polytechnic college**
Email: baofang@mail.jypc.org
And
Yonghui Pan [1,2]
[1]**School of Information Technology, Southern Yangtze University**
[1]**No. 1800, Lihudadao, Wuxi Jiangsu 214122, China**
[2]**Jiangyin polytechnic college**
And
Wenbo Xu[1]
[1]**School of Information Technology, Southern Yangtze University**
[1]**No. 1800, Lihudadao, Wuxi Jiangsu 214122, China**

ABSTRACT

In this paper, a training algorithm of neural network based on QPSO is proposed. Via training neural network with the quantum particle swarm optimization, the optimized combination of weights and bias of the network is attained. The application of function approximation and data classification is performed using the algorithm in the paper. The experiment result shows that the algorithm has better convergent speed and better global convergent characteristic compared with traditional neural network training algorithm.

Keywords: Neural Network Training, QPSO, Optimized Combination of Weights and Bias, Global Convergent Characteristic, Better Convergent Speed.

1. INTRODUCTION

Along with the trend of information blast in the whole society, the demand for information technology is higher up, so that learning potential pattern or evolving process from abundant original data becomes an absolute necessity. Neural network is widely used in machine learning field. Neural network could approximate such mapping by the muster of specified transfer function and weights and bias, as long as there is any kind of mapping relation between original input data and target output data. Thus neural network can construct correct learning pattern to perform machine learning, which makes neural network be widely used in many fields such as function approximation, pattern classification, image process, etc.

The standard back propagation neural network (BP network) achieves the approximation of the mapping function between input and output through the muster of specified transfer function between layers and the weights and bias of neurons. The training method in BP relies on the modification of weights and bias in the alternant process of forward deviation computation and backward weights computation. Grads descend method is the most frequently used method for executing the backward weights computation. Theoretically, BP neural network will converge ultimately if the descend speed is slow enough, but the convergent process will be so long that the practicability is greatly damaged. If we speed up descend pace, the convergent process is shorter but the local convergence and network shake is easily caused.

Recently, many researches have explored the applicability of evolutionary algorithms to training of the neural network. The superiority of evolutionary algorithm is that it can deal with such instance as the transfer function without differential coefficient or no grads information, when the traditional training method can't solve. Thus, evolutionary computation extends the application domain of neural network.

This paper proposes an algorithm of training neural network using the quantum particle swarm optimization (QPSO)[1]. First, to ascertain the structure of input layer, hidden layer, output layer of the neural network and the fitness function, according to the essence of the specified application. Then to train and modify the weights and bias of all the neurons to ensure the convergence of neural network using the proposed QPSO based algorithm. The experiment result shows that to train the neural network using QPSO has better global convergent character when the original data set is more dispersed and can speed up the convergent pace of the network.

The paper is structured as follows. In next section, we present the QPSO algorithm, and the algorithm of neural network training based on QPSO is proposed. In section 3, we introduce the application of function approximation and data classification using the algorithm of neural network training based on QPSO, and the compared experiment result is shown. The paper is concluded in section 4.

2. THE APPLICATION OF QPSO IN THE TRAINING OF NEURAL NETWORK

Particle swarm optimization (PSO) algorithm is originally introduced by Kennedy and Eberhart[3] in 1995. It is a population-based evolutionary computation technique. The algorithm is motivated by the behavior of animal organization such as migration and prey. Individual make the decision of next step according to the best experience of itself and the best experience of the colony, thus moving to the optimized position step by step. The PSO algorithm initializes a group of random particles (random solution), each particle's state is determined by the position vector X and velocity vector V, and each particle also has a fitness value settled by the fitness function that should be optimized. Then the algorithm search for the optimized solution with

the optimized fitness value through the overlap iterations, in each overlap, each particle updates itself by tracking two extremum, one is the best value found by itself, which named as *pBest*, another is the best value found by the colony, named as *gBest*. After these two values have been found, particle updates its position and velocity according to the following formulas.

$$v = v + c1 * rand(0,1) * (pBest - x) + c2 * rand(0,1) * (gBest - x) \quad \text{Eq. (1)}$$

$$x = x + v \quad \text{Eq. (2)}$$

Where *c1* and *c2* are the learning parameters.

2.1 Summarize of QPSO

QPSO algorithm defines the particle in a quantum space determined by the probability-density function, thus the scope of particle is greatly extended, and the range of problems that can be solved is more generalized.

Particle's position is presented by the wave function ψ (x, t):

$$\int_{-\infty}^{+\infty} |\Psi|^2 dxdydz = \int_{-\infty}^{+\infty} Qdxdydz = 1 \quad \text{Eq. (3)}$$

Q is the probability that the particle appears in position (x,y,z) on the time *T*. Via status transform through monte-carlo method, we can change the quantum status of particles into traditional status, thus the position overlap formula of the particle is proposed as follows:

$$p = (p1 * pBest + p2 * gBest)/(p1 + p2) \quad \text{Eq. (4)}$$

$$L = (1/g)|xi - p| \quad \text{Eq. (5)}$$

$$x = p \pm L/2\ln(1/u) \quad \text{Eq. (6)}$$

Where x is the newly updated position of particle, p1, p2 and u are the learning parameters randomly distributed in the range of 0~1. It has been proved that, only if $g > \ln\sqrt{2}$ is satisfied that the algorithm will converge.

QPSO algorithm initializes a group of random particles (random solution), finds the best value of itself and the best value of the swarm through the fitness function determined by the essence of the application. Then the overlap iterations begin. In each overlap, each particle's position updates according to the formulas (4) and (6), the newest best value of itself and the newest best value of the swarm are computed. Once the maximum overlap iterations or the minimum fitness value is attained, the algorithm is finished and the best value of the colony is the optimized solution of the application. It's very clear that the only parameter needs to be adjusted is g.

The QPSO can be implemented in the following procedure:

Step 1: Randomly initialize each particle's each dimension.

Step 2: Repeat the overlap iteration until terminate criterion is matched. In each iteration, fitness function value *f(xi)* is calculated via the position of each particle, thus finding out the best value of each particle *pBesti* via the fitness value, and finding out the best value of whole colony *gBest* via all the *pBesti*. Each particle's new position is updated through the following formulas:

$$g = (1 - 0.5) * (M - iteration)/M + 0.5 \quad \text{Eq. (7)}$$

M is the total iteration numbers.

$$p1 = rand(0,1), p2 = rand(0,1), u = rand(0,1) \quad \text{Eq. (8)}$$

$$p = (p1 * pBest + p2 * gBest)/(p1 + p2) \quad \text{Eq. (9)}$$

$$L = (1/g)|xi - p| \quad \text{Eq. (10)}$$

$$x = p \pm L/2\ln(1/u) \quad \text{Eq. (11)}$$

In the above procedure, the convergent condition $g > \ln\sqrt{2}$ is always satisfied.

2.2 The Algorithm of Training Neural Network using QPSO

BP neural network achieves the convergence of network depending on the muster of specified transfer function between layers and the weights and bias of all neurons. The training method in BP relies on the modification of weights and bias in the alternant process of forward deviation computation and backward weight computation. So, we focus on how to capture the optimized weights and bias of the neural network effectively.

First, we must determine the structure of the neural network according to the specified problem at hand, and then make the set of all the weights and bias as the dimensions of each particle, hence each particle represents a group of weights and bias in the neural network. The granted value of each dimension (the granted value of each weight or bias) is determined by the practical situation of the application.

Second, randomly initialize all the particles according to the pre-specified dimensions and their specified granted value, and begin the overlap iterations according to the QPSO algorithm. In the process of overlap iteration, fitness function needed for optimization is adopted as the deviation between the actual output and the target output of the neural network. Thus the process of optimization equals with finding out the position that endows the minimum fitness value. We can deduct that the particle which has the minimum fitness value represent a set of weights and bias that can result in the least error between the neural network's actual output and target output when the network converges.

Now we acquire the target of network convergence through training its weights and bias using QPSO. It's obvious that the algorithm of training neural network using QPSO omits the process of the forward deviation computation and backward weight computation, simply focuses on the optimization of weights and bias. Some practical details of the algorithm are listed as follows:

1) the structure of neural networks: A BP neural network that has 3 layers is proposed. In the network, the number of neurons in input and output layers is chosen according to the practical situation of the application. Suppose that the input layer have *N* neurons and the output layer have *M* neurons, based on the known experience, the number of neurons in hidden layer are set as *H*= $\sqrt{N+M}$ +(2~10). We set the number of bias in hidden layer and output layer as 1 respectively, and thus the neural network have 2 biases. The transfer functions between layers are no longer needed here.

2) the dimensionality of particle and the fitness function: The dimension number of each particle is the sum of the number of the weights and bias, i.e. *H***N*+*M***H*+2. So,

the proposed algorithm is optimized by the global output deviation instead of the local output deviation. That is to say, the fitness function of the algorithm is the global output error of the neural network.

3) the training process: With the number of particles often set between 20 and 50, the algorithm randomly initializes the dimensions of each particle according to their pre-specified value range, and finds out the initial best value of itself and the initial best value of the colony via the fitness function. The overlap iterations is often set between 1000 and 5000, then the overlap iteration begins, in each iteration, each particle's position is updated according to the formulas (4) and (6), the newest best value of itself and the newest best value of the swarm are computed. Once the overlap loops end, the best value of the colony is the muster of weights and bias that made the neural network converge.

The compute flows of the algorithm are as follows:

1) Ascertain the structure of the neural network according to the specified application, and the total sum of the number of weights and bias are ascertained, so are the dimensionalities of particle.

2) Ascertain the granted value of each dimensionality according to the specified application, initialize each dimension of each particle in the colony.

3) Begin with the initial colony, using the training data set, train the particles with QPSO algorithm, finally, the particle having best fitness value is found, namely the optimized muster of weights and bias of the neural network be attained.

4) Test the neural network with the optimized weights and bias, using testing data set.

3. EXPERIMENT RESULTS AND DISCUSSION

To test the performance of the algorithm, we apply QPSO-trained BP neural network to function approximation and data classification, and, in turn, we compare the results with those generated from the traditional BP neural network train algorithm-grads descend method.

3.1 The Experiment of Data Classification

According to the order amount, order price, and order forward scheduled time, we can find out the minimum order bankroll engross rate of a specified period of run time, and educe a threshold of the order bankroll engross rate. Thus we can design a neural network, and classify all the order forms into 2 classes based on the threshold.

We use a 3-layer BP neural network for the classification, in which there are 3 neurons designed in input layer, 2 neurons in output layer, 4 neurons in the hidden layer. Thus the dimensionalities of each particle are 4*3+2*4+2=22, the granted value range of each dimensionality is selected according to the practical order amount, price and forward scheduled time of all the order forms, so the value range are very dispersed. The swarm size is 20, and the overlap iterations are 1000. The fitness function is chosen as the amount of the mistake classified individuals. The set size of training data set and testing data set are the same, i.e. 96. The result of data classification using QPSO and grads descend method to train the neural networks with the same overlap iterations 1000 are shown in the following table, when the overlap iterations makes more large, such as 2000 or 3000, if the granted value range of each particle is not changed, namely still very dispersed, the convergent result are almost the same, so we just list out one result set:

Table 1.the compare of convergent result using QPSO and grads descend

Target convergent threshold	Actual convergent value(QPSO)	Actual convergent value (grads descend)
0.235849	0.242321	0.184500
0.000000	0.000000	0.194537
0.700949	0.663986	0.580716
0.762752	0.745427	0.999857

We can see as the order form's data is located in a highly dispersed search space, the traditional neural network training method tends to converge to some local optimized position, while training the neural network with QPSO avoids such trend. It converge to the global optimized position much stable than the traditional method.

3.2 The Experiment of Function Approximation

The function used here is:

$$f(x) = (1 - x + x^2)e^{-x^2} \quad \text{Eq. (12)}$$

According to the nature of the function, we decided to use a 3-layer BP neural network whose input layer have 6 neurons and output layer has 1 neuron and number of neurons in hidden layer being 4. Thus the dimensionalities of particle are 4*6+1*4+2=30, the granted value range of each dimensionality is between 0 and 1. The algorithm employs 20 particles and the overlap iterations are 1000. The fitness function is chosen as the mean square error of the actual output and target output of the network. The set size of training data set and testing data set are the same, i.e. 96. The result of function approximation using QPSO and grads descend method to train the neural networks are shown below. The line with '+' represents the original function, and the smooth red line represents the approximate function:

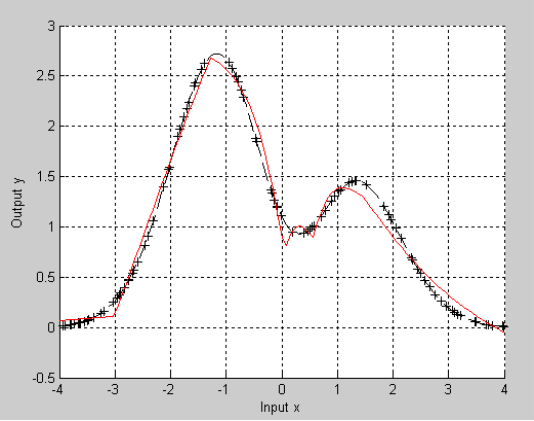


Fig. 1. The approximate result using QPSO

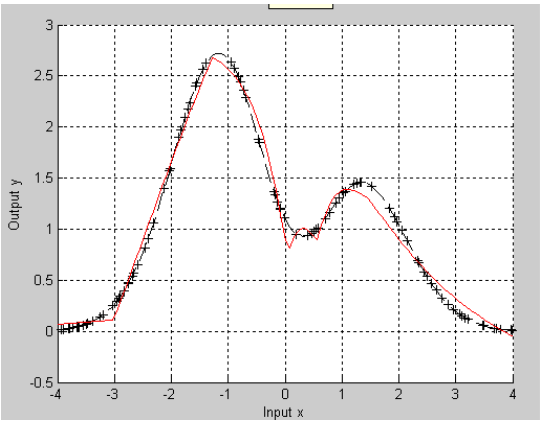


Fig. 2. The approximate result using grads descend

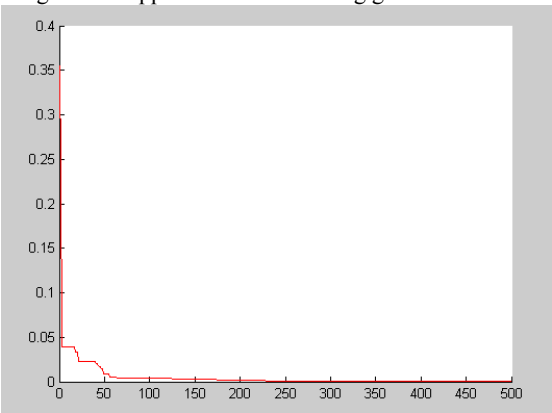


Fig. 3. Down trend of fitness function using QPSO

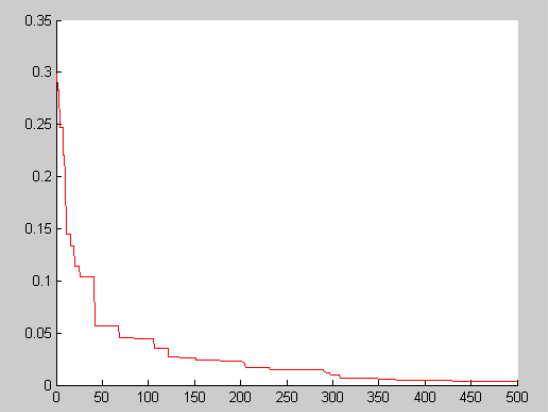


Fig. 4. Down trend of fitness function using grads descend

From the above results, we can conclude that, when applied to the function approximate fields, QPSO-trained neural network almost has the same approximate efficiency as the traditional grads descend method. It has only little superiority in convergent speed.

Thus we conclude that, when applied in the problem where the mapping relation between input and output are more fixed, it's to say the particles are in the traditional space, the algorithm based on QPSO can converge a little quickly, but the convergent efficiency has no superiority compared with the traditional training method. However, when the particles are located in a highly dispersed space because of the diversity of some specified application, it' to say the particles are in the quantum space, the algorithm based on QPSO has obvious superiority in global convergence to the other.

4. CONCLUSIONS

In this paper, we propose the training algorithm of neural network based on QPSO algorithm. QPSO-trained BP neural network can speed up the convergent pace of the neural network to a certain extent because of the process of the forward and backward computation is omitted.

More importantly, when dealing with some various and inconstant specified case, that is to say the data are distributed in the quantum space, the proposed algorithm overcomes the local convergent shortage of traditional neural network training method, it is likely to converge to the global optimized position. The experiment result shows that the algorithm has better global convergent characteristic and better convergent speed.

5. REFERENCES

[1] J.Sun,B.Feng et al, Particle Swarm Optimization with Particles Having Quantum Behavior, IEEE int. Conf. on evolutionary computation,2004, 325-331.

[2] J.Sun, W.B.Xu et al, A Global Search Strategy of Quantum-Behaved Particle Swarm Optimization, IEEE int. Conf. on Cybernetics and Intelligent System, 2004, 111-116.

[3] J.Kennedy, R.Eberhart, Particle Swarm Optimization. IEEE int. Conf. on Neural Network, 1995, 1942-1948.

[4] Y.Shi, R.Eberhart, Parameter Selection in Particle Swarm Optimization. Seventh annual Conf. On Evolutionary Programming, March 1998, 591-601.

[5] R.Eberhart,Y.Shi,ParticleSwarmOptimization:develop ments, applications and resources. IEEE int. Conf. On Evolutionary Computation, 2001, 81-86.

[6] Hecht-Nielsen.R, Kolmogorov's Mapping Neural Network Existence Theorem. First International Conference on Neural Networks, vol. III, 11-14.

[7] S.Becker,G.E.Hinton, A self-organizing Neural Network that Discovers Surfaces in Random-dot Stereograms. Nature, vol.355, 161-163.

[8] J.Kennedy, The Behavior of Particles. The 7th annual Conf. On Evolutionary Computation, 1997, 303-308.

Evolvable Hardware for Fuzzy Logic Controllers Design

Dabin Zhang [1, 2] , Yuanxiang Li [1]
[1] School of Computer, Wuhan University, Wuhan 430072, Hubei, China
[2] Dept of Information Management, HuaZhong Normal University, Wuhan 430079, Hubei, China
Email: zdbff@yahoo.com.cn

ABSTRACT

Fuzzy Logic Controllers (FLCS) are rule-based systems that successfully incorporate the flexibility of human-decision making by means of the use of fuzzy set theory. This paper provides an overview on evolutionary learning methods for the automated design and optimization of fuzzy logic controllers. A three-stage evolution framework that uses Genetic Programming (GP) and Genetic Algorithms (GAS) evolves rule-base and membership function parameters of FLCS. For hardware implement of FLCS, We propose an Evolvable Hardware (EHW) platform for the design of fuzzy logic controllers. This platform, which can be used for the implementation of the designed fuzzy system, is based on a PAMA (Programmable Analog Multiplexer Array). The performance of a fuzzy system in the control of both a linear and a nonlinear function is evaluated. The results obtained with these functions show the applicability of this model in the design of fuzzy control systems.

Keywords: Fuzzy logic controllers, Evolvable Hardware, Evolutionary algorithm.

1. INTRODUCTION

Fuzzy Logic controllers are usually designed by interviewing an expert and formulating her implicit knowledge of the underlying process into a set of linguistic variables and fuzzy rules. In particular for complex control tasks, obtaining the fuzzy knowledge base from an expert is often based on a tedious and unreliable trial and error approach.

Genetic Algorithms (GAS) is heuristic optimization procedures based on the principles of natural evolution. They combine the concepts of adaptation and survival of the fittest to produce suboptimal solutions for an arbitrary problem. Genetic Programming (GP) is a subset of the GA paradigm in which each individual in the population is a computer program. These programs try to solve a certain task, given as a set of examples expressing the desired behavior. Programs are built up by the composition of some functions (internal nodes of the tree) operating on a set of terminal symbols (leaves). This paper provides a general overview on genetic fuzzy systems and describes a framework for the evolutionary tuning and optimization of fuzzy control systems.

Fuzzy Logic Controllers have been used successfully in numerous control systems. The most common implementations of FLCS rely on microprocessors; however, since FLCS admit a high degree of parallelism, an analog solution is also suitable. Unfortunately, analog circuit design is not always an easy task. Recently, Genetic algorithms have been used as agents that program a reconfigurable platform in order to carry out circuit design, promoting a novel area of interest known as Evolvable Hardware (EHW).

This paper is divided into five additional sections. Section 2 introduces a brief review of fuzzy logic controllers. Section 3 describes the main concepts about Evolvable Hardware, FLCS and hardware implementations. Section 4 presents the proposed evolutionary fuzzy system and the platform for future evolution of the fuzzy hardware. Section 5 presents experimental test results and section 6 concludes the work.

2. FUZZY LOGIC CONTROLLERS

Fuzzy Logic Controllers are rule-based systems that successfully incorporate the flexibility of human-decision making by means of the use of fuzzy set theory. Natural-language terms are used in fuzzy rules. Their ambiguity is modeled by membership functions, Intends to represent human expert's conception of the linguistic terms. A membership function gives a measure of the confidence with which a precise numeric value is described by a linguistic label [1].

Rules take the form IF [conditions] THEN [actions], where condition and actions are linguistic terms describing the values of input and output variables (e.g. IF pressure IS low THEN valve-opening IS small). The fuzzy rule-base of the FLC is composed of produce precise output values according to actual input values. This control process is divided into three stages (Fig. 1.) [9, 10]:

(1) fuzzification: calculate fuzzy input, i.e. evaluate input variables with respect to the corresponding linguistic terms in the condition side.

(2) fuzzy inference: calculate fuzzy output, i.e. evaluate activation strength of every rule and combine their action sides.

(3) defuzzification: calculate actual output, i.e. convert fuzzy output into a precise numerical value.

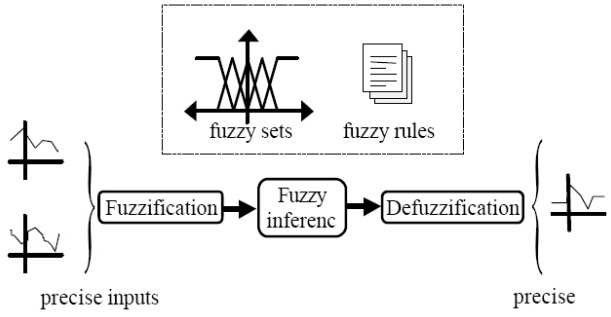


Fig. 1. The fuzzy control process

The fuzzy interpreter used in this work performs fuzzification via triangular membership functions, uses the min intersection operator, the max-min method for fuzzy inference and the centroid procedure for defuzzification.

3. EVOLVABLE HARDWARE

Evolvable hardware (EHW) refers to hardware that can change its architecture and behavior dynamically and autonomously by interacting with its environment. EHW makes synthesis less dependent on human ability and is very useful for finding new and unconventional topologies. Results of various applications show the potential of this technique as compared to the traditional ones [2].

At present, almost all EHW uses an evolutionary algorithm as their adaptive mechanism. Circuits are generally represented by a binary code, the chromosome. An initial population is created and all individuals (circuits) are evaluated. The best individuals have a higher chance of being selected for reproduction in order to form the next population; this is performed until the objective is reached.

Two main methods have been established for applying artificial evolution to the design of hardware systems. These are Extrinsic and Intrinsic. The EHW platform used is named FPGA in digital chips [7, 8], is named PAMA (Programmable Analog Multiplexer Array) in analogue chips. PAMA is classified as Field Programmable Analog Array. FPAAS have just recently appeared, and most projects are being carried out in universities and research centers [3].

4. EVOLUTIONARY FUZZY SYSTEM DESIGN

When we try to design a FLCS, two problems arise: first, how to establish the structure of the controller and, second, how to set numerical values of the controller parameters. The learning algorithm has been successfully applied in these problems, learning controller structure as well as tuning controller parameters.

Import aspects to be considered in development of a fuzzy hybrid system are [4]:

(1) interpretability in terms of linguistic (fuzzy) rules;
(2) capability of adjustment to a given set of data;
(3) rule-base creation and parameters adjustment.

The learning algorithm is illustrated in Fig. 2. Initially the type of membership functions, fitness evaluation and stopping criterion are defined. The policy of alteration in the number of membership functions is also defined.

The learning algorithm is based on the concept of evolutionary cycle; each cycle consists of three stages.
In the first stage, partitioning of input and output spaces takes place. This is in accordance with the number of membership functions assigned to each variable in the cycle. In the second stage a GP searches for the best set of rules among all possible ones. Implication and aggregation operators and defuzzification methods are also taken into account here. Eventually, structure identification is concluded: definition of the rule-base, choice of implication and aggregation operators and defuzzification method.

In the third stage parameter tuning takes place. Membership functions are adjusted so that the most satisfactory fuzzy system for the given numbers of membership functions is obtained. If the requirements are met, the fuzzy system design is completed. If not, the number of membership function may be altered and a new cycle begins.

In the second stage, a single rule can be easily represented as a binary tree: the root being an IF node, and left and right branches representing the conditions and actions can be expressed as trees. On the one hand, a variable paired with a fuzzy set can be represented as a tree with an IS root-mode: the variable name in the left branch and the fuzzy set in the right. On the other hand, a conjunction of such terms can be expressed as a tree with an AND root-node and two branched representing nested conjunctions or pairs (variable, fuzzy set). Fig. 3. shows an example

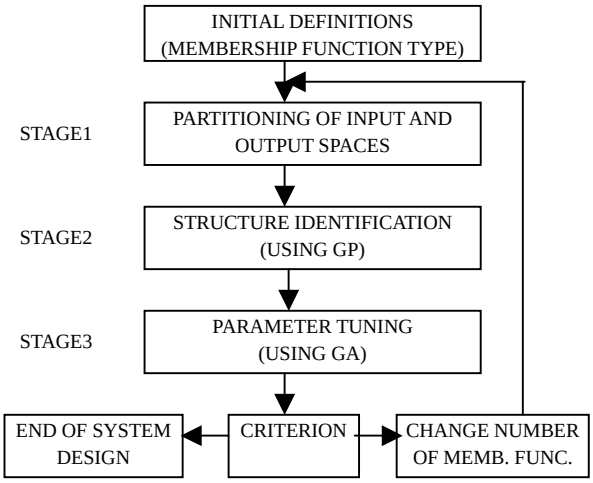
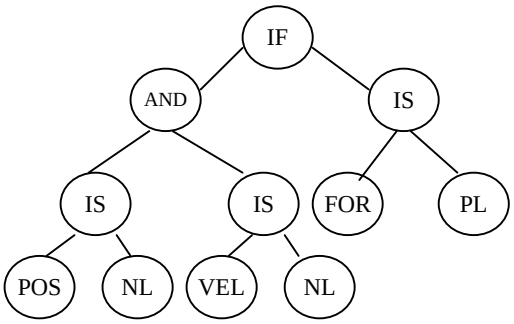


Fig. 2. Learning algorithm



IF *pos* **IS** NL **AND** *vel* **IS** NL **THEN** for **IS** PL

Fig. 3. A syntactic tree and
the rule it represents

A list of rules might be represented as a list of trees. However, it is more appropriate to use a single tree representation because it allows rules to be tightly connected and therefore facilitates the propagation of good functional blocks. Two additional symbols are required. The EL symbol represents an empty list. The RLIST symbol combines rules in the same fashion that AND symbol joins terms, i.e. a list of rules is a tree with an RLIST root-node and two branches representing single rules or analogous lists of rules.

This is a very flexible and powerful representation of FLC rule-based. It deals well with fact that an FLC may have rules with very different structures among them as well as a variable number of rules. This aspect contrasts with other GA-based approaches in which the structure or the number of rules in the FLC is arbitrarily prefixed.

The chromosome used in the third stage, depends on membership functions shape, triangular fuzzy sets ware defined by three characteristic points, trapezoidal fuzzy sets by four points and Gaussian fuzzy sets by center and width. In this example, the parameters identify the vertices of

triangular membership functions; the chromosome is shown in Fig. 4.

The stopping criterion can be a time restriction (fixed number of generations) or a quality restriction (the solution is better than the desired one). Evolution may be halted when there is no decrease in error within a certain number of generations.

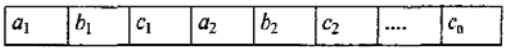


Fig. 4. Chromosome of membership functions

Once the system has been designed and validated, hardware implementation can be started.

Once the system has been designed and validated, hardware implementation can be started.

There are two basic approaches to implement an analog fuzzy system: monolithic and modular.

The first considers a monolithic block implementation to map the inputs to the output. The design via a monolithic block (for implementation of a fuzzy controller, for example) makes use of less hardware than the modular approach does. On the other hand, the design may not be easy due to the complex nonlinear multi input single output mapping.

In the modular approach, the implementation reproduces the traditional configuration of fuzzy logic systems; three main function blocks are considered: fuzzifier, fuzzy inference engine and defuzzifer. These analysis and understanding of the entire system are made easier. On the other hand, many analog circuits are necessary for implementation (each membership function is implemented by a single circuit, for example).

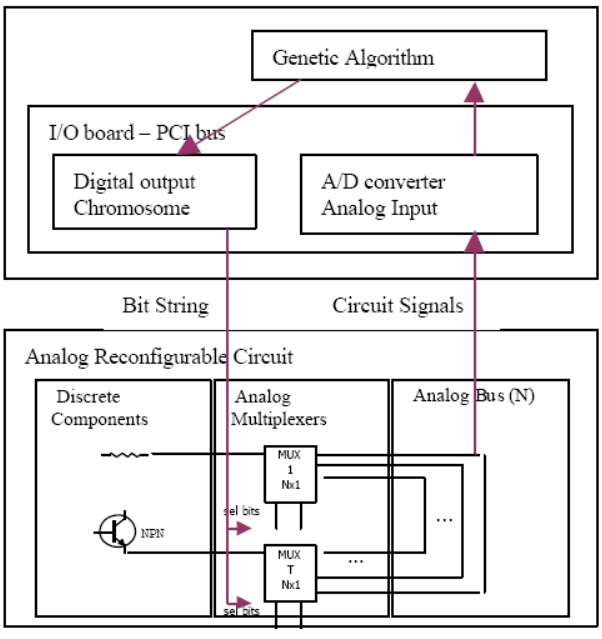


Fig. 5. PAMA block diagram

In a future work a new evolutionary platform shall be used for implementation of the entire system. It is based on a PAMA (Programmable Analog Multiplexer Array), PAMA is classified as Field Programmable Analog Array, is an analog platform based on analog multiplexers/ demultiplexers. These multiplexers are fixed elements and are responsible for the interconnections of the different discrete components that can be plugged into the board. The PAMA supplies an environment to evolve generic analog circuit based on discrete components, without the need of simulators. PAMA is a flexible platform with respect to granularity: its building blocks can be chosen by the designer, from transistors to high level analog components, such as operational amplifiers. Fig. 5 depicts PAMA block diagram [5, 6].

The platform is used to perform intrinsic evolution of analog circuits through a traditional Genetic Algorithm. Each gene configures the select input signals of a particular analog multiplexer. As shown in Fig. 5, a multifunction I/O board connected to the PC bus is responsible for the A/D conversion and for the chromosome download.

5. EXPERIMENTS

Only the fuzzy system design is addressed here; implementation of the electronic circuit is still under development. The membership functions of input and output variables are defined by three parameters, called geometric proportional factors, adding with the scaling factors, which are adjusted to maximize the performance index by using the GA [11].

Key of the Genetic Algorithm is encoding methods and selection of adaptive function. Real-number encoding is used to reduce space dimension. For selection of adaptive function, we use following function as Eq. (1):

$$J(ITAE) = \int_0^\infty t|e(t)|dt = \min \qquad (1)$$

This function can estimate dynamic and static capability of control system as a whole. In order to realize simply by computer, we need discretize data. The type of discretization is as following Eq. (2):

$$J(ITAE) = \sum_{k=1}^{N} t|e(t)|T \qquad (2)$$

where T is time of sampling, N is number of sampling.

Following example is two inputs and one output Fuzzy system, Input of Fuzzy Logic Controllers is e and c, output is u. Their bound are （-E，E），（-Ec，Ec），（-U，U） which are tuned by scaling factors k_e， k_d， k_u. Their values are tuned by geometric proportional factors k_1， k_2， k_3. Fuzzy Logic Controllers is optimized by tuning the six parameters,. In order to demonstrate the effect of scaling factors and geometric proportional factors, a Matlab simulation are made by controlling a transfer function and a nonlinear function. The transfer function is given by Eq. (3). The nonlinear function is given by Eq. (4).

$$H(s) = \frac{2}{s^2 + 4s + 3} \qquad (3)$$

$$y = y(t) + \sqrt{y(t)} + u(t) \qquad (4)$$

The GA parameters of following experiments are:

- Population: 80
- Reproduction: 0.95
- Crossover: 0.75
- Mutation: 0.2

Eq. (3) is a second order linear function, where select T=0.01, input of fuzzy control are excursion e and

excursion rate of change e_c, output of fuzzy control is control variable u, Bound of input and output are [-6，6]. Linguistic variable of input and output has seven levels, membership function is triangular, and control rules refer to references [12]. Scaling factors k_e, k_d, k_u and Geometric proportional factors k_1, k_2, k_3 are adjusted to optimize fuzzy control by using the Genetic Algorithm. Initial values of scaling factors k_e, k_d, k_u and geometric proportional factors k_1, k_2, k_3 are defined 1.5, cycle of sampling is 1000. Simulate result is , $k_e = 59.079$, $k_d = 1.076$, $k_u = 43.466$, $k_1 = 1.0201$, $k_2 = 1.1127$, $k_3 = 1.1054$, $F = 0.1301$. Input/output responses is shown in Fig. 6. .

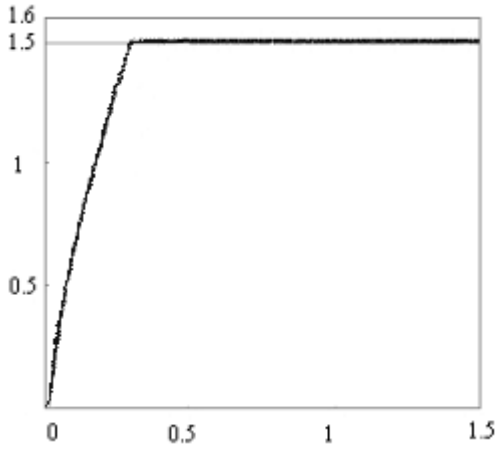


Fig. 6. I/O curve of second order linear function

Eq. (4) is a nonlinear function，Where select T=0.01, k_e, k_d, k_u and k_1, k_2, k_3 are defined 1.5, cycle of sampling is 1000. Simulate result is , $k_e = 1370.7$, $k_d = 1.2$, $k_u = 1$, $k_1 = 1.02$, $k_2 = 0.0244$, $k_3 = 0.3199$, $F = 0.0541$. Input/output responses is shown in Fig. 7 ..

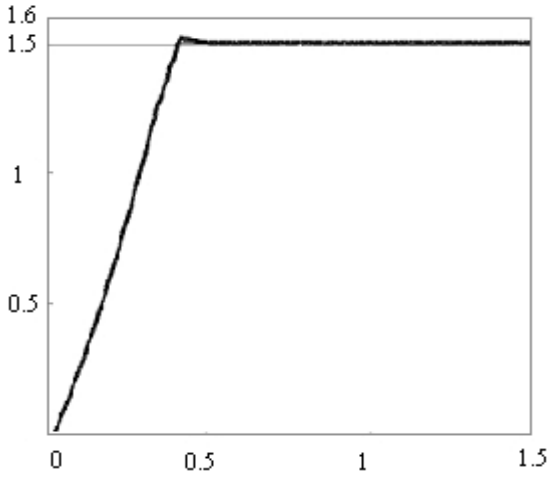


Fig. 7. I/O curve of nonlinear function

6. CONCLUSIONS

From the results of the I/O curve it can be concluded that the new GA method is effective for linear or nonlinear function in fuzzy control optimization.

The results obtained through simulated experiments have been satisfactory for the controllers tested. The quality of control depends upon the algorithm's capability of generating an adequate set of rules and the restrictions imposed during the parameter adjustment stage.

With respect to the new evolutionary model, it should be emphasized that it is not a completely automatic learning fuzzy system; it requires monitoring by the user. How to realize rule-base and membership function parameters to evolve in parallel? This is our important task next stage.

The platform for implementation of Fuzzy Logic Controllers though EHW ideology has been proposed, this platform can use a PAMA and a genetic algorithm to achieve intrinsic evolution. Further work is the implementation of some fuzzy analog circuits by this platform.

7. REFERENCES

[1] C. Carlos, A. Enrique, et al, Evolutionary Design of Fuzzy Logic Controllers, Proceedings of the 1996 IEEE International Symposium on Intelligent Control, Dearborn, MI. September 15-18, 1996，127-132.

[2] X. Yao, T. Higuchi, Promises and challenges of Evovable Hardware [J]. IEEE Trrans. On Systems Man and Cybernetics, Part C: Applications and Review ws, 1999,29(1):87-97.

[3] C.S. Cristin, Z. Ricardo Zebulum, Pama - Programmable Analog Multiplexer Array, The Third NASA/DoD Workshop on Evolvable Hardware, 2001, 0-36.

[4] J.F.M. Amaral, R. Transcheit, et al, M.M.R. Vellasco. Evolutionary Fuzzy System Design and Implementation. Proceedings of the 9th International Conference On Neural Information Processing(ICONIP'02),Vol.4，1872-1876.

[5] J.F.M Amara, Evolvable Building Blocks for Analog Fuzzy Logic Controllers. . 2002 NASA/DOD Conference on Evolvable Hardware(EH'02)，123.

[6] J.F.M. Amara, Towards Evolvable Analog Fuzzy Logic Controllers. 2003 NASA/DOD Conference on Evolvable Hardware(EH'03)，101.

[7] C.H. Pauline, T. Gunnar, an Evolvable Hardware FPGA for Adaptive Hardware (2000). Proceeding of the 2000 Congress On Evolutionary Computation CEC00，553-560.

[8] J.H. Li, H.L. Meng, et al, An Intrinsic Evolvable and Online Adaptive Evolvable Fuzzy Hardware Scheme for Packet Switching Network. 2004 NASA/DOD Conference on Evolvable Hardware(EH'04)，109.

[9] F. Herrera, M. Lozano, et al, A Learning Process for Fuzzy Control Rules using Genetic Algorithms. Fuzzy Sets and Systems, vol. 100, 1998, 143-158.

[10] H. Frank , Evolutionary Algorithms for Fuzzy Control System Design. Proceedings of the IEEE, 2001.

[11] X. Lei，A.B. Zhang, A New GA-Based Optimization Method for Fuzzy Controller Design,

[12] Techniques of Automation &Applications, 2005, 24(10):1-3.

Fuzzy Inference Engine for Capturing Spray Target

Fajun Zhang
College of mechanical & material engineering, Three Gorges University
Yichang, Hubei (Province) 443001, China
Email: doctorzfx@126.com
Cangxiu Chai
Mechanical & electrical college, Wuhan University of technology
Wuhan, Hubei (Province) 430070, China
Email: chaicxiu@mail.whut.edu.cn

ABSTRACT

A basic inference method of precision spraying (PS) in the PCO (Pest Control Operation) market was presented based on fuzzy control system as follows, which has overcame the difficulties found in this profession field of the intelligence spraying fog with truck-mounted machine currently. The fuzzy inference engine (FIE) involved the construction of capturing target and control direction on spraying from fuzzy relations, such as the parameters of move-speed, tree-crown and target-fractal etc, and its logic rule has been introduced according to more import and more output (MIMO) systems emphatically. The results of some related experimental applications about these parameter analysis procedure were given and it was demonstrated that the proposed technique in the fuzzy inference engine for capturing spray target being accurate. It was effectively solved the problem of environmental pollution and the waste of physic drug from blind sprayed and it was concluded that this method would be used in precision spraying for plant protection on the intelligence spraying machinery.

Keywords: Fuzzy Target Fractal Capture Spray.

1. INTRODUCTION

It is severe that the diseases and insect pests damage the constraining forestation and the ecological environmental construction, and yet the technology of drug spraying still laggard presently in our country. The objective reality method of blind spray regardless of the target and non-target objects, it did not take into account the density of plants trees and the area of tree-crown in size, didn't even consider the distribution patterns of individual differences in pesticide demand, the pesticide droplet made the pollutants increased and the ecological environment deterioration. Precision spray (PS) is a response to the situation in new technologies, facing with the target particularity such as types of trees, size of tree-crown, tree-row spacing and distance factors etc, the majority of the information is uncertain or ambiguous, which makes the whole system more complex in intelligent control [1]. In addition, because of the complexity of the object, it is difficult for the export system (precision of the target spray) to accurately describe as input variables (real-time monitoring information) quantitative function, and there are no easy to resolve this problem using conventional means [2]. Taking into account the decision-making spray conditions of uncertainty and ambiguity, this article give a control model of intelligence spray about some fuzzy rules based on the products 6WC-30Y-G truck-mounted machine made by Shenzhen Longrui Technology Co., Ltd (Fig. 1. shown).

2. FUZZY CONTROL ALGORITHMS

On the part of the characteristics of spray machines truck-mounted, taking into account the character of more import and more output (MIMO) of the decision-making system, we Separate it into more import and single output (MISO) system firstly, after have discussed it then made a synthetic algorithm with the iteration theory to the final [3,4].

The variables vector $X = (X_1, X_2 \cdots X_n)^T$ is the importation parameter of the spray system, and the variable of the export system is *Y*, the subscript variable *n* is a dimension of the importation system in which.

Each domain variables constitute a theory in the domain, and each of the number domains can be defined into some subsets, such as the definition a fuzzy subset S_i from the variables vector X_i, and marked $A_{ki}(i = 1,2,\cdots n, k_i = 1,2,\cdots S_i)$, the corresponding rules for example number P paragraph of fuzzy inference may be expressed as:

If $x_1 \in A_{k1}$ ***and*** $x_2 \in A_{k2}$ ***and ..... and*** $x_n \in A_{kn}$

then $y^p = c_0{}^p + c_1^p x_1 + c_2^p x_2, \cdots c_n^p x_n$ (1)

The letter *P* is an order of the rules, x_i is an importation variables of the vector x_i and $x_i \in A_{ki}$, the K_i is fuzzy subsets, y^p is export function expressed by the rules of number paragraph *P*.

When the system requirements to be a group of vector $(x_1, x_2, \cdots x_n)^T$, the total exports of system can be geton from the corresponding export y^p by the weighted average. If there are *N* rules, the system output *y* can be expressed as follows:

$$y = \frac{\sum_{p-1}^{N} \omega^p y^p}{\sum_{p-1}^{N} \omega^p} \quad (2)$$

In which, the ω^p is the real value corresponding number P rules base on the vector $(x_1, x_2, \cdots x_n)^T$ input, the calculation type:

$$\omega^p = \prod_{i=1}^{n} A_{ki}(x_i) \quad (3)$$

In the expression above, the symbol $\prod$ is an operator; the membership function A_{ki} is according to the projects

to choice, it is available to get it from determining the actual situation through data sample group in many cases.

After getting the exported from various single variables, the entire system export response can be constructed by iteration theory from many parameters export in the devices of spray [5].

3. TARGET CAPTURED MODEL

3.1 Target characteristic

Getting the tree-crown images with complex backgrounds by computer visual technology principles, extraction the contours of tree-crown images and studying it with fractal technology to calculate the fractal dimensions of tree-crown, the RGB vector parameters of trees, the distance from the spray machines to the target, the vertical amplitude of tree-crown, and the central coordinates of tree-crown shaped.

Different tree-crown correspond to a different of fractal dimensions in value, this feature can be used for the signals to open electromagnetic valve in flow rate. At the same time the color RGB vector is taken into account for discrimination the interference effects of the advertising billboards and the green belts in urban. In the parameters of feature, the distance parameter from target tree-crown to the vehicle spray is the mainly as signals to control spray pressure, the main parameters of the tree-crown amplitude size and centroid coordinates of tree-crown shaped is the input variables to adjust the tracking of spray target.

3.2 Adjust the target

The nozzle is an outlet for medicine; it is important for the nozzle sprays when and where, so the control signals from the target information is indispensable. Adjusting the target means the nozzle jet is controlled to toward the target all the time while spray. It is particularly important strikes in the process of vehicle dynamics with this point. Accurately we need to analyze how to use the collection of information adjusting nozzle position and operating the valve in flow rate and changing the jet pressure of medicine. The tree-crown vertical amplitude and the centroid coordinates of tree-crown shaped used as the original signal-driven, taking into account the calibration of machine visual, adjust the spray nozzle to aim at the centered position of the tree-crown with the processed electrical signal.

4. FUZZY CONTROL MODEL

There are two aspects of the target recognition in the system, the first is collection the information of target characteristics, and the second is the control method of nozzle towards the target. To simplify the problem, the subject of the collection and identification systems to the target without specialized in study here. The target capture system can abstract for the car speed V, the fractal dimension F, the RGB color parameters C, the distance parameters D, the tree-crown length L, the tree-crown width W and the tree-crown center position P as some import variables, the entire export function is a assembly of the target-control Y_1, the valves-control Y_2 and the pressure-control Y_3, the module of captured fuzzy system as shown in Fig. 2.

Fig. 1. 6WC-30Y-G spray machine

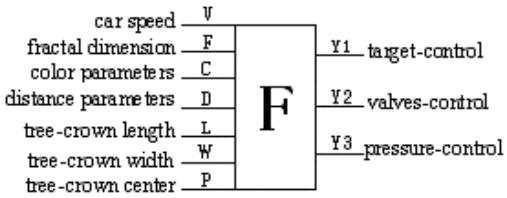


Fig. 2. The fuzzy control system

This fuzzy system contains the following elements taking into account the practical application of spray machine in precision forest industry.

1) Fuzzy the import volume: expressed certain variables value as an appropriate language of the ambiguous situation subsets so as to enable computing;

2) Reasoning machine and the corresponding rules: abstraction fuzzy control rules through a group of fuzzy conditions, and to choose the appropriate function;

3) Calculation the export of fuzzy reasoning machine and getting the accurate control volume of the fuzzy system;

4.1 Import volume precision

Taking into account the features of seven parameters in the system, we divided the system variables into three subsets, in order to improve the dynamic speed of the arithmetic operations. Specifically, the speed parameters is divided into the faster (V_F), the middling (V_M) and the slow (V_S); Using the same method, the distance parameters is divided into the long (D_L), the middling (D_M) and the near (D_N); the parameters of tree-crown amplitude in length is divided into the high (L_H), the middling (L_Z) and the short (L_S); the parameters of tree-crown amplitude in width is divided into the wide (W_W), the middling (W_Z) and the narrow (W_N). The three parameters remaining of the sub-domain F, C and P are not divided into subset, since their influence factor is high to target identification and target adjusting, we make them as a direct importation of the original system and not to deal with them fuzzy.

4.2 Fuzzy model

The role of fuzzy model is to identify state variables as a domain of a fuzzy theory assembly. It can divide in to two ways generally: single value fuzzy models and more value fuzzy models, consider the actual application of the project the more value fuzzy models is given as follows:

$$\mu_A = \begin{cases} 1 \cdots \cdots \cdots x > \overline{x} \\ 0 \cdots \cdots \cdots x = \overline{x} \\ -1 \cdots \cdots \cdots x < \overline{x} \end{cases} \tag{4}$$

The letter $\overline{x}$ is the speed value of the highest frequency in using, in generally the highest frequency speed of the spray machine is $4\,km/h$, we can view $\overline{x} = 4\,km/h$, when $x > 4\,km/h$, the μ_A is equal

to 1, then the move parameter is corresponding to the faster speed (V_F) of the spray machine; When $x < 4\,km\ /\ h$ the μ_A is equal to -1, the move state corresponding to the slow speed (V_S) also; obviously, when $x = 4\,km\ /\ h$, the μ_A is equal to 0, the move parameter is corresponding to the middling speed (V_Z). Identically the other domain parameters *D*, *L* and *W* are analyzed with the same method corresponding.

4.3 Export function

The target function Y_1 for control nozzle of the machine is consist of the variables such as move speed *V*, tree-crown length *L*, tree-crown width *W* and tree-crown-shaped coordinate *P* and it is formed by a certain logic relations, the coordinates demarcation of the computer visual is also considered here [6]. The move speed *V* affects the level angle of the nozzle deflection, the tree-crown length *L* and tree-crown width *W* affect directly the nozzle rotation scope in size too, but the tree-crown-shaped coordinates *P* is the key parameters in tracking target, thus we serve it as the jet center at the actual coordinates of the location, that is to say after the jet premise to aim at the center coordinate of the heart in tree-crown shaped, we also take into account the influence of parameters *V*, *W* and L to the function Y_1, the function expression :

$$Y_1 = f_1(V, L, W)\big|_{P(x,y)} \tag{5}$$

It is not necessarily a spray target when detected the green vector *G* in RGB color parameter, the key factor of judgment is also depending on the dimension value of the target, after the algorithm of logical AND with *G* and F then judged the detection object whether it is real target or not, but one thing for sure, it is not a spray target if testing no green vector *G*. If the variables of *F* and *C* are not empty sets, the valve size of drug liquid is corresponding to the machine speed *V* mainly, because the synthetic speed will affect the condition of opening and the match time in memory system. Otherwise, it is difficult to ensure that the liquid jet to the target from computer visual system, the function expression:

$$Y_2 = f_2(V)\big|_{F<>0,C<>0} \tag{6}$$

To the control function blocks of jet pressure, it is based on the distance from target to nozzle; we can control the rotational speed of engine by the three states of distance parameters *D*, and get the resultant of different jet pressures. It is noteworthy that the adjust changes on engine rotational speed is under the right conditions of the targets correct tracking; otherwise it is senseless for the practical application of engineering. After the target sub-dimensional parameters *F* and the RGB color parameters *C* are not empty sets, the different distances corresponding to the different pressures, so as to reduce the drug floating in the air, the function expression:

$$Y_3 = f_3(D)\big|_{F<>0,C<>0} \tag{7}$$

5. FUZZY CONTROL RULES

According to the definition of the import and export variables, the *y* subset should contain complete sets of rules based on fuzzy inference engines (FIE), the number of the function for control target Y_1 are 27 rules (3*3*3); the number of function for control medicine valves Y_2 are 3 rules; and the number of function for control spray pressure Y_3 are 3 rules too. The corresponding rules expressed as follows:

R^1: ***if*** (V ***is*** V_F ***and*** L ***is*** L_H ***and*** W ***is*** W_W ***and*** F ***is not*** Φ ***and*** C ***is not*** Φ) ***then*** $y_1^1 = g_1^1(v,l,w)\big|_{P(x,y)}$ ***and*** $y_2^1 = g_2^1(v)\big|_{F<>0,C<>0}$

⋮

R^{27}: ***if*** (V **is** V_S ***and*** L ***is*** L_S ***and*** W ***is*** W_N ***and*** F ***is not*** Φ ***and*** C ***is not*** Φ) ***then*** $y_1^{27} = g_1^{27}(v,l,w)\big|_{P(x,y)}$ *and* $y_2^{27} = g_2^1(v)\big|_{F<>0,C<>0}$

R^{28}: ***if*** (V **is** V_F ***and*** F ***is not*** Φ ***and*** C ***is not*** Φ) ***then*** $y_2^{28} = g_2^{28}(v)\big|_{F<>0,C<>0}$

⋮

R^{30}: ***if*** (V ***is*** V_S ***and*** F ***is not*** Φ ***and*** C ***is not*** Φ) ***then*** $y_2^{30} = g_2^{30}(v)\big|_{F<>0,C<>0}$

R^{31}: ***if*** (D ***is*** D_L ***and*** F ***is not*** Φ ***and*** C ***is not*** Φ) ***then*** $y_3^{31} = g_3^{31}(d)\big|_{F<>0,C<>0}$

⋮

R^{33}: ***if*** (D ***is*** D_N ***and*** F ***is not*** Φ ***and*** C ***is not*** Φ) ***then*** $y_3^{33} = g_3^{33}(d)\big|_{F<>0,C<>0}$

Inferred further the rule P:

R^P: ***if*** $v \in V_i^p$ ***and*** $l \in L_j^p$ ***and*** $w \in W_k^p$ ***and*** $f \notin \Phi$ ***and*** $c \notin \Phi$ ***then***

$$\begin{cases} y_1^p = y_{10}^p + y_{11}^p + y_{12}^p + y_{13}^p \\ y_2^p = y_{20}^p + y_{21}^p + y_{22}^p + y_{23}^p \\ y_3^p = y_{30}^p + y_{31}^p + y_{32}^p + y_{33}^p \end{cases} \tag{8}$$

In which the Φ is the empty set, $p = 1 \cdots 33$, $y_{10}^p, y_{20}^p, y_{30}^p$ are expressed the initial value of the corresponding function, which takes into account the factors such as computer calibration and systematic error etc.

Here we should pay attention to this situation that if we change certain parameters of them there are not able to change all of the export function, such as changing the distance parameters *D* from the nozzle to target object, then one of the rules R^{31}, R^{32} and R^{33} searched by the fuzzy inference engines (FIE), and its corresponding function changing only, the other export function in the system maintains the original export volume without the memory replacement.

At the same time，the export parameters in the fuzzy inference engines $y_1^p = y_{10}^p + y_{11}^p + y_{12}^p + y_{13}^p$ 、$y_2^p = y_{20}^p + y_{21}^p + y_{22}^p + y_{23}^p$ and $y_3^p = y_{30}^p + y_{31}^p + y_{32}^p + y_{33}^p$ are collected from the sample $(v_i, l_i, w_i, d_i, y_{1i}, y_{2i})$, in which $i = 1, \cdots, S$, S is the number of the samples. Thinking of the system specificity, it is difficulties for collecting the desirable sample, the method of this paper is to been sought the sample collection through computer simulation.

When an imported variable belong to linguistic meaning, the output function required that it is no longer this types of variables but a specific figures variables [7,8]. The export variables can be gotten according to the series of formula (2), (3), (4) and (8) by rigorous

calculation.

$$y_1 = \frac{\sum_{p-1}^{N} \omega_1^p y_1^p}{\sum_{p-1}^{N} \omega_1^p}$$

$$y_2 = \frac{\sum_{p-1}^{N} \omega_2^p y_2^p}{\sum_{p-1}^{N} \omega_2^p} \quad \text{and}$$

$$y_3 = \frac{\sum_{p-1}^{N} \omega_3^p y_3^p}{\sum_{p-1}^{N} \omega_3^p}$$

In which

$\omega_1^p = \omega_2^p = \omega_3^p = A_{Vi}^p(v) \wedge A_{Li}^p(l) \wedge A_{Wi}^p(w) \wedge A_{Di}^p(d)$, the symbol $\wedge$ is an operator for the min value calculation.

6. RELATED EXPERIMENT

For acquisition the fractal demission from tree-grown is the critical parameters in this fuzzy system, in our work we have use the software MATLAB to get it. The following section presents an experimental result that the fractal dimension had been obtained from an outline of spray target only [9]. The experiments are carried out with PC, PIV 2.4, 256M memories and the CCD device, 10MOONS SDK2000. The size of input image is 240×320 pixels.

Fig. 3. shows the ash degree picture and the results of tree-crown recognition; the experiments have shown a good result visually in recognition about the spraying target.

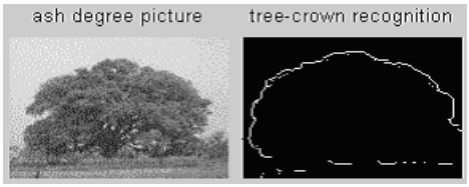


Fig. 3. Original image and partition

We are very convenient to get the fractal dimension of the trees-grown according the box-dimension definition in the Fig. 4. It shows the center-points of the spray target and the course of calculation the fractal dimension from an outline of the tree-crown. There are a wide variety of algorithms for estimating the fractal dimension of a structure outline, such as the fractal of box counting, modified-box-counting, and fractional-Brownian-motion etc. The algorithm for box counting estimates how many boxes are taken up by the fractal structure. An arbitrary grid is placed over the structure to be measured, and the number of boxes in the grid that are filled by the fractal structure is counted.

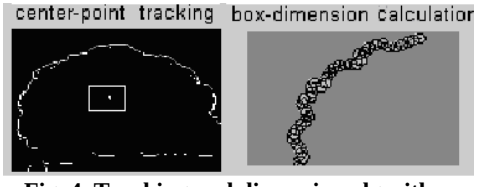


Fig. 4. Tracking and dimension algorithm

7. CONCLUSIONS

Under the precision input of the sub-dimensional, the color vector and the heart-shaped coordinates of the spray target, a decision-making model for the spray machine based on the fuzzy inference system has been expounded. For the first time introduced the fractal dimensions as a signal variable to control the drug spraying, it has given a new method to research the decision-making problem about the market of PCO (Pest Control Operation).

Throughout the system, a large number of specific terms could be gotten on separation line and then embedded the system models later [10]. This system could better meet the vehicle spray requirement in the intellectual spray system the system with a good timeliness and robustness. I am set about my work to use the technology of the fuzzy system for the equipment 6WC-30Y-G truck-mounted machine made in Shenzhen Longrui Technology Co., Ltd. Up to now, the progress in development has shown that it has an important worth for this technology to apply for the field of precision fog in agriculture.

8. REFERENCES

[1] K. Sankar, S. Mitra, et al, Fuzzy Sets and Classification, IEEE Trans. NNS. 1992，3（5）: 683-697.

[2] J. Lawton, R. Wbeard, et al, Approximation of nonlinear Optimal Attitude Control, ACC99-IEEE0220

[3] Kristin L.Makovec, A nonlinear magnetic Controller for Three Axis Stability of Nan satellites, thesis, July 23, 2001.

[4] H. R. Berenji, S. Saraf, et al, Pitch Control of the Space Shuttle Training Aircraft, IEEE Transactions on Control Systems Technology, 9（3）, May 2001.

[5] V. Wu, R. Manmatha, Document Image Binarization and Cleanup, Proc. of SPIE/EI'98, San Jose, CA, Jan., 1998，23-30.

[6] T.Taniguchi, K.Tanaka, et al, Fuzzy descriptor systems and nonlinear model following control, IEEE Transaction on fuzzy systems, 2000, 8(4): 442-432.

[7] DesouzaGN, KakAC, Vision for mobile robot navigation: a survey, IEEE Transactions on Pattern Analysis and MachineIntelligence, 2002, 24(2): 237-267.

[8] R.K. Srihari, Z.F. Zhang, Finding Pictures in Context, Proceedings of the International Workshop on Multimedia Information Analysis and Retrieval, Hong Kong, August 1998.

[9] A. Penn, S. Thompson, et al, Bolinger L (2000) Fractal discrimination of MRI breast masses using multiple segmentations, Hanson KM (ed) Medical imaging 2000: Image processing, SPIE 3979: 959-966.

[10] Y.G. Dong, X.Z. Chen, et al. Simulation of Fuzzy Reliability Indexes, KSME (The Korean Society of Mechanical Engineers) International Journal, 2003, 17(4): 492-500.

Machine learning and Artificial Knowledge Emergence

Fan He, Zhongxiong He
College of Computer Information and Technology, Beijing Jiaotong University,
Beijing, 100044, P.R. China
Email:he_eden@126.com

ABSTRACT

A Knowledge-Activation and self-increasing model is built. First, on the base of the All Set and the Extenics theory, Matter-Element is the foundational description of knowledge, and Knowledge-Iterative-Extension is achieved to improve the convergence of knowledge-space. Then the function of knowledge matching is set up by the Set Pair Analysis, which is excelled at solution matching. The Matter-Element equation is used to get the transitive closure of answers by setting several parameters, the process of Knowledge-Activation is carried out and it makes the knowledge-library increasing spontaneously. Finally, the knowledge emergence is measured by the criterion defined in the end; a better model of Machine Learning is completed.

Keywords: Machine Learning, All Set, Extenics, Set Pair Matching, Knowledge Activation, Artificial Knowledge Emergence, and Autopoiesis Theory.

1. MACHINE LEARNING: FROM REASONING MACHINE TO ARTIFICIAL KNOW LEDGE EMERGENCE

Machine learning is the process that machine simulates human learning methods for improving system performance automatically. Even though many quite mature technologies have appeared, such as inductive learning, discovery learning, genetic algorithm, they are not smart enough. The limitation of knowledge learning is: weak expandability, weak capacity of convergence for knowledge space, weak extraction and analysis for asymmetric information, less ways of learning system and so on. Brand new practical subjects in the field of intelligent control was introduced in the paper on the basis of classical knowledge learning theories: all set, extenics, set pair analysis, problem equation, knowledge activation and so on, aims at integrating multiplicity machine learning methods and setting up new learning system.

2. MACHINE LEARNING AND ALL SET

All set theory

All set is a brand new set theory, which is the united set form of the current set theories including fuzzy set, extension set, vague set, rough set and so on. An all set is comprised of four parts, that is Eq. (1)

$$S = (A, B, F, J) \tag{1}$$

S is all set, *A* is the universe of the problem discussed, one of the elements in *A* can be described by an element of *B*, *F* is the map from *A* to *B*, and *J* restricts *F*. Fuzzy set fixes a membership grade for every object and widely applies to treating under uncertainty. Rough set classifies discuss-field based on equivalence relation and make boundary indistinct classification for subject investigated. Extension set, which is used to deal with contradictory problems, describes the changeability of things; Set pair analysis theory represents the relationship between two sets by identical, discrepancy and contrary. All set is a smart one, not only indicates set in possession, but also generate need set under special circumstances.

2.1 Application of extenics in knowledge acquisition

A perfect learning system needs abundant knowledge as the background for advanced learning. Extenics, which is the interdisciplinary subject of thinking science, system science and mathematics, is introduced here. It permeates artificial intelligence and its theoretical pillars are matter-element theory and extension set theory. Enhancing knowledge acquisition is realized through defining extension set, extension field and regulation transformation of knowledge base. Extension representation of knowledge base was presented in document [3]. Further perfect definition of knowledge matter-element is as follows, *C*, *X* are knowledge matter set, knowledge characteristic set and knowledge value set, respectively. *P* (*M*), *P* (*C*), *P* (*X*) is corresponding power set. Then, knowledge matter -element *R*= (*m, c, x*) samples value at *Z* = *P* (*M*) *P* (*C*) *P* (*X*), *R* is called matter-element variable of *Z*. If the classification is based on potential of *M*, *C*, *X*, when the potential of certain set is 1, it is believed that corresponding variable is constant. We emphasize on the characteristic set *C* and value set *X* of knowledge matter- element variable. When the potential of *C*, *X* set are greater than 1, *R*= (*m, c, x*) is called as value and characteristic multi-transformed matter which is noted as *R*=(*m*0*, c, x*) Likewise, According to different knowledge convergence condition, we can define its matter, value, or matter, characteristic or total transform matter element, which can contribute to matter element extension by various practical matter element function and extension methods. From extension, discovering knowledge will be mentioned in the following chapter.

2.2 Extension set of knowledge

Two different forms of extension set can represent knowledge set.

1) Knowledge matter-element extension set: According to the definition of extension set and relevant function, given that $u \in U$ has an corresponding real number $K(u) \in (-\infty, +\infty)$.

$$\tilde{A} = \{(u, y) \mid u \in U, y = K(u) \in (-\infty, +\infty)\} \tag{2}$$

Eq. (2) is a knowledge extension set on universe of discourse *U* (*U* is the basic element set of all information

and knowledge), $y = K(u)$ is the relevant function [4] of this set. Here, in Eq. (3), A is the positive universe of knowledge extension set $\check{}A$. In Eq. (4), M is the positive extension universe (represent the set that can be transformed into element of knowledge extension set by transformation T) of $\check{}A$ with respect to transformation T Eq. (5).

$$A = \{u \mid u \in U, K(u) \geq 0\} \tag{3}$$

$$M = \{u \mid u \in U, K(u) \leq 0, K(\mathrm{T}u) \geq 0\} \tag{4}$$

$$T = \begin{cases} e & u \in A \\ \text{extension transformation} & u \in \overline{A} \end{cases} \tag{5}$$

$$(e \text{ represents no extensics transformation})$$

2) Extension space with complete information: Given that the definition of basic extension space of knowledge matter-element R [3] is Eq. (6):

$$W_0(R) = R \cup \{R_{\mathrm{M}}\} \cup \{\mathrm{R_C}\} \cup \{\mathrm{R_X}\} \cup \{\mathrm{R}\sim\} \cup \{\mathrm{R}\Rightarrow\} \cup \{\mathrm{R}\leftrightarrow\} \tag{6}$$

Right-hand side of equal sign are: matter-element, matter-element set with same element, matter-element set with same characteristic, matter-element set with same value, correlation matter-element set, implication matter-element set and conjugate matter-element set of R from left to right. Iterating this space for N times to get complete information extension space, Eq. (7) is:

$$\Omega(R) = \lim_{\mathrm{n}\to\infty} \mathrm{W_n} \tag{7}$$

$\mathrm{W_0(R')}$ is the nth-order of extension space.

Knowledge R is the rule that some matter-elements are jointed with coupling word, the above defined basic extension space and complete information extension space can be utilized. The complete information space of knowledge is infinite; certain restriction should be given for this space when expanding. This restricted complete information space is called extension space of knowledge. The definition of knowledge matter-element extension set, similarity and correlation degree between knowledge and so on can be given in space with extension space of knowledge. Traditional extension method can deal with the process of continuation and formation of knowledge; detailed methods are referred to Ref [2].

The emphasis of these two methods is differently; the first model is suitable for normality knowledge, while the second model is suitable for empirical and common sense knowledge. Their extension modes for matter-element are different and may lead to the diversity of their knowledge convergent mode and efficiency.

2.3 Set pair analysis and knowledge matching

We represent guiding and regularity knowledge by matter-element, Eq.(8) is :

$$K = (I_1 \wedge I_2 \wedge ... \wedge I_{\mathrm{n}}) \cup (C_1 \wedge C_2 \wedge ... \wedge C_{\mathrm{n}}) \Rightarrow (R_1 \wedge R_2 \wedge ... \wedge R_{\mathrm{n}}) \tag{8}$$

K-knowledge, C-matter-element describing the precondition, R-matter-element describing conclusion, $\wedge$-coupling word conjunction, $\Rightarrow$-coupling word implication. Whether we obtain exactly *C and I* for solving problem form knowledge library may have direct influence on the executive unit efficiency and result of machine learning system.

We apply set pair matching (SPM) for realization，it is a way of solving fuzzy, randomization, uncertainties consider two sets which have certain relations as a set pair and give admit for objectively uncertainties, meanwhile, see the certainties and uncertainties as a identical、discrepancy、contrary system, that is in Eq.(9):

$$\mu = a + bi + cj, i \in [-1, 1] \tag{9}$$

generally, j=− 1*, a + b + c* = 1*,a, b, c* ∈ [0*,* 1]

Seeking optimal matching scheme [5] based on traditional and fuzzy matching. Then we use WAS [6] of knowledge library to get value of weighted set pair when searching for problem matching. Finally, we make the matching strategy R by comparing values of weighted set pair and referring the magnitude of set pair potential.

3. KNOW LEDGE ACTIVATION AND AUTOPOIESIS THEORY

3.1 Knowledge activation and matter-element equation

Knowledge theory contains two parts: knowledge creation theory and transformation from knowledge to intelligence. The latter can be reckoned as the process of knowledge activation and the former is the precondition- prerequisite of knowledge activation. In the aspect of knowledge creation, we mainly study how to set up knowledge library for solving problems. However, in the aspect of knowledge activation, the key point is to generate solving strategy based on specific solving problems, environment constraint condition and solving goal. According to different levels of knowledge library, the activation paths are as follows:

1) For normality knowledge: First, making the goal, inducing and reasoning correlative knowledge in the library under environmental constraint, generating possible strategy. Finally, through evaluating the effect, confirming whether it needs to be operated.

2) For empirical knowledge: employing empirical structure (such as nervous network), then, establishing its intelligent strategy by the method of direct training of experimental sample.

3) For commonsense knowledge: Here, the intelligent strategy is completely equated to commonsense, as long as apperceiving the input, just react according to the action settled by commonsense.

The three methods above have represented mechanism of symbolic logical school based on functionalism, mechanism of nervous network school based on structuralism and mechanism of perception school based on behaviorism. Setting an example for the process of activation from normality knowledge to intelligence, the expression applying matter-element Eq.(10) is:

$$(Q \wedge K \wedge E) \circ R = P / A \tag{10}$$

Q- problem, A- effect, P - goal, E-environment constraint, R- strategy, K-knowledge, ∘-operator, represents relational composition, /- matching, represents the degree of difference between effect and goal. Better matching effect is based on less difference between effect and goal. To search for strategy R and get the effect under the conditions of problem correlated knowledge and environment constraint. Determining demand set Q, which contains expected effect and goal, and then making information set I, which contains environment constraint. Demand and information are

multi-sided, they can be divided into Q = {Q1, Q_2 ... Q_n}, I={I_1, $I2 \ldots I_n$}, the set (Q_i, I_i) is called the direct-multiplication of Q and I, $R^0=Q\times I$ can be determined by using expert group and calling weight analysis system (WAS). This equation can be solved by successive approximation, transitive closure of R, Eq. (11):

$$R = R^0 \cup R^1 \cup R^2 \cdots = \bigcup_{i=0}^{n} R^i \tag{11}$$

It is solved to get $\left|R - R^n\right| \le \varepsilon$ and make R successively approximate to Rn. Problem state may have changed, then, get effect C according to the degree of difference between new state B and goal P, then, employing set pair matching (SPM) to identify whether the goal P and the effect C match.

3.2 Autopoiesis theory and cognitive process

Two biologists from Chile created Autopoiesis Theory [8] by the research of cell beings. The definition of Autopoiesis Theory is as follows: it's a dynamic system and an entity synthesized by network generating form components; the entity meets the following terms:

To generate network by themselves with reproduction of interacting iteration.

To implement this network by an entity in space. This entity must be able to generate boundary separate with its interacting background.

The elementary model of machine learning can be described as the process of obtaining knowledge, accumulating experience, improving performance, discovering regularities and adapting environment.

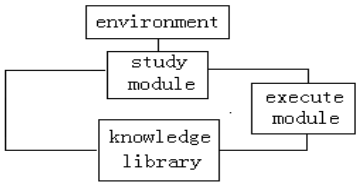


Fig. 1. Simple model of machine learning

From the aspect of model and composition, the model satisfies some key points of Autopoiesis system: separable boundary; components built-up by entities can be described; the entity is a mechanical system; unitized element of boundary is determined by "preferential neighborhood interaction", boundary and entity are generated by the system itself. When we complicate this learning system and provide more intelligent modules or apply Agent theory for simulation, it would be a system with obvious Autopoiesis, which acquire extracts external information continuously, improves knowledge library and develops its system through repeated execution and learning. Then, we can study the coupling between Autopoiesis system and environment.

With the time pass by, the interacting nondestructive disturbance between system and environment may make changes on themselves. Here, the reaction of system to environment is in a "proper way". According to Autopoiesis Theory, "transformation" may generate between system and environment automatically. The "transformation" can be considered as environment, such as that is information and background knowledge, making influence on some components of learning system for internal change and then give feedback. We can call it an implicit "guide" for learning system rather than traditional guide. This "spontaneous" and "continuously" disturbance has the effect that manual intervention cannot reach. What is mattered, this structure coupling also can apply among different Autopoiesis systems. The coupling can be considered as the power of generating new knowledge and intelligence by learning system. The change made by Autopoiesis System would improve the limitation in the field of knowledge innovation; likewise, it makes great sense to the aspect that getting new knowledge relying on suspicion previously and hardness of knowledge analysis. We could build a new learning mode by setting up proper and complete Autopoiesis system, all set, traditional machine learning theories and the machine learning model above which set up by extenics and set pair theory. Now, the model is in the test stage, author will give further discussion in the following parts.

4. MACHINE LEARNING AND ARTIFICIAL KNOW LEDGE EMERGENCE

4.1 Artificial knowledge emergence

Artificial knowledge emergence employs computer to extract its regularities for individual from every modeling and simulation of cognitive object system then interact in computer to make machine emergence possible. Artificial emergence is on the level of strategy, but not a simulation of a large-scale system. First, setting up a model for a guide to explore expected regularities and restraints. Then, choose information, especially for asymmetric information which always contains a large number of practical factor, its fuzziness and extension can be extracted and converted to the symbolic information by using extenics. Finally, determine the weight and combine information. Here, the extensics force [11], proceeding transformation of matter-element and so on are the basic ideas of studying extensics characteristic of strategy, they are turning contradictory problems into compatible and realizing artificial emergence.

Firstly, based on elementary model of machine learning, we take extension method like transformation of matter-element into the process of machine learning. Secondly, apply SPM to deal with matching problem coming from the whole process and generate decision scheme by knowledge activation, meanwhile, introduce dynamic deterministic module to estimate and improve the learning and executive process. Thus, repeat the above process up to knowledge emergence. The standard of knowledge emergence will be discussed in the following parts.

4.2 Measuring for artificial knowledge emergence

To estimate the knowledge artificial emergence, we introduce expert estimate system which contains expert system, WAS and extension method set in learning system. Based on it, the problem equation is defined as Eq. (12):

$$S = f(K, T, W_F) \,\square\, r \tag{12}$$

K - The weight value of knowledge in calling knowledge library

T - The time of machine learning

WF - The weight value of extensics force

r - The goal matter-element

Here, f is a function of K、T、WF, the set of its value is conditional matter-element, r is the goal matter-element, this

problem equation is calculated through employing methods such as the merit rating in extensics theory and set pair matching by expert estimate system, then, estimating the knowledge emergence based on the time of machine learning.

5. FLOW DIAGRAM

Explanation for some steps:
4-calling learning element
6-proceeding extension transformation and building transforming bridge
7-calling execution element
8-proceeding set pair matching
9-building matter-element equation and proceeding knowledge activation
11-calling fuzzy extension force and proceeding conditional improvement
12-calling extension decision 7-4-3 [7] to deal with decision contradiction and conflict
13-employing extension innovation 3-4-4 [13] to set down more proper scheme15-employing expert rating system to calculate the degree of knowledge emergence
17-handling inconsistent phenomenon in scheme by eliminating error15-6-3 [14] method
19-building transforming bridge [2] and analyzing over again
20-recording information during learning process and S value in every step

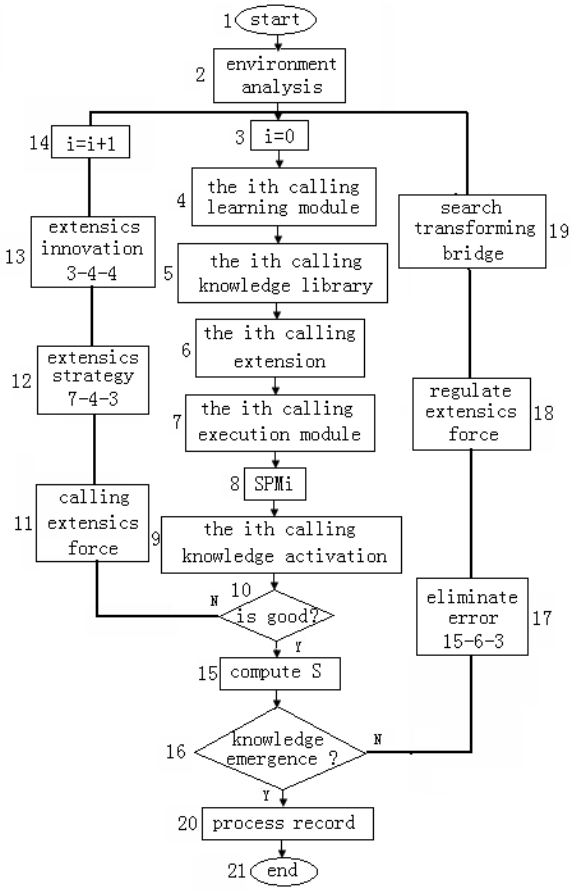


Fig. 2. Knowledge diagram of artificial knowledge emergence

6. EXPECTATION: MACHINE INSIGHT THOUGHT OF ARTIFICIAL KNOW LEDGE EMERGENCE

In this paper, we have employed all set and extenics as the basic methods and have tried to extract the mature thoughts in these theories for every key structure of machine learning system, specific application of extenics, set pair analysis, problem equation and so on for the system have been presented primarily. We consider a learning system whose structure and function is complete as a complex one and combine the theory of traditional machine learning for enhancing the ability of handling background knowledge and asymmetric information. In the future research, we will employ the methods such as autopoiesis theory and traditional nervous network to optimize learning system and hammer at self-continuation and recessive guide of knowledge, build and consummate an enlightenment of thinking machine insight thought model.

7. REFERENCES:

[1] J.Zhang, H.Lin, et al, “All Set and Artificial Intelligence”, Engineering Science, 2002,vol.4 (3): 40-47.
[2] W. Cai, Method of the Extension Engineering. Beijing: science press, 1997.
[3] J.Zhang, Z.X.He, “Introduction to extension machine”, CAAI-9.Guangdong, 1056-1059.
[4] F. He, Z.X. He, “The Extenics In the Artificial Life”, Proceedings of 2005 National Conference on Artificial Intelligence, Beijing: Beijing University of Posts and Telecommunications Press, 2003, 1463-1467.
[5] X.Y. Zhou, Z.X. He, “Set Pair Matching Decision Support System And Application to Matching Talents”. Proceedings of 2005 National Conference on Artificial Intelligence.Beijing, BUPT Publishing House, 2003, 1484-1487.
[6] H. Yan, Z.H. He, “Weight Value Analysis System”, System engineering and electronic technology, vol. 4, 1992,41-46.
[7] Y.Q. Liu, K.Z. Guo, The Collision of Large Scale System and Theory Application of Error. Wu Han: Wu Han SCUT Press, 2000.
[8] J.H. Li, J. Zhang, Digital Creative Century -New Science of Artificial Life, Beijing: Science Press, 2006.1, 308-311.
[9] Y.X. Zhong, “Unified Theory of Information, Knowledge and Intelligence”. Chinese Journal of Electronics, 2003,vol. 6: 56-61.
[10] Y.X. Zhong, “Frame of Knowledge and Theory”, Engineering Science, 2002, vol.9: 50-64.
[11] D.Wang, Z.X.He, “Introduction in Fuzzy Extension Force”. Journal of Liaoning Technical University, 2001,vol. 10:583-585.
[12] L. Dong, Z.X. He, “Control and Artificial Emergence of Complex Large-Scale Systems”, Computing Technology and Automation, 2005, vol. 24:22-25.
[13] W. Cai, C.Y. Yang, Extension Marketing. Beijing: Science and Documents Press, 2000.
[14] K.Z. Guo, S.Q. Zhang, Introdution of Eliminating Error, Guangdong: SCUT Press, 1995.

Simulation of SARS based on Cellular Automata

Lei Yu, Huifeng Xue, Gang Li
Institute of Resources and Environmental Information Engineering,
Northwestern Polytechnical University, Xi'an, Shaanxi, 710072
Email:yul525@163.com

Abstract

This paper proposes on a epidemic model based on Cellular Automata (CA). We developed a two-dimensional CA model. Based on this model, we simulated the spread process of SARS patients in Beijing from April, 21, 2003 to July, 1, 2003. The simulation results are consistent with the actual data, and this shows that the model has practical application value, and can provide support for epidemic control decision-making. The CA model has more advantages than conventional differential equation model in epidemic simulation.

Keywords Cellular Automata, epidemic model, computer simulation, SARS.

1. Introduction

Human history has repeatedly faced extremely serious threat of infectious disease. In fact, the description and prediction of infectious diseases is a long-standing study. Current mathematical models used to help forecast infectious diseases. In mathematical models to study infectious diseases, we can put them into two categories: decisive models and network dynamics models.

Since forties in the 20th century, the decisive model based mainly on differential equation paid attention to gradually and still has a very important academic position till now. The most influential models are SIR and SIS model. The decisive model can often provide the result accorded with the real statistics pretty well, however, this kind of method is determinism, that is if all parameters and all initial value of the variable can be known accurately in the model equation, epidemic disease of moment any evolve situation will certainly at all confirm and may be incomparably prophesied accurately. This is clearly inconsistent with the facts, because spread of infect disease is full of many accidental influence. With the development of artificial intelligence computer technology, the network dynamics model has become the new research focus at present.

Network dynamics is considered to be one of non-linear scientific research. It includes cellular automata, Boolean network, neural network and L-system, etc. The basic structure of dynamics of network can adopt the picture in mathematics to describe. Now cellular automata and artificial neural network model study epidemic mathematical model mainly.

2. Cellular Automata (CA)

According to definition, CA can be regard as by a cellular space and the definition is composed in this spatial transition function. [1]

A CA consists of a regular uniform n-dimensional array, usually of infinite extent. At each site of the cell, a physical quantity takes on values. This physical quantity is the global state of the CA, and the value of this quantity at each site is its local state. Each cell is restricted to local neighborhood interaction only and, as a result, it is incapable of immediate global communication. The neighborhood of a cell is taken to be the cell itself and some (or all) of the immediately adjacent cells. The states at each cell are updated simultaneously at discrete time steps, based on the states in their neighborhood at the preceding time step. The algorithm used to compute the next cell state is referred to as the CA local rule. Usually, the same local rule applies to all cells of the CA.

A CA is characterized by five properties:

(1) The number of spatial dimensions (n);

(2) The width of each side of the array (w). w_i is the width of the j th side of the array, where i =1, 2, 3, … , *n*;

(3) The width of the neighborhood of the cell (d). d_i Is the width of the neighborhood along the i th side of the array;

(4) The states of the CA cells;

(5) The CA rule, which is an arbitrary function F. [2,3,4,5,6]

3. Description of the model

The population over which the epidemic propagation will be modeled is assumed to exist in a two dimensional space. Furthermore, it is assumed that the population distribution is homogenous. The two dimensional space is divided into a matrix of identical square cells, with side length a and it is represented by a CA. Each CA cell includes a number of individuals living there. The number of spatial dimensions of the CA array is n =2. The widths of the two sides of the CA array are taken to be equal. The user of the model defines the size of the array and it is a compromise between accuracy and computation time. A large array size results in a large number of CA cells with small side length a, thus, increasing the model accuracy, but it also results in large computations times. The width of the neighborhood of a CA cell is taken to be =3 in both array sides. The state $C_{i,j}^t$ of the (i, j) CA cell at time t is a flag. The value of this flag indicates the state of the individuals located in the (i, j) cell at time t.

1)If $S_{i,j}^t$ =0, then the population in the (i, j) cell is susceptible to the disease;

2)If $S_{i,j}^t$ =1, then the population in the (i, j) cell is infected;

3)If $S_{i,j}^t$ =2, then the population in the (i, j) cell is

immune to the disease.

The time duration of the disease is user-defined and it can be assumed to be equal to t_{in}. After that time the population has recovered from the disease and has acquired a temporal immunity to this disease. After t_{in} time steps, the flag $S_{i,j}^t$ will change its value from 1 to 2.

Each CA cell starts as susceptible with probability p and becomes infected if some CA cell in its neighborhood is infected, the flag $S_{i,j}^t$ will change its value from 0 to 1. It remains infected for time t_{in} and then it becomes immune, the flag $S_{i,j}^t$ will change its value from 1 to 2. The immune population loses its immunity after time t_{im}, it becomes susceptible to the disease again and the value of the flag $S_{i,j}^t$ becomes equal to 0. The probability p is done according to:

$$p = k*(S_{i-1,j}^t + S_{i,j-1}^t + S_{i,j+1}^t + S_{i+1,j}^t)/4 + l*(S_{i-1,j-1}^t + S_{i-1,j+1}^t + S_{i+1,j-1}^t + S_{i+1,j+1}^t)/4 \quad (1)$$

The probability p is affected by the states of all eight cells in its neighborhood at the present time step t. In $Eq.(1)$, the adjacent nearest neighbors of the (i, j) cell, i.e. the neighbors that have a common side with the (i, j) cell and the diagonal adjacent neighbors are grouped, respectively, together. The effect of the adjacent nearest neighbors is multiplied by k, whereas the effect of the diagonal adjacent neighbors is multiplied by l. It is expected that the (i, j) cell will be infected more quickly; if it has an infected adjacent nearest neighbor than if it has an infected diagonal adjacent neighbor, because of the more extensive contact between populations. Therefore, it is always $k > l$. [7]

4. Process of epidemic propagation

The spreading of the epidemic is homogeneous if the initial properties of the population of CA cells are the same for all cells. Consider that the epidemic process starts at a point in the center of the two dimensional space, where the population exists. The algorithm, in this particular case, was applied to a matrix of $100*100$ cells and it was found that, for the model to produce circular epidemic fronts, the values of the parameters k and l of the local CA rule should be 0.5 and 0.1, respectively. In this case, the central CA cell is assumed to be infected and that it spreads the disease to its neighborhood, the epidemic process finishing after $T = 100$ time steps. The infection and immune times are chosen to be $T_{in} = 5$ and $T_{im} = 20$.

In $Fig.1$, simulated data and actual data are relatively close. Actual data are SARS patients in Beijing from April 21, 2003 to July 1, 2003. This shows that the simulation of the epidemic spread process based on Cellular Automata model is viable.

$Fig.2$ Shows the two-dimensional state maps of the SARS spread in 100 days. In $Fig.2$, SARS patients (grey points) were surrounded by immune people (blank points) and stop infections. Because infected and susceptible persons (blank points) all exist, it has unusual disease spread possibilities. This image is difficult to determine given the theory models, and it cannot be received by decisive model.

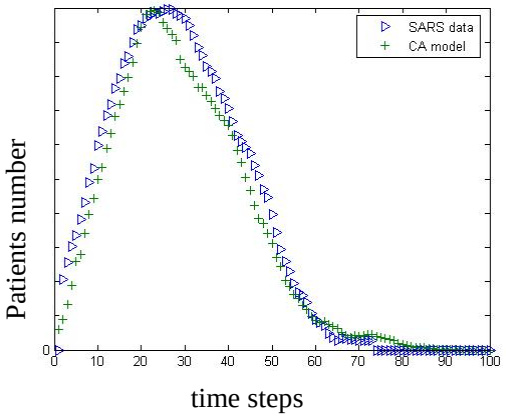


Fig. 1. SARS simulation in Beijing

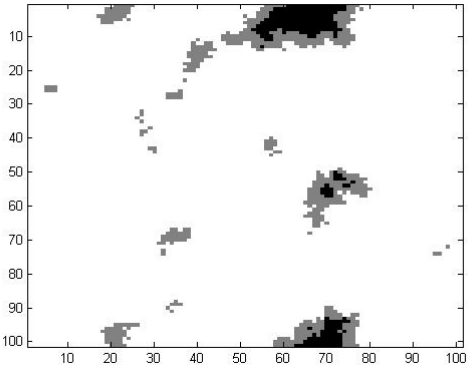


Fig. 2. two-dimensional state maps of the SARS spread in 100 days

5. Conclusions

1) This paper proposes on a epidemic model based on Cellular Automata. Based on this model, we simulate the spread process of SARS. The simulation results are consistent with the actual data, and this shows that the model has practical application value, and can provide support for epidemic control decision-making.

2) Because the simulation of CA model based on microcosmic study, it has more advantages than conventional differential equation model. It has the characteristic of computing mode simple and the analogue result direct viewing, and because of the model structure flexible, it can change the control strategy in the evolutionary process. [8]

3) Because of the CA spatial distribution restriction, we will have further explorer to the study of large-scale epidemic model.

6. References

[1] S. Wolfmanm, “Theory and application of Cellular Automata [M]”, World Scientific, Singaporem 1986.
[2] A. Johansen, “A Simple Model of Recurrent Epidemics [J]”, Journal of Theoretical Biology, 1996,178(1): 45-51
[3] G. Ch. Sirakoulis, I. Karafyllidis, et al “A cellular automaton model for the effects of population movement and vaccination on epidemic propagation [J]”, Ecological Modelling, 2000,133(3); 209-223.
[4] A. Kleczkowski, T. B. Grenfell, “Mean-field-type equations for spread of epidemics”: the “small world” model[J], Physica A, 1999, 274(1), 355-360.
[5] L. de Arcangelis, H. J. Herrmann, “Self-organized criticality on small world networks [J]”, Physica A, 2002, 308(1), 545-549.
[6] M.L. Martinsa, G. Ceottoa, et al, “A cellular automata model for citrus variegated chlorosis”. Physica A 295 2001,42 – 48.
[7] M.A. Fuentes1, M.N. Kuperman, “Cellular automata and epidemiological models with spatial dependence” Physica A 267, 1999, 471.
[8] G.Ch. Sirakoulis, I. Karafyllidis, et al. “A cellular automaton for the propagation of circularfronts andits applications”.

Chaotic Time Series Forecasting with PSO-Trained RBF Neural Network

Bin Feng, Wei Chen, Jun Sun
School of Information Technology, Southern Yangtze University
No. 1800, Lihudadao Road, Wuxi,
214122, Jiangsu, China

Abstract

Radial Basis Function (RBF) networks are widely applied in function approximation, system identification, chaotic time series forecasting, etc. To use a RBF network, a training algorithm is absolutely necessary for determining the network parameters. In this paper, we use Particle Swarm Optimization (PSO), a evolutionary search technique, to train RBF neural network and therefore apply PSO-trained RBF network in chaotic time series forecasting. The proposed method was test on Mackey-Glass model, and the results show that it can predict the time series quickly and precisely.

Keywords: RBF Network, Training algorithm, Particle swarm, Chaotic Time Series

1. Introduction

In order to use a Radial Basis Function network we need to specify the hidden unit activation function, the number of processing units, a criterion for modeling a given task and in turn, a training algorithm for finding the parameters of network. Finding the RBF weight is called network training [1] [2] [4]. If we have at hand a set of input-output pairs, called training set, we optimize the network parameters in order to fit the network outputs to the given inputs. The fit is evaluated by means of a cost function, usually assumed to be the mean square error. After training, the RBF can be used with data whose underlying statistics is similar to that of the training set.

The most widely used training algorithms for RBF network include Orthogonal Least Squares (OLS) algorithm, clustering and gradient-based algorithm, etc. These algorithms, however, possess their shortcomings, which will be mentioned in Section 2. Evolutionary algorithms are a class of population-based search techniques, which have strong global search ability and robustness and could be used to training RBF and other neural networks. They can solve difficult problems with objective functions that do not possess "nice properties" such as continuity, differentiability, satisfaction of the Lipshcitze Condition, etc. Due to these excellences, evolutionary algorithm becomes promising training algorithm for neural networks.

Particle Swarm Optimization is a newly proposed evolutionary approach, which differs from other evolution-motivated evolutionary computation in that it is motivated from the simulation of social behavior. PSO can be easily implemented but is computationally inexpensive. The method requires only the function value, and does not require gradient information of the objective function of the global optimization problem under consideration. On the other hand only primitive mathematical operators are used. Hence the method requires low memory and small computational requirement.

In this paper, we will apply PSO in training RBF neural network and thus use PSO-trained RBF network to predict chaotic time series. The paper is structured as follows. In Section 2 and 3, RBF network model and parameter selection problem are introduced. Section 4 describes PSO algorithm. In Section 5, we propose our PSO-Trained RBF network model. The problem of chaotic time series forecasting is described in Section 6. Section 7 gives the experiments results of the proposed model on a well-known testing problem. Finally, the paper is concluded in Section 8.

2. Structure of RBF Neural Network

RBF Neural Network is structured by embedding radial basis function a two-layer feed-forward neural network. Such a network is characterized by a set of inputs and a set of outputs. In between the inputs and outputs there is a layer of processing units called hidden units. Each of them implements a radial basis function. The architecture of RBF network is shown in Fig.1.

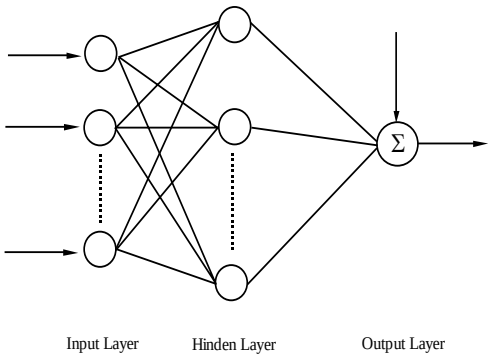


Fig. 1. RBF network in time series modeling

Mathematically the RBF network can be formulated as

$$g(x) = \sum_{k=1}^{m} \lambda_k \varphi_k \left(\left\| x - c_k \right\| \right) \quad (1)$$

Where m is the neuron number of hidden layer, which is equal to cluster number of training set. $\left\| x - c_k \right\|$ Stands for the distance between the data point *x* and the RBF center c_k. λ_k is the weight

related with RBF center c_k. Therefore, the RBF neural networks output is a weighted sum of the hidden layer's activation functions. Various functions have been tested as activation functions for RBF networks. In this paper, we adopt the most commonly used Gaussian RB functions as basis functions, then in the formula (1),

$$\varphi_k(x) = \frac{R_k(x)}{\sum_{i=1}^{m} R_i(x)} \quad (2)$$

$$R_k(x) = \exp\left(-\frac{\|x - c_k\|^2}{2\sigma_k^2}\right) \quad (3)$$

In formula (3), σ_k indicates the width of the *k*th Guassian RB functions. One of the σ_k selection methods is shown as follows.

$$\sigma_k^2 = \frac{1}{M_k} \sum_{x \in \theta_k} \|x - c_k\|^2 \quad (4)$$

Where θ_k is the *k*th cluster of training set and M_k is the number of sample data in the *k*th cluster.

3 Parameter Selection of RBF Neural Network

The neuron number of the hidden layer, i.e., the cluster number of training set, must be determined before the parameter selection of RBF neural network. In this paper, we adopt an efficient method, Rival Penalized Competitive Learning (RPCL) [11], to decide the cluster number.

If the neuron numbers of hidden layer has been decided, the performance of RBF depends on the selection of the network parameters. There are three types of parameters in a RBF neural network model with Gaussian basis functions:

1. RBF centers (hidden layer neurons),
2. Widths of RBFs (standard deviations in the case of a Gaussian RBF)
3. Output layer weights

Different strategies exist for training of RBF neural network models. By means of training, the neural network models the underlying function of a certain mapping. In order to model such a mapping we have to find the network weights and topology. There are two categories of training algorithms: supervised and unsupervised. RBF networks are used mainly in supervised applications. In a supervised application, we are provided with a set of data samples called training set for which the corresponding network outputs are known. In this case the network parameters are found such that they minimize a cost function. In unsupervised training the output assignment is not available for the given set.

A large variety of training algorithms has been tested for training RBF networks. However, the existing training algorithms mentioned have their own shortcomings respectively. In the existing literature, many attempts have been made to employ evolutionary computing approaches, such as Genetic Algorithm (GA) and PSO, to train RBF as well as other neural networks. For instance, in [11], Juang propose a hybrid of GA and PSO for the design of recurrent neural network.

4. Particle Swarm Optimization

Particle Swarm Optimization (PSO) algorithm is a population-based optimization technique originally introduced by Kennedy and Eberhart in 1995 [7]. A PSO system simulates the knowledge evolvement of a social organism, in which each individual is treated as an infinitesimal particle in the *n*-dimensional space, with the position vector and velocity vector of particle *i* being represented as $X_i(t) = (X_{i1}(t), X_{i2}(t), \cdots, X_{in}(t))$ and $V_i(t) = (V_{i1}(t), V_{i2}(t), \cdots, V_{in}(t))$. The particles move according to the following equations:

$$V_{ij}(t+1) = V_{ij}(t) + c_1 \cdot r_1 \cdot (P_{ij}(t) - X_{ij}(t)) + c_2 \cdot r_2 \cdot (P_{gj}(t) X_{ij}(t)) \quad (5)$$

$$X_{ij}(t+1) = X_{ij}(t) + V_{ij}(t+1)$$
$$i = 1,2,\cdots M; j = 1,2\cdots, n \quad (6)$$

Where c_1 and c_2 are called the acceleration coefficients. Vector $P_i = (P_{i1}, P_{i2}, \cdots, P_{in})$ is the best previous position (the position giving the best fitness value) of particle *i* known as the *personal best position* (pbest); vector $P_g = (P_{g1}, P_{g2}, \cdots, P_{gn})$ is the position of the best particle among all the particles in the population and is known as the *global best position* (gbest). The parameters r_1 and r_2 are two random numbers distributed uniformly in (0,1), that is $r_1, r_2 \sim U(0,1)$. Generally, the value of V_{ij} is restricted in the interval $[-V_{\max}, V_{\max}]$.

Many revised versions of PSO algorithm are proposed to improve the performance since its first introduction in 1995. Two most important improvements are the version with an Inertia Weight [10], ***w***, and a Constriction Factor [3], ***K***. In the inertia-weighted PSO the velocity is updated by using

$$V_{ij}(t+1) = w \cdot V_{ij}(t) + c_1 \cdot r_1 (P_{ij}(t) - X_{ij}(t)) + c_2 \cdot r_2 \cdot (P_{gj} - X_{ij}(t)) \quad (7)$$

While in the Constriction Factor model the velocity is obtained by using

$$V_{ij}(t+1) = K \cdot (V_{ij}(t) + c_1 \cdot r_2 \cdot (P_{ij}(t) - X_{ij}(t)) + c_2 \cdot r_2 \cdot (P_{gj} - X_{ij}(t))) \quad (8)$$

Where

$$K = \frac{2}{\left|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}\right|}, \quad \varphi = c_1 + c_2, \quad \varphi > 4 \quad (9)$$

There exists another general form of particle swarm, referred to as the LBEST method in [9]. This approach divides the swarm into multiple "neighbourhoods", where each neighborhood maintains its own local best solution. This approach is less prone to becoming trapped in local minima, but typically has slower convergence. Kennedy has

taken this LBEST version of the particle swarm and applied to it a technique referred to as "social stereotyping" [8].

6 PSO-Trained RBF Neural Network

When training RBF neural network by PSO, a decision vector represents a particular group of network parameters including c_k , λ_k and c_k $(k=1,2,\cdots,m)$. Thus each particle flies in a 3m-dimensional search space with $X_i=(c_1,c_2,\cdots,c_m,\sigma_1,\sigma_2,\cdots,\sigma_m,\lambda_1,\lambda_2,\cdots,\lambda_m)$ denoting its position. Initialization of the population involves generating randomly the position vector X_i $(i=1,2,\cdots,M)$ and setting the personal best position $P_i=X_i$ $(i=1,2,\cdots,M)$.

Since a component of the position corresponds to a network parameter, a RBF network is structured according the particle's position vector. Training the corresponding network by inputting the training samples, we can obtain an error value computed by the following formula.

$$E=\frac{1}{2Q}\sum_{j=1}^{Q}\sum_{s=0}^{c}[y_{s,j}(x_j)-g_{s,j}(x_j)]^2 \qquad (10)$$

where $y_{s,j}(x_j)$ and $g_{s,j}(x_j)$ are the actual response (output) and network's predicted response (output) at output unit *s* on x_j, respectively. Q is the number of the training sample and *c* is the number of output units. The particle is evaluated by the obtained error value (fitness value), by which it can be determined whether P_i and P_g need to be updated. In a word, the error function (17) is adopted as the objective function to be minimized in PSO-based RBF neural network.

There are two alternatives for stop criterion of the algorithm. One method is that the algorithm stops when the increment of objective function value is less than a given threshold ε; the other is that it terminates after executing a pre-specified number of iterations. The following is the description of PSO-Trained RBF neural network algorithm:

(1) Initialize the population by randomly generate the position vector X_i and velocity V_i of each particle and set $P_i=X_i$;

(2) Structure a RBF neural network by treating the position vector of each particle as a group of network parameter;

(3) Training each RBF network on the training set;

(4) Evaluate the fitness value of each particle by formula (10), update the personal best position P_i and obtain the global best position P_g across the population;

(5) If the stop criterion is met, go to step (7); or else go to step (6);

(6) Update the velocity and position vectors of each particle according to equation (5) and equation (6);

(7) Output the P_g as a group of optimized parameters;

6 Chaotic Time Series Forecasting by PSO-Trained RBF

Assume that $x(k),k=1,2,3,\cdots$ is a chaotic time series, the purpose of chaotic time series forecasting is to determine $x(k+p)$ when $x(k-m+1)$, $x(k-m+2),\cdots,x(k)$ are given. In this paper, the problem is reduced to, based on the given sample data $x(1),x(2),...,x(M)$, constructing $M-m-p+1$ pairs of input-output data: $[x(M-m-p+1),...,x(M-p);x(M)]$, $[x(M-m-p),...,x(M-p-1);x(M-1)]$, $[x(1),...,x(m);x(m+p)]$, that is, constructing input-output pairs $[x(n);d(n)],n=1,2,...,N$, where $x(n)=(x_1(n),...,x_m(n))\in R^m$ represent *m* inputs of the system under consideration, $d(n)\in R$ is the expectant response of the system and *N* is the number of sample data (sample size).

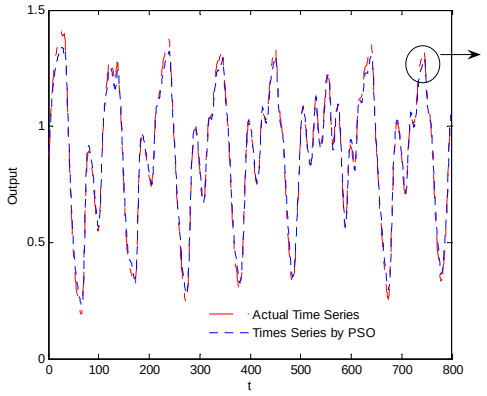
In the process of forecasting chaotic time series by RBF network, the first *l* input-output pairs in the sample data are employed as training set to establish the chaotic time series forecasting model based on RBF network and the last *N-l* pairs are used as testing data to test the times series.

7 Experiments

In this paper, we used Mackey-Glass Model described following as the testing system.

$$\frac{dx(n)}{dn}=\frac{0.2*x(n-\tau)}{1+x^{10}(n-\tau)}-0.1*x(n) \qquad (11)$$

The system comes into chaotic state when $\tau>17$. In our experiment, we set $\tau=30$ and $x(0)=0.6$. Therefore we may work out chaotic time series numerically. In the process of forecasting as described in the above section, we set $m=4$ and $p=1$. When testing Mackey-Glass Model, we set $M=1004$ and thus construct N=1000 sample data pairs, of which the first 800 pairs are used as training set and the last 200 as testing data.



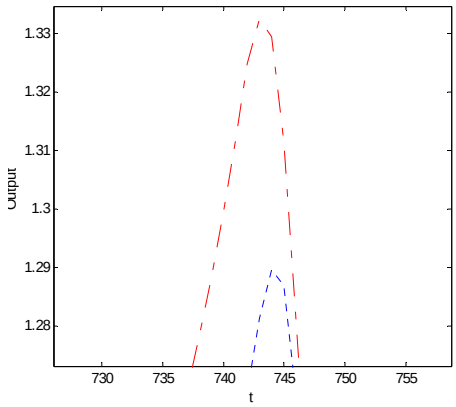


Fig. 2. Experiment results on training set.

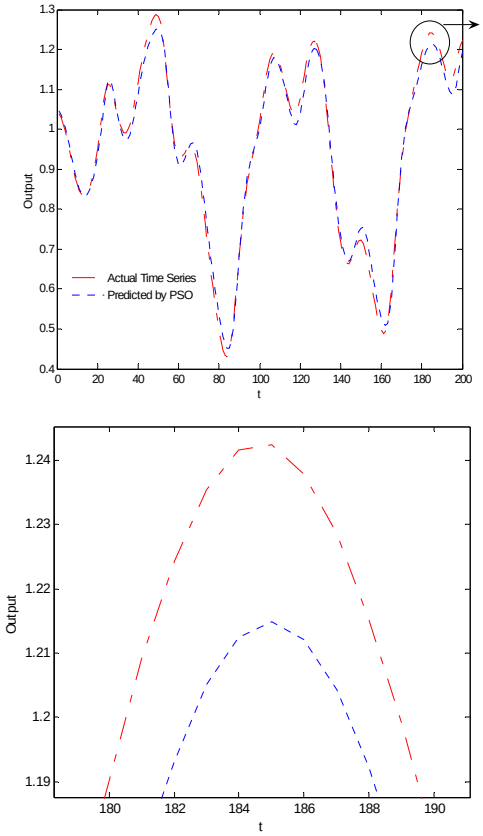


Fig. 3. Experiment results on testing data

We use PSO as training algorithm for RBF network respectively for performance comparison. The experiment configuration is as follows. For PSO, the inertial weight ω varies linearly from 0.9 to 0.4 over the running of the algorithm, the acceleration coefficients c_1 and c_2 are both set to 2 and V_{max} is 15; the training algorithm uses 20 particles and execute for 200 iterations. The RBF network used in our experiment has 4 input neurons, 8 neurons in hidden layer and one output neuron. The experiment results are shown in Figure 2 and Figure 3, where Figure 2 is the visualization of the results on training set and Figure 3 is that of the results on testing data. It can be seen that RBF network trained by PSO can predict the chaotic time series with high precision.

8 Conclusion

In this paper, we employ PSO to train RBF network and use PSO-trained RBF neural network to forecast chaotic time series. The experiment results of PSO-Trained RBF network on the testing problem show that it can predict the chaotic time series more quickly and precisely.

References

[1] M. Casdagli, “Nonlinear Prediction of Chaotic Time Series”, Physica D, vol. 35, pp. 335-356.
[2] S. Chen, C.F.N. Cowan, et al, “Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks”, IEEE Trans. On Neural Networks, 1991, vol.2, no.2, pp.302-309.
[3] M. Clerc, “The Swarm and Queen: Towards a Deterministic and Adaptive Particle Swarm Optimization,” Proc. Congress on Evolutionary Computation 1999, pp. 1951-1957.
[4] T.K. Kohonen, Self-Organization and Associative Memory. 1989, Berlin: Springer-Verlag.
[5] S. Haykin, “Neural Networks: A Comprehensive Foundation”, Upper Saddle River, NJ: Prentice Hall.
[6] C.F. Juang, “A Hybrid of Genetic Algorithm and Particle Swarm Optimization for Recurrent Network Design”, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, VOL. 34, NO. 2, APRIL 2004, pp. 997-1006.
[7] J. Kennedy, R. C. Eberhart, “Particle Swarm Optimization,” Proc. IEEE Int’l Conference on Neural Networks, IV. Piscataway, NJ: IEEE Service Center, 1995, pp. 1942-1948.
[8] J. Kennedy, “Sereotyping: Improving Particle Swarm Performance with cluster analysis,” in Proc. 2000 Congress on Evolutionary Computation, pp. 1507-1512.
[8] P. N. Suganthan, “Particle Swarm Optimizer with Neighborhood Operator,” Proc. 1999 Congress on Evolutionary Computation, pp. 1958-1962.
[9] Y. Shi, R. C. Eberhart, “A Modified Particle Swarm,” Proc. 1998 IEEE International Conference on Evolutionary Computation, pp. 1945-1950.
[10] L. Xu, “Rival Penalized Competitive Learning for Clustering Analysis, RBF Net, and Curve Detection”, IEEE Trans on Neural Networks, 1993, 4(4), pp. 636-649.

Research on the Algorithm of Interval Numbers Reinforcement Learning

Guoqiang Xiong
School of Business Administration, Xi'an University of Technology
Xi'an , P.R. China, 710054
Northwestern Polytechnical University
Xi'an , P.R. China, 710072
Email: xgq168@163.com
Quan Pan, Hongcai Zhang
Northwestern Polytechnical University
Xi'an , P.R. China, 710072

ABSTRACT

Aiming at the problem that in agent reinforcement learning process, it is generally difficult to represent the environmental information and specialist experience with precise value, this paper proposes an interval numbers Q-learning algorithm according to the traditional idea of the Q-learning algorithm which is based on the numerical value information. First of all, the paper explains the reinforcement learning which involves the environmental information of interval numbers, and then offers the steps of the Q-learning algorithm and the two principles to define the optimum strategies, both of which are based on the information of interval numbers. This kind of algorithm is characterized by the fact that in the reinforcement learning the utility function and the reward signal can be expressed as the interval numbers and the two principles offered show that the algorithm proposed has the convergence. In the algorithm, it is suggested that the combination Boltzmann of mechanism with the experiential inference can effectively expand the exploration so as to avoid the partial optimum. This is helpful for the learning machine to accumulate experience and to maintain the stability of fine strategies.

Keywords: Machine Learning, Interval Numbers Q-learning, Agent Intelligent System.

1. INTRODUCTION

Reinforcement learning is an important machine learning method. After Minsky put forward the concept of reinforcement learning in 1954, scholars such as Sutton and Watkins have done researches on the mathematical principle and the algorithm of reinforcement learning. Then, methods such as Monte Carlo method[1], Temporal Differences[2], Q-learning algorithm[3] and Adaptive Heuristic Critic Algorithm[4] are proposed. Reinforcement learning is the learning from environmental situation to action mapping. It is different from surveillance learning. Reinforcement learning must rely on the experience itself, through which intelligent body can obtain knowledge in the action-evaluation environment and improve the action plan to adapt to the environment. Nowadays, reinforcement learning is used widely in the areas of industry control, robots and multi-intelligent body system, etc[5,6]. Q-learning algorithm is a common method of reinforcement learning which is introduced by Watkins[7] and is extended by Rummery and Jing P[8,9]. But, it needs to point out that most of Q-learning algorithm and other methods concerned deal with the precise value on the basis of the concept of real numbers set R (Q value is sum of the future rewards). However, in many practical problems, it is hard to make environmental information accurate because of system state, action, state transition and reward signal collected by Agent, so it is usually in uncertain forms (like fuzzy numbers and interval numbers). At present, researches on this problem have become more and more important, but literatures on Q-learning algorithm which involves interval numbers information are still less. In order to solve the Q-learning problem, the paper introduces a new reinforcement learning method. In contrast with the traditional Q-learning algorithm, utility functions and immediate rewards are in the form of interval numbers, and it is applicable to Non-spot value environmental information. In addition, the two principles offered in the paper show the convergence of the method.

* Corresponding author Phone: 029-6266-0205.

2. BASIC Q-LEARNING ALGORITHM

The living environment of Agent is a limited discrete Markov process, the state event set of Agent perception environment is called state space S. Action set A(s) consists of action $s(t) \in S$ which can operate arbitrarily. For a discrete time sequence t=0, 1, 2, 3, …, Agent and the environment can have mutual effects in each time step-size. Under each sate s_t, action a_t is taken, then Agent will get a real value reward r_t and the next basic state s_{t+1} is observed. Learning assignment of Agent is to learn a controlling strategy π(s, a): $S \times R \to A$, which make the expected value of the reward maximized.

We can treat the controlling strategy as the objective function to learn. Q-learning method denotes the objective function to learn by using utility function $Q(s, a)$. The value of utility function is the immediate rewards when state s_t takes action a_t plus the value that follow the optimum strategy afterward:

$$Q(s_t, a_t) = r_t + \gamma \max\{Q(s_{t+1}, a_t) | a_t \in A\}, \quad (1)$$

where $0 \le \gamma \le 1$ is the discount coefficient and r_t is the immediate rewards.

We can improve learning capability by ceaseless response learning, and utility function Q renews according to the following equation after getting rewards each time when action is taken:

$$Q_t(s_t, a_t) = \begin{cases} (1-a_t)Q_{t-1}(s_t, a_t) + \alpha_t[r_t + \gamma V(s_{t+1})], s = s_t, a = a_t \\ Q_{t-1}(s_t, a_t), \quad \text{otherwise} \end{cases}, \quad (2)$$

where α_t is learning rate which controls learning speed.

$V(s_{t+1})$ is defined as:

$$V(s_{t+1}) = \max_{a \in A} \{Q_{t+1}(s_t, a)\} \quad (3)$$

In the realization of Q-learning method, we usually apply Boltzmann probability distribution method to choose action. Given time step i, the state is s_i, action $a_i \in A$, then the controller choose action a_i with the following probability:

$$\tilde{p}_1(a_t / s) = \frac{e^{Q(s_t, a_t)/T}}{\sum_{a_k \in A} e^{Q(s_t, a_k)/T}} \quad (4)$$

where T is temperature parameter in the process of anneal, which gradually declines along with learning.

According to stochastic process and Markov theory, the literature [3] has proved that $\{Q_t(s_t, a_t)\}$ converge to optimum strategy Q^* with probability 1 when the sequence $\{a_t\}$ satisfies certain condition.

If the environmental information is accurate (numerical value), then it is easy for utility function Q to renew and to choose optimum action. But in many practical problems, the environmental information perceived by Agent usually has uncertainties because of imprecise estimation and measurement errors. Meanwhile, the behavior evaluation in the process of Agent is a multi-attribute and fuzzy judgment. It is hard to give accurate value of reinforcement signal, and adopting a classic real value to reward usually increases the inaccuracy of the decision results. But it is relatively easy to give a possible range. Therefore, to solve this problem, the paper proposes interval numbers learning method to deal with the system state, action and the reward value which are difficult to describe precisely.

3. INTERVAL NUMBERS Q-LEARNING ALGORITHM

3.1 Correlative Conceptions of Interval Numbers Comparisons

$[a^L, a^U]$ is called a interval number, marked as $\tilde{a}$, and a^L is called left interval value, a^U is called right interval value ($a^L \le a^U$). Especially, if $a^L = a^U$, then $\tilde{a}$ becomes a real number spot value.

Definition 1. Let two interval numbers $\tilde{a} = [a^L, a^U]$, $\tilde{b} = [b^L, b^U]$. When $a^L = b^L$ and $a^U = b^U$, $\tilde{a}$ is equal to $\tilde{b}$, and marked as $\tilde{a} = \tilde{b}$. Especially, when $a^L = a^U = b^L = b^U$, $\tilde{a} = \tilde{b}$ is spot value equality.

Definition 2[10]**.** Let two interval numbers $\tilde{a} = [a^L, a^U]$ and $\tilde{b} = [b^L, b^U]$, then the algorithm of interval numbers is:

$$① \tilde{a} + \tilde{b} = [a^L + b^L, a^U + b^U]; \quad (5)$$

$$② k\tilde{a} = [ka^L, ka^U], \ k > 0 \quad (6)$$

Definition 3. Let two interval numbers $\tilde{a} = [a^L, a^U]$ and $\tilde{b} = [b^L, b^U]$, marked as $l_a = a^U - a^L$, $l_b = b^U - b^L$, then

$$p(\tilde{a} \ge \tilde{b}) = \min\{\max[\frac{a^U - b^L}{l_a + l_b}, 0], 1\} \quad (7)$$

is called possibility degree of $\tilde{a} \ge \tilde{b}$.

For a group of given interval numbers $\tilde{a}_i = [a_i^L, a_i^U]$ ($i \in N$), by taking pairwise comparisons, we can get relevant possibility degree $P(\tilde{a}_i \ge \tilde{a}_j)$ according to the equation above, marked simply as $P_{ij}(i, j \in N)$, and establish possibility degree matrix $P = (P_{ij})$. As the matrix includes all possibility degree information about pairwise comparisons of the interval numbers $\tilde{a}_i$., the problem of taking comparison to interval numbers can turns into arrangement problem of $\tilde{a}_i$. according to possibility degree matrix.

Principle 1. If possibility degree matrix about comparisons of a group interval numbers $\tilde{a}_i$ is $P = (P_{ij})_{n \times n}$, then we can deduce arrangement vector $v = (v_1, v_2, \ldots, v_n)$ of interval numbers according to P, and it satisfies[11]:

$$v_i = \frac{1}{n(n-1)}(\sum_{j=1}^{n} p_{ij} + \frac{n}{2} - 1), \ i \in N. \quad (8)$$

3.2 Learning Algorithm

In the consideration of Q-learning problem, the Q-learning which involves interval numbers information consists of the follow parts: discrete state set $s_t \in S$, discrete action set $a_t \in A$ and a reinforcement value set $\tilde{r}_t \in R$. In each time step i, Agent improves learning capability by ceaseless response learning utility function $\tilde{Q}_t(s_t, a_t)$. Given that $\tilde{Q}_t(s_t, a_t) = [Q_t^L, Q_t^U]$ is interval number, the immediate reward signal by taking action each time is also interval number, so the next problem needs to solve is: to interval number information, giving a algorithm by using circle action sequence to learn how to choose the optimum action for an Agent in Markov environment and ensure the convergence of Q-learning.

1) Initialize Q-learning system. Initializing all $\tilde{Q}_t(s_t, a_t)$ value under the state s_t, we can give value arbitrarily in theory. But at the beginning, because the information providing by external environment is less and the state information is difficult to describe accurately, to improve efficiency of reinforcement learning, we can get knowledge by using experts' experiences, then give value to Q-learning initial state.

2) Choose action. In time step i, observing current state s_t, the controller choose action a_t with the following possibility $\tilde{p}(a_t / s)$:

$$\tilde{p}(a_t / s) = \lambda \tilde{p}_1(a_t / s) + (1 - \lambda)\tilde{p}_2(a_t / s), \quad (9)$$

where λ can be chosen as a fixed value or a subsection function; $\tilde{p}_1(a_t / s)$ is Boltzmann probability distribution:

$$\tilde{p}_1(a_t / s) = \frac{e^{\varphi(s_t, a_t)/T}}{\sum_{a_k \in A} e^{\varphi(s_t, a_k)/T}} \quad . \quad (10)$$

In Eq. (10), $\varphi(s_t, a_t) = (Q_t^L + Q_t^U)/2 + \varepsilon(Q_t^U - Q_t^L)$ is mapping function that translates interval number $\tilde{Q}_t(s_t, a_t) = [Q_t^L, Q_t^U]$ into spot value, where ε is a constant and $|\varepsilon| \le 0.5$. T is annealing temperature value and it determines the random degree, the higher T is, the more randomness is.

In the initial stage of learning, T takes a higher value, and the T value reduces gradually in the learning process. In Eq. (9), $\tilde{p}_2(a_t / s)$ is determined by the following equation:

$$\tilde{p}_2(a_t / s) = \frac{m(s_t, a_t)}{M_t}. \quad (11)$$

In Eq. (11), $m(s_t, a_t)$ is the times of state s_t and action a_t, up to time step t; M_t is the numbers of action set $A(s_t)$. $\tilde{p}_2(a_t / s)$ calculates the emergence frequency of action a_t in the learning process. Along with $\tilde{p}_2(a_t / s)$ increasing, it is favorable to find out the fine strategy, and then keep it. In Eq. (9), action a_t with probability $\tilde{p}(a_t / s)$ embodies the following idea: pay more attention to the feedback of environment at the beginning, apply stochastic choosing strategy to expand the exploration so as to avoid partial optimum; along with learning advancement and experiences accumulation of learning machine itself, it gradually translates into the integration of exploration and inference. It is favorable to keep the fine strategy and to solve premature convergence.

3) Learning. Observe next state s_{t+1} and get immediate rewards $\tilde{r}_t = [r_t^L, r_t^U]$, then renew the $\tilde{Q}_t$ value according to the following equation :

$$\tilde{Q}_t(s_t, a_t) = \begin{cases} (1-\alpha_t)\tilde{Q}_{t-1}(s_t, a_t) + \alpha_t[\tilde{r}_t + \gamma \tilde{V}(s_{t+1})], \text{if}(s,a) = (s_t, a_t) \\ \tilde{Q}_{t-1}(s_t, a_t), \text{ otherwise} \end{cases}, \quad (12)$$

where, learning rate is

$$\alpha_t = \frac{1}{1 + m(s_t, a_t)}. \quad (13)$$

In Eq. (10), the calculation of interval numbers is taken according to Eq. (5) and Eq. (6); $\tilde{V}(s_{t+1})$ is defined as:

$$\tilde{V}(s_{t+1}) = \max_{a \in A} \{\tilde{Q}_{t+1}(s_t, a)\}. \quad (14)$$

In order to evaluate the $\tilde{Q}$ value of all state-action combination, according to Eq. (7) we can take comparisons to $\tilde{Q}_{t+1}(s_t, a)$ ($a \in A$) that is got by circle action, then possibility degree matrix $P=(P_{ij})$ is gained. According to comparison information of all $\tilde{Q}$ value, we can deduce arrangement vector $v=(v_1, v_2, \ldots, v_n)$ of interval number $\tilde{Q}_{t+1}(s_t, a)$ by using Eq. (8), and $\max_{a \in A} \{\tilde{Q}_{t+1}(s_t, a)\}$ corresponds to the maximum arrangement value.

4. THEORETICAL JUSTIFICATION

In order to prove the convergence of interval numbers Q-learning algorithm, firstly we introduce Definition 4, Definition 5 and Lemma 1.

Definition 4. Let $\tilde{a} = [a^L, a^U]$ and $\tilde{b} = [b^L, b^U]$ be two interval numbers, let

$$d(\tilde{a}, \tilde{b}) = \left|a^L - b^L\right| \vee \left|a^U - b^U\right|, \quad (15)$$

where, "$\vee$" denotes the maximum value, and $d(\tilde{a}, \tilde{b})$ is called the distance between interval numbers $\tilde{a}$ and $\tilde{b}$.

Definition 5. Let $\{\tilde{a}_k\}$ be a interval numbers sequence, where $\tilde{a}_k = [a_k^L, a_k^U]$ (k=1, 2, …). To interval number $\tilde{a} = [a^L, a^U]$, if any $\varepsilon > 0$, there is $K \in N$, then when $k > K$

$$d(\tilde{a}, \tilde{b}) < \varepsilon. \quad (16)$$

It is said that interval numbers sequence $\{\tilde{a}_k\}$ converge to $\tilde{a}$, marked simply as $\tilde{a}_k \to \tilde{a}$ ($k \to \infty$).

Principle 2. Let $\{\tilde{a}_k\}$ be a interval numbers sequence, where $\tilde{a}_k = [a_k^L, a_k^U]$ (k=1, 2, …), then the sufficient and necessary condition of that $\{\tilde{a}_k\}$ is a convergence sequence $\{a_k^L\}$ and $\{a_k^U\}$ is also convergent.

Proof: Necessity. Let $\{\tilde{a}_k\}$ be a convergence sequence, to $\forall \varepsilon > 0$, there is $k>0$ and interval number $\tilde{a} = [a^L, a^U]$, when $k>K$, then

$$d(\tilde{a}_k, \tilde{a}) < \varepsilon,$$

that is

$$\left|a_k^L - a^L\right| \vee \left|a_k^U - a^U\right| < \varepsilon,$$

it is to say

$$\left|a_k^L - a^L\right| < \varepsilon, \ \left|a_k^U - a^U\right| < \varepsilon,$$

so $\{a_k^L\}$, $\{a_k^U\}$ is convergent.

Sufficiency: If $\{a_k^L\}$, $\{a_k^U\}$ is convergent, then to $\forall \varepsilon > 0$, there is $K_1>0$, when $k > K_1$,

$$\left|a_k^L - a'\right| < \varepsilon,$$

and there is $K_2 > 0$, when $k > K_2$,

$$\left|a_k^U - a''\right| < \varepsilon,$$

so, when $k > K = \max\{K_1, K_2\}$, there is

$$d(\tilde{a}_k, \tilde{a}) = \left|a_k^L - a'\right| \vee \left|a_k^U - a''\right| < \varepsilon,$$

according to Definition 5, $\{\tilde{a}_k\}$ converges to $\tilde{a}$.

□

Lemma l[12]**.** Let F_t be an increasing sequence of σ, $\alpha_t > 0$ and α_t is an observable variable of F_t, in the following supposition:

① Each state and action is involved in algorithm infinite circle;

② $\sum_{t=1}^{\infty} \alpha_t = \infty, \ \sum_{t=1}^{\infty} \alpha_t^2 < C < \infty;$

③ $\exists \gamma, \ 0 \le \gamma \le 1.$

Then, iterative process

$$Q_t(s_t,a_t)=(1-a_t)Q_{t-1}(s_t,a_t)+\alpha_t[r_t+\gamma\max_{a\in A}(Q_{t+1}(s_t,a))]$$

inevitably converges to Q^*.

Principle3. Interval numbers Q-learning algorithm inevitably converge to optimum strategy $\widetilde{Q}^*$.

Proof: According to interval numbers algorithm Eq. (5) and Eq. (6), the iterative process Eq. (12) of interval numbers Q-learning algorithm:

$$\widetilde{Q}_t(s_t,a_t)=(1-a_t)\widetilde{Q}_{t-1}(s_t,a_t)+\alpha_t[\widetilde{r}_t+\gamma\max_{a\in A}(\widetilde{Q}_{t+1}(s_t,a))],$$

is that

$$\widetilde{Q}_t(s_t,a_t)=[Q_t^{\ L}(s_t,a_t),Q_t^{\ U}(s_t,a_t)], \qquad (17)$$

where

$$Q_t^{\ L}(s_t,a_t)=(1-a_t)Q_{t-1}^{\ L}(s_t,a_t)+\alpha_t[r_t^{\ L}+\gamma\max_{a\in A}(Q_{t+1}^{\ L}(s_t,a))]; \quad (18)$$

$$Q_t^{\ U}(s_t,a_t)=(1-a_t)Q_{t-1}^{\ U}(s_t,a_t)+\alpha_t[r_t^{\ U}+\gamma\max_{a\in A}(Q_{t+1}^{\ U}(s_t,a))]. \quad (19)$$

Obviously, sequence $\{\alpha_t\}$ denoted by Eq. (13) is incomplete geometry progression,

$$\sum_{t=1}^{\infty}\alpha_t=\infty, \quad 0<\alpha_t<1,$$

and sequence $\{\alpha_t^2\}$ is convergent, that is

$$\sum_{t=1}^{\infty}\alpha_t^{\ 2}<C<\infty;.$$

Therefore, according to Lemma 1, operator sequence $Q_t^{\ L}$ of iterative process Eq. (18) is convergent, marked as Q^{L*}. Similarly, operator sequence $Q_t^{\ U}$ of iterative process Eq. (19) is convergent, marked as Q^{U*}. So, according to Principle 2, interval numbers sequence Eq. (12) converge to $\widetilde{Q}^*$,

$$\widetilde{Q}^*=[Q^{L*},Q^{U*}].$$

□

5. CONCLUSIONS

Reinforcement learning is important in multi- intelligent body system research, and it is an important machine learning method. Aiming at the problem that environmental uncertainty information of reinforcement learning algorithm is hard to describe accurately, this paper proposes an interval numbers Q-learning algorithm. In each time step, both the utility function and the reward signal are interval numbers, not spot-value data, so the algorithm explores the optimum strategy on the basis of interval numbers information. In contrast with basic Q-learning method, it broadens the environmental application of intelligent body reinforcement learning and it is helpful for dealing with none-spot-value information. This paper proposes an exploration strategy which combines Boltzmann mechanism with experiential inference. This helps to expand effectively and solve the premature convergence. It is favorable to keep the fine strategy, and also to strengthen the reliability.

6. REFERENCES

[1] R.S. Sutton, A.G. Barto, Reinforcement Learning: an Introduction, MA，MIT Press, 1998.

[2] R.S. Sutton, “Learning to Predict by the Methods of Temporal Difference”, Machine Learning, Vol.3, 1988, 9-44.

[3] C.J.C.H. Watkins, P. Dayan, “Q-learning”, Machine Learning, 1992, 8（3）: 279-292.

[4] R.S. Sutton, Temporal Credit Assignment in Reinforcement Learning, PhD Thesis: University of Massachusetts, Amherst, 1984.

[5] L.P. Kaelbling, M.L. Littman et al, “Reinforcement Learning: a Survey”, Journal of Artificial Intelligence, No.4, 1996, 237-285.

[6] Y.Gao, S.F.Chen, et al, “Research on Reinforcement Learning Technology: a Review”, Acta Automatica Sinica, 2004, 30（1）: 86-100.

[7] C.J.C.H. Watkins, Learning from Delayed Rewards, PhD Thesis: Psychology Department, Cambridge University, England, 1989.

[8] G. Rummery, M. Niranjan, “On-line Q-learning Using Connectionist Systems”, Technical Report: CUED/F-INFENG/TR166, Engineering Department, Cambridge University, 1994.

[9] P. Jing, J.W. Ronald, “Increment Multi-step Q-learning”, Machine Learning, Vol.22, 1996, 283-291.

[10] G. Facchinetti, R.G. Ricci, et al, “Note on Ranking Fuzzy Triangular Numbers”, International Journal of Intelligent Systems, Vol.13, 1998, 613-622.

[11] Z.S. Xu, “Algorithm for Priority of Fuzzy Complementary Judgment Matrix”, Journal of Systems Engineering, 2001, 16（4）: 311-314.

[12] M.L. Littman, C. Szepesvari, A Generalized Reinforcement Learning Model: Convergence and Applications, Portland: Proc of the 13th International Conference on Machine Learning, 1996, 310-318.

A Novel Binary-valued Hopfield Network Based Iterative Approach to Stereo Matching

Lili Lin
College of Information and Electronic Engineering, ZheJiang Gongshang University, Hangzhou, Zhejiang, 310035, China
Email: lili_lin@mail.zjgsu.edu.cn
And
Wenhui Zhou
Department of Information Science and Electronic Engineering, ZheJiang University, HangZhou, Zhejiang, 310027, China
Email: sunshine@hzcnc.com

ABSTRACT

Traditional neural network algorithms for stereo matching often suffer the poor performance and slow or premature convergence. In order to avoid these problems, a novel Hopfield network based iterative stereo matching approach is proposed. Firstly, the optimal search problem of disparity map is transformed to an iterative convergence process of binary-valued neural network, whose maximal number of neurons is the size of image. Secondly, the disparity pre-labeling based on local matching method is used to initialize the weights of neural network. Moreover, according to the implicit assumption in local matching algorithms, two evaluation criteria are proposed to determine the disparity range of each pixel, which can reduce the number of active neurons in each iteration. Experiments indicate this approach is much better than traditional algorithms in performance and convergence speed.

Keywords: Stereo Matching, Hopfield Network, Iterative Approach, Energy Minimization, Disparity Pre-labeling.

1. INTRODUCTION

Binocular stereo vision problem is to recover the depth information from two simultaneous images, which are taken from two difference viewpoints. The key stage of this problem is stereo matching, which can be treated as energy minimization problem.

Hopfield neural networks have been successfully applied to solve hard optimization problems over roughly the last two decades. However, these applications suffer the local minimum and convergence problems. In addition, the complexity of Hopfield network based algorithm is related to the size of network. Moreover, its computation time lies on how long it will take to converge that is sensitive to initial conditions.

There are many approaches proposed for stereo matching based on Hopfield type neural networks. Many feature matching models incorporated with artificial neural networks have been proposed [1,2,3,4]. But their main drawback is the produced disparity maps are sparse, because the features are irregularly distributed. If a dense disparity map is required, an extra surface fitting stage is needed. E.Binaghi etc. proposed an adaptive neural model related to zero normalized cross correlation (ZNCC) [5], which fails to get rid of the disadvantages of local area-based matching methods. The approach proposed by Hu [6] is analogous to the method of dynamic programming. Hence, Hu's approach fails to incorporate the smoothness constraints of the horizontal and vertical directions strongly. And the number of neurons in the neural network equals the product of the number of pixels of the left and right epipolars. In reference [7], the global match based neural network method has prodigious structure, where every disparity value of each pixel is modeled as a neuron. So the total number of neurons is the size of whole disparity search space. Obviously, this algorithm has high computation complexity, and it converges slowly.

This paper proposes an iterative stereo match algorithm based on Hopfield network. Firstly, it transforms the optimal search problem of disparity map to a repeat convergence process of binary-valued neural network, whose maximal number of neurons is the size of image. Then, a local matching based disparity pre-labeling strategy method is used to set the initial weights of the neural network. Furthermore, to reduce the number of active neurons in each iteration, two evaluation criteria of local matching method are proposed, which can greatly accelerate the convergence speed. Hence, this algorithm not only has good performance but also converges rapidly.

2. ITERATIVE SOLUTION FOR ENERGY MINIMIZATION

It is well known that stereo matching can be naturally formulated in terms of energy minimization. In this framework, not only the principle of stereo vision can be represented clearly, but also many constraint conditions can be added easily.

A standard form of the energy function is

$$E(d)=\sum_{i\in I} w_i(d_i)+\sum_{\{i,j\}\in N} w_{i,j}(d_i,d_j) \qquad (1)$$

Where $d=\{d_i \mid i\in I\}$ is the disparity of image I, N is the set of all pairs of neighboring pixels, w_i is a data penalty function, and smoothness term $w_{i,j}$ is interaction potential. The energy function in Eq.(1) can be justified on Maximum a Posteriori Probability estimation of Markov Random Fields [8,9].

Therefore, the major difficulty with energy minimization for stereo matching lies on the enormous computation cost. Especially when the energy functions have thousands of

local minima (e.g. in order to get good results at discontinuities, non-convex energy functions with discontinuity-preserving are often used), the global minimization of these energy functions is still a NP-hard problem even using the simplest Potts model [10,11,12]. The energy function with smoothness term described by Potts model is used as followed. In order to minimize the energy function in Eq.(1) by binary Hopfield network, we need to use the method that minimizes an energy function with non-binary variables by repeatedly minimizing an energy function with binary variables. This is analogous to the α expansion move algorithm [10,11,12]. Kolmogorov *etc*. have already derived and demonstrated the convergence and graph-representable of the energy functions involving binary-valued variables [12]. Obviously, this kind of combination optimization problems can be solved by Hopfield type neural network.

Same as the graph construction method of Kolmogorov, a two-dimension neural network with the scale $N_w \times N_h$ is created. Each pixel is modeled as a neuron, which connects only with its neighboring neurons. Let x_i is the state of the ith neuron, x_i =0 represents "inactive" that means disparity value $f(x_i)$ keep unchanged, while x_i =1 represents "active" that means to update the disparity value $f(x_i)$ as new disparity value. Therefore, the optimal disparity search problem can be implemented with the iterative convergence process of a binary-valued neural network.

3. COST FUNCTION FOR HOPFIELD NEURAL NETWORK

3.1 Energy function construction for Hopfield network

The neural network based approach need to transform the energy function and constrain conditions of stereo matching into the cost function of Hopfield network. The energy function of neural network can be formulated as followed.

$$E = -\sum_i \sum_{j,\{i,j\}\in N} c_{i,j} x_j x_i - \sum_i c_i x_i \tag{2}$$

$$c_{i,j} = \begin{cases} \lambda \cdot \left[w_{i,j}(k, d_j) - w_{i,j}(d_i, d_j) \right], & x_j = -1 \\ \lambda \cdot w_{i,j}(d_i, d_j), & x_j = 1 \end{cases} \tag{3}$$

$$c_i = w_i(d_i) - w_i(k) \tag{4}$$

Where N is a set of the nearest four-neighbor in the image, x_i is the state of the neuron i, and k is the candidate disparity.

Let $I(i)$ is the intensity of the pixel i, and the disparity of the pixel i is denoted as $f(x_i)$. Hence, a typical choice for w_i is the $L2$ distance.

$$w_i = [I(i) - I(i + f(x_i))]^2 \tag{5}$$

To achieve good results at discontinuities, it is important that the energy function must be discontinuity preserving. One of the simplest examples is the following Potts model,

$$w_{i,j} = \lambda \cdot T(f(x_i) \neq f(x_j)) \tag{6}$$

Where $T(\cdot)$ is 1 when its argument is true and 0 otherwise, λ is regularization parameter.

3.2 Convergence proof of the Hopfield network

Since Eq.(2) can be reformulated as followed,

$$E = -\sum_i \left\{ \left[\sum_{j,\{i,j\}\in N} c_{i,j} x_j + c_i \right] \cdot x_i \right\} \tag{7}$$

And the state of the ith neuron is

$$x_i = \begin{cases} 1, & \left[\sum_{j,\{i,j\}\in N} c_{i,j} x_j + c_i \right] \geq 0 \\ -1, & \left[\sum_{j,\{i,j\}\in N} c_{i,j} x_j + c_i \right] < 0 \end{cases} \tag{8}$$

The energy function E always descends or keeps unchanged when the state of the ith neuron x_i changes.

4. LOCAL METHOD BASED PRE-LABELING TECHNIQUE

It is well known that local correspondence methods are very efficient, but they are sensitive to locally ambiguous regions in images. Moreover, they perform poorly near object boundaries (i.e., depth discontinuities). The reason is that such methods implicitly assume that all the pixels within a window have similar disparities. However，local methods can achieve good results in the rest regions. And the ZNCC method performs even better, because it is relatively insensitive to radiometric gain and bias.

Since the selection of initial weights directly affects the convergence of neural network, we use the disparity results of ZNCC to implement the pre-labeling, and select the initial weights. According to the evaluation of disparity results of ZNCC, the pixels with reliable disparity are partitioned off, and they do not need to update their disparity values by neural network. So the size of neural network is reduced effectively.

According to the implicit assumption in the local match algorithm, we create a sliding evaluation window, which is of the same size with match window, on the disparity map obtained by ZNCC algorithm. In practice, sliding evaluation window is often a little larger than match window to achieve better robustness. The following two criteria are used for evaluation.

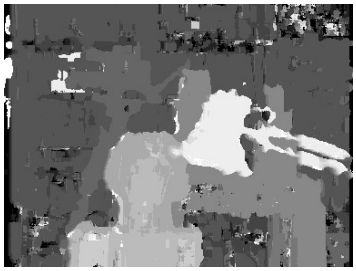
Criterion 1: If all pixels in the evaluation window have the same disparity value, then the disparity value of the center pixel is assumed correct. For example, Fig. 1(a) is the result of ZNCC, and Fig. 1(b) is the evaluation of Fig 1(a). In Fig. 1(b), the pixels in white regions can be viewed as having reliable disparities.

Criterion 2: If all pixels in the evaluation window do not have identical disparity value, then the average bias between the center pixel and the other pixels are calculated. If the average bias is smaller than the threshold, then according to the smoothness constraint, it can be assumed the disparity value of center pixel is in the disparity range of evaluation window.

So the maximal and the minimal disparities in the evaluation window are instructive for the range of center pixel. For robustness purpose, this disparity range should be enlarged appropriately. Apparently, criterion 1 is a special case of criterion 2.

For pixels that do not satisfy either criterion 1 or criterion 2, it can be supposed there are dramatic disparity changes between these pixels and their neighboring pixels. These pixels usually locate at ambiguous regions or near object

boundaries and their disparity values should be searched in the whole disparity range. Fortunately, the number of these kind pixels is only a small percent of the whole image.



(a) Result of ZNCC

(b) Evaluation of the disparity map

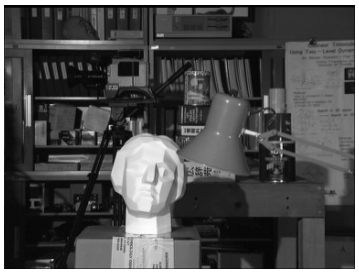
Fig. 1. Result and evaluation of ZNCC algorithm

5. EXPERIMENT RESULTS AND CONCLUSIONS

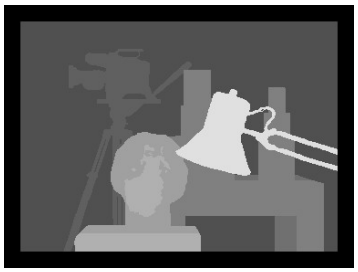
The standard test stereo pair "Tsukuba" is chosen for the performance experiments. The size of images is 384×288, and the maxima of disparity is D_{max} = 16. The size of match window of ZNCC is 7×7, and the size of evaluation window is 9×9. The match and evaluation results are shown in Fig. 1.

In Fig.2, (a) is the left image of the Tsukuba stereopsis, (b) is the Tsukuba ground truth, and (d) shows the result of the proposed method. For comparison, (c) gives the result of Boltzmann machine, whose structure is $N_w \times N_h \times (D_{max}+1)$. The ZNCC results are used to initialize the weights of Boltzmann machine, but no evaluations are made.

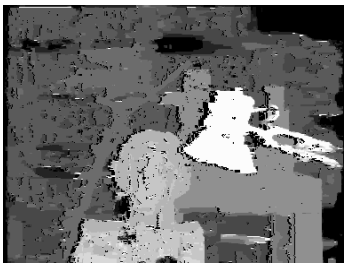
Experiments indicate the result of the method proposed in this paper is much better than that of Boltzmann machine obviously. Moreover, the speed of our method is 3~5 times faster than that of Boltzmann machine on average. After 40 iterations, the result of the proposed method shown in Fig.2(d) indicates this method possesses good performance and can escape from the local minima quickly.



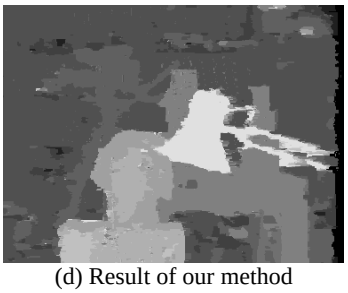
(a) Left image of Tsukuba



(b) Ground truth



(c) Result of Boltzmann machine



(d) Result of our method

Fig. 2. Performance comparison of Boltzmann machine and our method

6. REFERENCES

[1] N.M. Nasrabadi, W. Li, et al, "Use of Hopfield network for stereo vision correspondence", IEEE International Conference on Systems, Man and Cyberbetics, Cambridge MA, USA, 1989, pp.429-432.

[2] S. Nichani, "Solving the correspondence problem using a Hopfield network", IEEE International Conference on Neural Networks, Orlando, Florida, USA, June 1994, pp.4107-4112.

[3] C.H. Huang, J.H. Wang, "Stereo correspondence using Hopfield network with multiple constraints", IEEE International Conference on Systems, Man and Cybernetics, Nashville, Tennessee, USA, 2000, pp.1518-1523.

[4] R.Yassine, "Multilevel- and neural-network-based stereo-matching method for real-time obstacle detection using linear cameras", IEEE Transactions on Intelligent transportation System, 6(1), March 2005, pp.54-62.

[5] E. Binaghi, I. Gallo, et al, "Neural Adaptive Stereo Matching", Pattern Recognition Letters, 25(15), November 2004, pp.1743-1758.

[6] H.F. Hu, Y.G. Xiong, "A New Stereo Matching Approach Based on Hopfield Network", Journal of Image and Graphics, 9(6), June 2004, pp.729-736.

[7] Y.J.Xu, L.M.Du, et al, “An Improved Neural Network Utilized in Stereo Matching”, Journal of Image and Graphics, 3(10), October 1998, pp.845-848.
[8] S.Geman, D.Geman, “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 6(6), 1984, pp.721-741.
[9] S.Z.Li, Markov Random Field Modeling in Computer Vision, Tokyo, Japan: Springer-Verlag, 1995.
[10] Y.Boykov, O.Veksler, et al, “Fast approximate energy minimization via graph cut”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(11), 2001, pp.1222-1239.
[11] Y.Boykov, V.Kolmogorov, “An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Computer Vision”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(9), September 2004, pp. 1124–1137.
[12] V.Kolmogorov, R.Zabih, “What energy functions can be minimized via graph cuts?”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(2), February 2004, pp.147-159.

Research On Multi-examinations Method Fusion Based On Dempster-Shafer evidential theory And BP Artificial Neural Network On Failure Diagnosis For Military Electronic Equipment Module

Xinmin Huo, Linchun Xing
Self-propelled gun Department
WuHan Ordnance Noncommissioned officers Academy
WuHan 430075, China
Email: xlc86619759@163.com

ABSTRACT

The main repair means to military electronic equipment module is to replace the failure modules by new ones, according to this, by analyzing the characteristics of electronic module in the failure diagnosing process, a way which based on Dempster-Shafer (D-S) evidential theory And BP Artificial Neural Network was presented to fusing the characteristics information to find the failure electronic module. In this way, two methods and two fusing levels was introduced in the multi-examination methods fusion course of failure diagnosis for electronic equipment module. At last, an instance was given to certify that the method to the failure diagnosis for military electronic equipment module is feasible.

Keywords: multi-examination means fusion; electronic equipment; failure diagnosis; two level fusion.

1. INTRODUCTION

In conformity to the military electronic equipments servicing system， the main repair means to electronic equipment module is to replace the failure modules by new ones, according to this , during the failure diagnosing course, by analyzing the characteristics of electronic equipment module, a multi-detection methods fusion means was carried on by two fusion level. At the first level, the fusion result was worked out by fusing the outcome of each kind of detection methods on D-S evidence theory. At the second level, the result from the first fusion level was regarded as the input of an improved BP artificial neural network, by the second fusion, the final failure electronic module was found. At last, an instance about some kind of military computer was given to certify that the multi-detections method fusion based on D-S evidential theory and BP artificial neural network on failure diagnosis for military electronic equipment module was feasible.

2. THE ANALYSIS TO AN EXAMPLE

A certain kind of military computer is composed of 7 kinds of molectron module as main board, keyboard board, power transformer board, data adapter, liquid crystal screen, power supply board, GPS board and etc. In this example, we took two detection methods to finish the failure diagnoses, one was to measure the voltage signal, the other was to measure the data input and output signal. Fist, the D-S evidential theory was applied in the fist level of the fusion course; then according to the correlativity between the voltage detection and the data input and output detection, the fusion results of the fist fusion level was took into the BP artificial neural network arithmetic as the input to carry on the second fusion level, by this way to find out the failure electronic module.

By analyzing the 7 electronic modules of the computer, it could be found that the voltage signal was easy to be obtained, and the detection results was direct-viewing without any calculation, so it was not essential to be take into the fist fusion level. But to the data input and output signal, the signal of the main board is difficult to be gathered, and it only could be deduced by fusing the four detection results about data adapter, GPS board, keyboard board and liquid crystal screen. So to the data input and output detection methods, we only took the detection results about the data adapter, GPS board, keyboard board and liquid crystal screen as the samples of the first fusion level. After that, the voltage detection results and the data input and output fusion results were put into the second fusion level, then in conformity to the determinant rules to pick out the failure electronic module. In this example, the voltage signal is examined by a multi-meter and the data input and output signal is measured by software.

To the voltage detection method, the reliability of the failure state of the detection results was noted as 0 or 1.If the result isn't in the normal fields of the electronic module, it was noted as 1, otherwise, it was noted as 0. To the data input and output detection method, the reliability of the failure state of the detection results was noted as a reliability value in the first fusion level, the reliability value was confirmed by experts calculation and statistic of many failure samples. Before the second fusion level, the mapping from detection results space to the failure modules space was built in accordance with correlativity of the two detection methods to the same electronic module. For the ability of self adjustment and the disposal mapping, the Bp artificial neural network arithmetic was applied in the second fusion level, and all the results that were taken into the second fusion course were noted as 0 or 1. If the failure mode is appeared in the first fusion level, the failure state reliability is noted as 1, otherwise, if the failure mode is not appeared in the first fusion level, the failure state reliability is noted as 0. Only on this basis, the multi-detection methods fusion based on the D-S evidential theory and BP artificial neural network could be done.

3. THE FIRST FUAION LEVEL－FUSION BASED ON THE D-S EVIDENTIAL THEORY

3.1 The Fusion Arithmetic of D-S Evidential Theory

Some Concepts of The D-S Evidential Theory [1,2]： If each unattached possibility answer or the supposition is composed of a finite muster, the set will be noted as an identification frame Θ . All the possible musters of the Θ is note as 2^{Θ} , if the sum of elements are N in Θ , there

will be 2^N-1 elements in 2^{Θ}. On the assumption that Θ is a identification frame, if the muster function is accord with the term as:

$$\begin{cases} \text{The probabilit y of the impossibil ity case shall be} \\ m(\phi) = 0 \\ \text{The reliabilit y of all the elements in } 2^{\Theta} \text{ shall be} \\ \sum_{A \subset \Theta} m(A) = 1 \end{cases}$$

in the formula, m is noted as basic probability assignment; $\forall A \subset \Theta, m(A)$ is noted as the basic probability number of A, the function $Bel : m : 2^{\Theta} \to [0,1]$ defined by $Bel(A) = \sum_{B \subseteq A} m(B), \forall A \subset \Theta$ is noted as the reliability of A. The function $pl : 2^{\Theta} \to [0,1]$ which is defined by $pl(A) = 1 - Bel(\overline{A}) = \sum_{B \cap A \neq \phi} m(B), \forall A \subset \Theta$ is noted as the verisimilitude function of Bel, $\forall A \subset \Theta, pl(A)$ is noted as the verisimilitude degree of A, $pl(A)$ and shall be reflected as the implicit degree or credibility degree or verisimilitude degree of A.

3.2 The Synthesis Principle of D-S Evidential Theory [1, 3]

The synthesis principle of the D-S evidential theory is a principle which can reflect the evidence joint action, given some reliability function of the same identification frame and based on different evidences, if these evidences are not absolutely incompatible, the principles will be used to reckon a reliability function, and this function can be regarded as the function produced by joint action from those evidences.

1) The compages Rule About Two Reliability Functions: definition: in case of Bel_1 and Bel_2 are two reliability functions of the same identification frame Θ, m_1 and m_2 are their basic probability assignment, if $A \subseteq \Theta$ and $m(A) \rangle 0$, A is noted as the focus element. If there are some focus elements as $A_1, A_2, \cdots, A_k$ and $B_1, B_2, \cdots B_n$ presume that $\sum_{A_i \cap B_j = \phi} m_1(A_i) m_2(B_j) \langle 1$, the resultant basic probability assignment $m : 2^{\Theta} \to [0,1]$ will be denoted as:

$$m(A) = \begin{cases} 0, & A = \phi \\ \dfrac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1 - \sum_{A_i \cap B_j = \phi} m_1(A_i) m_2(B_j)}, & A \neq \phi \end{cases}$$

The reliability function given by m is noted as the straight summation of Bel_1 and Bel_2, and expressed as $Bel_1 \oplus Bel_2$. If the formula above is not exist, then the straight summation of Bel_1 and Bel_2 will not be exist

2) The Compages Rule About Many Reliability Functions: principle: suppose that $Bel_1, Bel_2, \cdots, Bel_n$ are reliability functions of the same identification frame Θ, $m_1, m_2, \cdots, m_n$ are their basic probability assignment, if $Bel_1 \oplus \cdots \oplus Bel_n$ is exist and the basic probability number is m ,then the synthesis of n reliability functions will be showed as: $Bel = \{[(Bel_1 \oplus Bel_2) \oplus Bel_3] \oplus \cdots\} \oplus Bel_n$, and the final evidence gathered from the compositive evidence has nothing to do with the order.

3.3 The Failure Determination Rules of Evidence Theory [4]

After acquired the reliability area $[Bel(F_j), pl(F_j)]$about all the proposition of the identification frame Θ from the evidences and the incertitude $m_j(\theta)$ of the evidences, we can make the conclusion according to the below rules:
①the determination failure style should be the maximal reliability function, and should be bigger than some threshold.② the margin of the reliability between the determination failure style and other failure styleshould be bigger than some threshold.③the reliability function value of uncertain failure style must be less than some threshold.④the reliability function value of the determination failure style should be bigger than other uncertain ones.

3.4 An Instance of Evidence Fusion Theory

It could be concluded from the above analysis, in this example, the first fusion level was mainly toward the data input and output methods about the data adapter, the GPS board, the keyboard and the liquid crystal screen of the military computer. The establishment of the identification frame, the determination of the evidences, and the assignment of the reliability was showed below:

1) The Establishment of Identification Frame: To the data input and output detection method to the electronic module, the identification frame will be confirmed as $\Theta = \{F_1, F_2, F_3, F_4, F_5\}$ by experts inference and probability statistic of many failure samples, in it, F_1 expresses the data adapter failure, F_2 expresses the GPS board failure, F_3 expresses the keyboard failure, F_4 expresses the liquid crystal screen failure, F_5 expresses the mainboard failure.

2) The Determination of Evidences: By analyzing the failure mechanism of the identification frame Θ, we choose the detection result information parameter of the data adapter, the GPS board, the keyboard and the liquid crystal screen from the data input and output detection methods, we regarded each detection result information parameter as a evidence, so, there were four evidences.

3) The Assignment of Evidences: Thinking of the corresponding relationship between the detection result information parameter of the four electronic module from the data input and output detection method and failure mode and layers of data transmission and the uncertainty of the detection result θ, we make a unitary disposal, the result is showed in the Table 1., the value in the table expressed the reliability assignment.

4) The Fusion and the Decision-Making of the Evidences: if all the evidences in the table 1 were not occur, it showed that the input and output of all the electronic module was normal. If only one evidence were occurred, and it showed that only the electronic module which opposite to the evidence was failure, this two case were not necessity to carry on the fusion on this level, we will go

straight into the next fusion level. If more than two kinds of the evidences were occurred, the first fusion level based on D-S evidence theory would be carried on. According to the failure determination rules of D-S evidence theory, in this example, we choose α=0.5，β=0.3，γ=0.1，then the reliability area and the diagnosis result of each evidence from the fusion would be expressed in the Table 2.:

It can be seen from Table 2. , if we done the state identification of the failure electronic module only by one or less detection result, the reliability was low. If we do the information fusion by more detection results, it would effectively enhance the exactitude identification rate, so in the first fusion level, we tried to use more detection result to do the discursion.

Tab1e 1. The assignment of the evidences

serial number	evidences	identification frame						Reliability summation
		data adapter failure	GPS board failure	Keyboard failure	liquid crystal screen failure	Main board failure	θ	
1	the input and output of the data adapter is abnormal	0.6	0	0	0	0.3	0.1	1
2	the input and output of the GPS board is abnormal	0	0.6	0	0	0.3	0.1	1
3	the reaction of the keyboard is abnormal	0	0	0.6	0	0.3	0.1	1
4	The display of the liquid crystal screen is abnormal	0	0	0	0.6	0.3	0.1	1

Table 2. The reliability area and the diagnosis result of each evidence

evidences	$m_i(\theta)$	$[Bel(F_i),pl(F_i)]$					diagnosis result
		F_1	F_2	F_3	F_4	F_5	
1&2	0.04	0.21 0.25	0.21 0.25	0 0.04	0 0.04	0.54 0.58	F_5
1&3	0.04	0.21 0.25	0 0.04	0.21 0.25	0 0.04	0.54 0.58	F_5
2&3	0.04	0 0.04	0.21 0.25	0.21 0.25	0 0.04	0.54 0.58	F_5
1&2&3	0.01	0.07 0.08	0.07 0.08	0.08 0.09	0 0.01	0.77 0.78	F_5
1&2&3&4	0.003	0.021 0.024	0.021 0.024	0.024 0.027	0.018 0.021	0.913 0.916	F_5

4. THE SECOND FUSION LEVEL BASED ON BP NEURAL NETWORK

On the first fusion level basis, in conformity to the fact, we could found the correlative relationship between the two detection methods, if the voltage of some kind of the electronic module were abnormal, the input and output of it must be abnormal, if the voltage of some kind of the electronic module were normal, the input and output of it would be uncertain normal or not. According this, we should carry on the second fusion level in this examples, an improved BP artificial neural network was applied in the

second fusion course. And the probability value of the reliability function was noted as 0 or 1, if the detection result were in the normal working scope of the electronic module, the probability value of the failure subjection degree would be noted as 0, otherwise, if the detection result were not in the normal working scope of the electronic module, the probability value of the failure subjection degree would be noted as 1. In a word, except the power supply board and power transformer board (because these two module could be recognized failure or nor direct from the voltage detection result), the correlation relationship about the detection result space and the failure mode space of other electronic modules can be expressed as: if the input vector is (0, 0), the output vector is 0; if the input vector is (1, 1), the output vector is 1; if the input vector is (0, 1), the output vector is 1.

4.1 The Establishment of BP Neural Network

The Establishment of The Input Nod Sum and Output Nod Sum And The Initialization of The Weight Value And The Limen Value [4]: The sum of BP network layer was selected as 3 , and the sum of the detection methods was regarded as the sum of the input neuron units, and the sum of the failure modes was regarded as the sum of the output neuron units .In this illustration ,the sum of the input neuron units was selected for N_1=2 , the sum of the output neuron units was selected for N_3=1 , the sum of the neuron units of the latent layer was selected for N_2=5 .The weight value and the limen value was initialized as any decimal fraction among the field of（-1，1）by random function when calculating by the c++ procedure of this example

1) The establishment of training samples for the inputed and outputed matrix: According to the ideas of multi-detection methods fusion, except the power supply board and power transformer board, the relationship between the detection result information obtained from the two kinds of detection methods and the training samples was illustrated as Table 3..

Table 3. The studying samples of the second fusion level

serial number	training samples		teacher signals
	Voltage detection	input and output detection	
1	0	0	0
2	0	1	1
3	1	1	1

So, the inputed matrix of raining samples was:

$$I = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$

the outputed matrix of raining samples was:

$$o = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

In the c++ procedure of the improved BP neural network algorithm (in which a momentum item was added to the standard algorithm), learning efficiency was selected for $\eta_1 = 0.6$, $\eta_2 = 0.7$, Momentum coefficient was selected for $\alpha = 0.2$; the error field was selected for E <0. 001, then take the training samples matrixes into the procedure to do the training.

2) The Analysis of Identification Results: substituting the sample into the trained procedure to make identification, the matrix resulted shall be shown as follows:

$$T = \begin{pmatrix} 0.0328 \\ 0.9671 \\ 0.9697 \end{pmatrix}$$

As the judgment principle of the failure module, it can be found from the result of accuracy rate of the identification of the training samples is 100%. It can be seen that a mathematics model established on BP neural network which can make a well disposal on the complex and nonlinear mutuality about the detection result space to the failure mode space.

5. CONCLUSION

There is a unique superiority on the expression and the processing the indefinite information aspect for D-S evidence theory, which can fetch up the indeterminacy of the single evidence by fusing each evidence of the same identification frame, and make the dependability of the evidence ratiocination enhanced, so it has been widely used in the information fusion area. Otherwise, The D-S evidence theory doesn't has the function of dealing with the complex and nonlinear mutuality about the detection result space to the failure mode space, but BP neural network can do this better, and BP neural network also has the ability about distributed memory, parallel disposal, self-organization and self-studying and has the advantages about self adjustment, contain error and robustness, so it has made wide attention in the failure diagnosis of the electronic area. When the sum of the diagnosis parameter and the diagnosis information is large, the network structure of Bp neural network will be too large to training. Because of this and under the direct of the idea to enhance the automatization and intelligentize level of maintain for military electronic equipment, in the failure diagnosis course of the electronic module, we choose the way about two level fusion by multi-detection methods based on the D-S evidence theory and a improved BP neural network. Proved by examples, this method can obtain a good diagnosis result, it is feasible.

6. REFERENCES

[1] H.W. Reng, F.Q. Deng, "Information Fusion on Failure Diagnosis Methods Based on The Evidence Theory", Systems Engineering and Electronic Technique, 2005, 27(3).

[2] G.J. Zhang, S.Q. Zhou, et al, "Synthetical Failure Diagnosis Methods Based on D-S Evidence Theory", Journal of Logistics Engineering, 2004, 20(4).

[3] F.T. Wang, X.J. Ma, et al, "The Application on Information Fusion About The Equipment Failure Diagnosis Based on D-S evidence theory", Journal of Dalian Science And Technique Engineering, 2003, 43(4).

[4] D.Q. Zhu, "The Principle and Practice of the Failure Diagnosis of Electronic Equipment", Beijing: Electronic Industry Press, 2004, 121-180.

Fault Diagnosis of Analog Circuits Based on Wavelet Neural Network

Guoming SONG, Houjun WANG, Hong LIU, Shuyan JIANG
School of Automation Engineering, University of Electronic Science and Technology of China
Chengdu, Sichuan 610054, P.R.China
Guoming SONG
Computer Engineering Department, Chengdu Electromechanical College
Chengdu, Sichuan 610031, P.R.China
Hong LIU
School of Computer Science and Technology, Changchun University of Science and Technology
Changchun, Jilin 130022, P.R.China
Email: guoming905@yahoo.com.cn

ABSTRACT

A fault diagnosis method for analog circuits was proposed. Wavelet neural network was used to detect the faults and classify default patterns in circuits under test. To construct this new network, wavelet functions were embedded in a neural network as neuron stimulating functions, which combined the multi-scale analysis merit of wavelet transform with nonlinear mapping capability of neural network. Wavelet multi-resolution decomposition was adopted to extract fault features after sampling the impulse response of the analog circuit. The features were applied to train the wavelet neural network for further fault patterns recognition. Diagnosis principle and procedure were elucidated. Simulation results on band pass filter circuit show that the presented approach has good capability of fault identification and diagnosis.

Keywords: Fault Diagnosis; Analog Circuit; Wavelet Neural Network; Fault Feature Extraction; Pattern Recognition.

1. INTRODUCTION

It's necessary and important for us to detect the faults by the aid of computer in microelectronic and semiconductor industry as the application of LSI technology grows rapidly. The study on fault diagnosis techniques is significant in the theory and application fields. Comparing with variously mature fault diagnosis techniques of digital circuits, the diagnosis techniques for analog circuits develop slowly, although they have been researched for a long time. As the fault diagnosis techniques evolved, the researchers noted that some difficulties impeded their fast development. The following facts are the main factors affect the progress of fault diagnosis techniques [4]: (1) It is complicated and difficult to construct fault detecting model because of the diversity characteristic of analog circuits, which results in the output response and component parameters varied continuously. (2) Many traditional diagnosis approaches cannot achieve testing veracity and stability since some factors will cause complex relationship between input and output and deviation of working properties, such as nonlinearity, component parameter tolerance and noise etc. (3) High integration of LSI and VLSI generates difficulty in circuit test and fault diagnosis. The actual circuits are multilayered and encapsulated, especially for integrated circuits, so that accessible ports and nodes are limited for measuring. The insufficient information used for fault detection will cause the unreliability and fuzziness of diagnosis.

Generally, two approaches are used for analog circuit diagnosis [1]: simulation before test (SBT) and simulation after test (SAT), which are classified according to the actual test occurring prior or posterior to the circuit test. The former contains fault dictionary and pattern recognition method, while the latter includes fault parameter identification and fault verification method. As the integration of analog and mixed circuits is increased and the topology of LSI or VLSI becomes more complex, the SBT has advantage of much less test time consuming [2]. As a practical method, the fault dictionary was employed in fault detection and diagnosis for a long time. The main idea of fault dictionary is to extract circuit features under various fault states and construct a fault dictionary firstly, and then fault detection and location are realized by comparing the feature of circuit response with fault dictionary in testing stage. But it is a tough work to construct a fault dictionary containing all faults because of the complicated fault mechanism of analog circuits, especially for large-scale analog circuits. Several neural network based methods used in this field have been proved robust, fast and suitable for real-time fault diagnosis in recent years for its strong capability in tackling classification and nonlinear problem [5,6,7,8]. More recently, wavelet transform is successfully employed in feature extraction and this presents some new approaches in fault diagnosis of analog circuits, especially the technique of combining wavelet transform with neural networks. Two forms of combination are expected to deal with fault detection: (1) Wavelet transform is used as a preprocessor to optimizing the input features of the neural network [6]; (2) Wavelet functions are embedded in the neural network to be stimulating functions of neurons or weight functions between two layers of neurons [3, 10]. In this paper, we adopt the second form to design a wavelet neural network (WNN) for analog circuits fault diagnosis. Multi-resolution analysis of wavelet transform and data normalization is employed to get optimal fault feature vectors after the output response signals are sampled. The WNN is trained by using the fault features as training patterns. A filter circuit was detected by the WNN in the simulation experiment. The results also mentioned that the WNN has good capability of fault detection and location in analog circuits, and high diagnosability is obtained as well.

2. FAULT FEATURE EXTRACTION BASED ON WAVELET ANALYSIS

Wavelet transform can provide more precise decomposition in whole frequency band of an original signal and is suitable for non-stationary signals. It is

applied generally to analyze the complicated signals in many fields, such as pattern recognition, image processing and fault detection because of good capability of feature extraction and localization in time and frequency domain. The main advantage of wavelets is that they have a varying window size, being wide for low frequencies and narrow for the high ones, thus leading to an optimal time-frequency resolution in all the frequency ranges.

A family of wavelet is derived from the translations and dilations of a single function. If *Ψ(t)* is the original function, referred to as the mother wavelet, the members of the family are given by

$$\Psi_{a,b}(t) = |a|^{\frac{1}{2}} \Psi\left(\frac{t-b}{a}\right), a, b \in R, a \neq 0 \qquad (1)$$

They are indexed by two labels a and b, where a indicates the dilation and b the translation of the mother wavelet *Ψ(t)*. Additionally, a scaling function, *ϕ(t)*, is used that can be translated and dilated in the same way. The wavelets are derived from a so-called mother wavelet by dilation and translation factors. The mother wavelet is normalized with zero average and meets the following admissible condition

$$C_{\omega} = \int_0^{\infty} \frac{|\Psi(\omega)|^2}{\omega} d\omega < \infty \qquad (2)$$

The continuous wavelet transform of a given signal $f(t) \in L^2(R)$ is defined as the inner product of the wavelet function and the signal, i.e.

$$W(a,b) = \left\langle f(t), \Psi_{a,b}(t) \right\rangle = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \Psi^*_{a,b}\left(\frac{t-b}{a}\right) dt \qquad (3)$$

where $f(t)$ is the signal to be analyzed and $\Psi_{a,b}(t)$ is the wavelet function.

If we select $a=2^j$ and $b=k2^j$, $(j, k) \in Z^2$ in equation (3), *W (a, b)* translates into a binary wavelet transform. MRA (multi-resolution analysis) is used to carry out the transform of $f(t)$, which is divided into two parts at different scales: the low frequency part and the high frequency part. The low frequency part is called an approximation of the original signal, while the high frequency part called details respectively. The decomposition equation is

$$\begin{cases} c_{j-1}(k) = \sum_{k \in z} h^*(k-2n) c_j(k) \\ d_{j-1}(k) = \sum_{k \in z} g^*(k-2n) c_j(k) \end{cases} \qquad (4)$$

where *h(k)* and *g(k)* are two-scale series; *h*(k)* and *g*(k)* are complex conjugate functions of *h(k)* and *g(k)*.

After getting the decomposition coefficients of the binary wavelet transform at various scale spaces and calculating the sum of each coefficient's absolute value, the fault feature vectors of analog circuit are constructed by ranking this sum value according to the scale sequence. It contains four main steps:

Step1. Decompose the sampling signal using binary wavelet transform with *N* level, ranging from the first level to the Nth level, and obtain *N+1* decomposition coefficients sequence {*d1,d2,…, dN, cN*}.

Step2. Calculate the absolute value sum of decomposition coefficients at various levels. Suppose *CN* is the absolute value sum of sequence *cN*, which refers to the low frequency decomposition coefficient at the Nth level; while *Dj* is the absolute value sum of sequence *dj*, which is the high frequency decomposition coefficient at the jth level, respectively.

Step3. Construct feature vectors {*D1, D2, …, DN, CN* }.

Step4. Data normalization. In order to avoid a strong deviation of input data, data must be normalized and limited in area [-1, 1] by implementing some transform processing.

3. FAULT DIAGNOSIS BASED ON WAVELET NEURAL NETWORK

The fault diagnosis means pattern recognition of circuit states. The artificial neural network has obvious advantage of pattern recognition, for the capability of associated memory, large scale parallel processing, real time computing and easy implementation on hardware. It is an effective method for diagnosing tolerance circuits because of the trait of error tolerance. In recent years, artificial neural network as a novel technique of the information soft processing is applied widely in the field of diagnosis. Neural networks are employed in pattern recognition of faults to realize the completely non-linear mapping between fault features and fault classification, but their accuracy relies heavily on the parameter effectiveness of fault features and parameter selection of network's framework. Wavelet transform is good for extracting fault features of non-stationary and time-variable signals for the characteristic of time-frequency domain localization, which specify the important information on circuit running state in details. It can acquire the feature vectors as few as possible on condition that it can show the fault behavior. If we combine the wavelet transform with neural networks to develop a new method for fault diagnosis, the veracity of fault diagnosis will be improved.

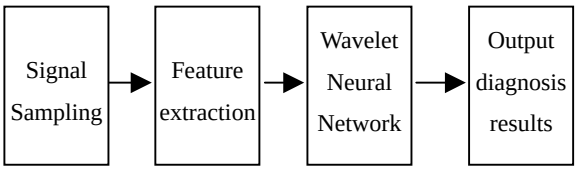


Fig. 1. Fault diagnosis flowchart based on WNN

The process of fault diagnosis in analog circuits based on WNN is illustrated in Fig.1. This diagnosis method identifies the fault class by the output of the wavelet neural network trained with various fault patterns.

The framework of WNN we proposed for fault diagnosis is shown in Fig.2. It is three-layer feedback architecture with wavelet functions, which allows converging at its global minimum in a short time. A wavelet base is employed in the hidden layer of WNN rather than a sigmoid function, which discriminates it from general back propagation neural networks.

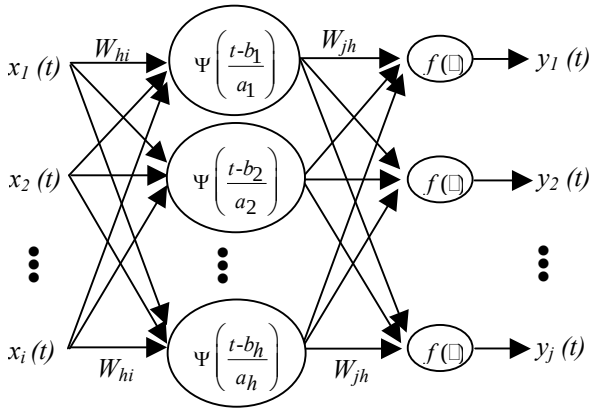


Fig. 2. Structure of wavelet neural network

The function of mapping from input vector space to output vector space can be expressed as [9]

$$y_j^p(t) = f\left(\sum_{h=1}^{H} w_{jh} \Psi\left(\frac{\left(\sum_{i=1}^{I} w_{hi} x_i^p(t) + b1_h \right) - b_h}{a_h} \right) + b2_j \right) \quad (5)$$

In equation (5), $\Psi(\square)$ and $f(\square)$ are the stimulating function of hidden layer and output layer, while $\Psi(\square)$ is a wavelet function and $f(\square)$ a tangent sigmoid function where $f(x) = \frac{1}{1+e^x}$. $x_i^p(t)$ and $y_j^p(t)$ are the ith input and the jth output at the pth pattern. $b1_h$ and $b2_j$ are the threshold of the hth hidden layer node and the jth output layer node, respectively. The sum square error performance function is expected to reach minimum by feeding information forward and feeding errors back, thus updating the weights and bias parameters according to its learning algorithm. A momentum factor and adaptive learning rule is adopted to reduce the sensitivity of the local details of error surfaces and shorten the learning time on speculating the experience. The training algorithm of WNN is elucidated as following steps:

Step1. Data acquisition and fault feature extraction: The output response voltage signals of test point are sampled in terms of the CUT running in different fault patterns. Then optimal features for training neural networks are obtained by wavelet decomposing coefficients and normalization according to the four steps described before.

Step2. Select the parameters of wavelet neural network: The dimension of fault feature vectors and the circuit fault pattern determines the number of input node and output node. The output number is equal to the number of fault classes, while input number is equal to the dimension of feature vectors. The number of neurons on hidden layer is set greater or equal to $\sqrt{M+N}+a$, where M and N is the node number of input layer and output layer and a is set 1~10.

Step3. Training of wavelet neural network: The fault feature vectors, as input vectors of training pattern, are used to train the wavelet neural network. The output vectors reflect the states of CUT. The gradient descent algorithm is employed to minimize the error function so as to adjust the weights of network. The error function is

$$E = \frac{1}{2} \sum_{p=1}^{P} \sum_{j=1}^{J} \left(y_j^p(t) - \bar{y}_j^p(t) \right)^2 \quad (6)$$

where P is the total number of training patterns; $\bar{y}_j^p(t)$ and $y_j^p(t)$ are the desired and real output associated with the jth feature for the pth neuron.

To minimize the sum square error function in Eq. (6), the weights and coefficients in Eq. (5) can be updated using the following formulas:

$$w_{jh}(t+1) = w_{jh}(t) + \eta \Delta w_{jh} + \alpha \left[w_{jh}(t) - w_{jh}(t-1) \right] \quad (7)$$

$$w_{hi}(t+1) = w_{hi}(t) + \eta \Delta w_{hi} + \alpha \left[w_{hi}(t) - w_{hi}(t-1) \right] \quad (8)$$

$$a_h(t+1) = a_h(t) + \eta \Delta a_h + \alpha \left[a_h(t) - a_h(t-1) \right] \quad (9)$$

$$b_h(t+1) = b_h(t) + \eta \Delta b_h + \alpha \left[b_h(t) - b_h(t-1) \right] \quad (10)$$

$$b1_h(t+1) = b1_h(t) + \eta \Delta b1_h + \alpha \left[b1_h(t) - b1_h(t-1) \right] \quad (11)$$

$$b2_j(t+1) = b2_j(t) + \eta \Delta b2_j + \alpha \left[b2_j(t) - b2_j(t-1) \right] \quad (12)$$

where η is the learning rate for updating weights/coefficients $w_{jh}, w_{hi}, a_h, b_h, b1_h, b2_j$; and α is the momentum factor, respectively.

Note that $f'(x) = f(x)(1 - f(x))$, thus we can deduce the following equations by using Eq. (5) and Eq. (6):

$$\Delta w_{jh} = -\frac{\partial E}{\partial w_{jh}} = \delta_j x_h^p \quad (13)$$

$$\Delta w_{hi} = -\frac{\partial E}{\partial w_{hi}} = \delta_h x_i^p(t) \quad (14)$$

$$\Delta b_h = -\frac{\partial E}{\partial b_h} = -\delta_h \quad (15)$$

$$\Delta a_h = -\frac{\partial E}{\partial a_h} = -\delta_h \frac{x_h^p - b_h}{a_h} \quad (16)$$

$$\Delta b1_h = -\frac{\partial E}{\partial b1_h} = \delta_h \quad (17)$$

$$\Delta b2_j = -\frac{\partial E}{\partial b2_j} = \delta_j \quad (18)$$

where $\delta_j = -(x_j^p - \bar{x}_j^p) \times x_j^p \times (1 - x_j^p)$

and $\delta_h = \sum_j \delta_j w_{jh} \Psi'\left(\frac{x_h^p - b_h}{a_h} \right) \times \frac{1}{a_h}$.

Step4. Fault diagnosis: Inputting the trained WNN with the measured data from the CUT due to the change in the feature values, the outputs of the WNN will show the fault pattern.

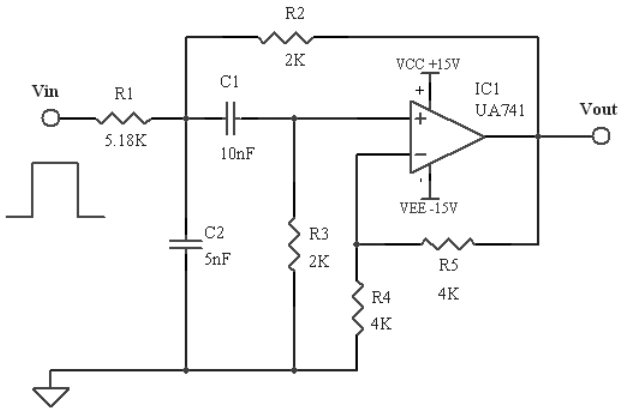


Fig. 3. Circuit under test

4. SAMPLE CIRCUIT AND FAULT

The circuit under test chosen is shown in Fig. 3. The nominal values of its components are given in the figure. The tolerance of resistors and capacitors is assumed as 5% and 10%. The simulated faults are divided into two categories: hard faults (or catastrophe faults), which refer to the short or open defect of components, and soft faults (or component parameter faults), which are derived from the component parameters exceeding permitted tolerance bound. The hard fault model is realized by using 0.0 1 Ω (short defect) and 100M Ω (open defect) bridge resistors. If we assume that R1, R3, C1 and C2 are 50% higher or lower than their respective nominal values shown in Fig. 3. Hence there are eight class soft faults. Additionally, open defect of R2 and short defect of R5 are supposed also. We have the fault classes R1↑, R1↓, R3↑, R3↓, C1↑, C1↓, C2↑, C2↓, R2 open and R5 short. The symbol ↑ represents higher than nominal values and ↓ lower, respectively. Considering the no fault state, eleven fault models are achieved. To generate training data for different fault class, we set a faulty component in the circuit and vary the other resistors and capacitors within their tolerances. Pspice software is used for circuit modeling and simulating. The output response signals of the circuit are sampled, and then 100 Monte Carlo analyses are conducted for every fault pattern of the sampled circuit with tolerance to extract candidate features. 70 of them as training patterns are used to train WNN and the rest 30 are employed for simulation.

5. DIAGNOSIS RESULTS

In this section the WNN is utilized to simulate fault diagnosis of analog circuit. A single impulse signal of height 5V and duration 10μs is used as input signal, and the response signals V_{out} of the filter circuit are simulated to train WNN after implementing feature extraction described in section2. The sampled data is decomposed in four levels with Harr wavelet to construct feature vectors. WNN architecture of 6-H-11 is adopted to diagnose the filter circuit. The network has 6 input neurons and 11 output neurons. H is the neuron number of hidden layer; here we set its value as 15 in accordance with the selection rules elicited before. Morlet wavelet is selected for hidden layer function because of its good difference property. The design and training of WNN were completed in MATLAB 6.5 simulation environment.

The percentages of correct detection for each type of fault are shown in Table 1. The no fault and hard fault class achieves a 100% detection rate, while other fault classification rate is satisfactory. Comparing with other neural network diagnosis methods, it presented a higher correctness of fault diagnosis and faster convergence speed, although the WNN proposed in this paper is a little more complicated.

Table 1. Diagnosis Results of WNN

Fault Class	Detection Rate
No Fault	100%
R1↑	90%
R1↓	100%
R3↑	100%
R3↓	97%
C1↑	100%
C1↓	93%
C2↑	97%
C2↓	100%
R2 Open	100%
R5 Short	100%

6. CONCLUSIONS

A novel fault diagnosis method for analog circuits is proposed in this paper. This technique uses wavelet neural network, in which wavelet functions are adopted as the stimulating functions of neural network, to realize correct fault classification of analog circuits. The advantages of good feature extraction and pattern recognition derived from wavelet transforms and neural network are combined effectively in fault diagnosis and produce high diagnosis correctness. The simulation results demonstrated that it is a promising method for fault diagnosis of analog circuits. The described approach can be developed for multiple fault diagnosis of analog circuits, although a single fault case is considered and applied in this paper. The further work will be focused on optimizing the training algorithm and selection of mother wavelet to improve the network's performance, including simplifying network architecture, faster training speed of network and higher fault detection rate.

7. REFERENCE

[1] S.Y. Yang, "Fault Diagnosis and Reliability Design of Analog Systems", Beijing: Tsinghua University Press, 1993.

[2] M. Catelani, A. Fort, "Soft fault detection and isolation in analog circuits: some results and a comparison between a fuzzy approach and radial basis function networks", IEEE Transaction Instrument & Measurement, 2002, 51: 196-202.

[3] B.Q. Lv, T.Z. Li, " A Fault Detection Approach Based on Wave-net", Journal Control Theory and Application, 1998, 15(5):802-805.

[4] J. Huang, Y.G. He, "The State-of-the-Art of Fault Diagnosis of Analog Circuits and Its Prospect", Journal

Microelectronics, 2004, 34(1): 21-25.
[5] F. Aminian, M. Aminian, et al, “Analog fault diagnosis of actual circuits using neural networks”, IEEE Transaction. Instrument & Measurement, 2002, 51(3):544-550.
[6] M. Aminian, F. Aminian, “Neural-network based analog-circuit fault diagnosis using wavelet transform as preprocessor”, IEEE Transaction Circuits and Systems II: Analog Digital Signal Process, 2000, 47(2):151-156.
[7] F. Aminian, M.Aminian, et al, “Analog fault diagnosis of actual circuits using neural networks”, IEEE Transaction Instrument & Measurement., 2002, 51:544-550.
[8] M. R. A shouri, “Fault detect ion of analog circuits using neural networks and Monte-Carlo analysis”, Proceeding 44th Midwest Symposium, 2001, 2:14-17.
[9] X. Wen, L. Zhou, et al, “Neural network application design based on MATLAB”, Beijing: Chinese Science Press, 2000.
[10] I. Statish, “Short-time Fourier and Wavelet transforms for fault detection in power transformers during impulse tests”, Proceeding IEEE Sci. Meas. Technol, 1998, 145(2):77-84.

An Improved ART1 Neural Network Learning Algorithm

Qingbing Sang [1], Ying Yin [2]
1. Information Engineering School, Southern Yangtze University,
Wuxi, China, 214122.
Email: sangqb@163.com
2. Science School, Southern Yangtze University,
Wuxi, China, 214122.
Email: sqbyy@sina.com.cn

Abstract

Since inadequacy exists in the traditional ART1 neural network, an improved calculating method of similarity is proposed. This method considers the corresponding position value of the two vectors at the same time, thus avoiding different results of similarity because of different inputting sequence of the two vectors in the ART1 neural network. In order to solve the pattern-drifting problem of ART1 neural network, the principle of minority subordinate to majority is proposed to reduce the occurrence of this type of problem, improving the application effect of ART1 neural network.

Keywords: neural network, Adaptive Resonance Theory, pattern classification, pattern drifting.

1. Introduction

Adaptive Resonance Theory, or ART for short [1], is a neural network pattern proposed by S.Grossberg and A.Carpenter of Boston University in the United States in 1976. ART instructs learning through the dynamic principle of biological neurons' self-excitation and lateral inhibition and makes the input pattern realize resonance through the identification and comparison of the network two-way link to finish its own memorization, and realize association in the same way. As a feedback artificial neural network for teacher-free learning, ART network can be used for on-line learning and is suitable for the classification of complicated systems. ART network is classified into three types, i.e. ART1, ART2 and ART3. ART1 network has a wide application and is a closed-loop self-adaptive system. It has the self-organizational, self-stable and self-learning ability, demonstrating good self-adapting ability and plasticity [2, 3].

2. The Operating Principle of the ART1 Network [4, 5]

The ART1 neural network consists of the input layer, the comparison layer and the identification layer. When the network operates, it receives the input pattern from the external environment and detects the matching level between the input pattern and all the patterns of the identification layer. For the patterns, which have the highest similarity, the network needs to further inspect the similarity between the typical vector of this pattern and the current input pattern. If it is higher than the threshold value of similarity, this pattern shall be put into this category. If it is not higher than the threshold value of similarity, it shall match the similarity among the rest nodes. If all the nodes cannot be matched for similarity, a new pattern shall be automatically added to match this input pattern.

1) Initialize the weight of the instar and outstar weight vector: Suppose t_{ij} represents the outstar weight vector: the vector from node j of the identification layer to node i of the comparison layer. Order t_{ij} (0) =1; suppose b_{ij} represents the instar weight vector: the vector from node i of the comparison layer to node j of the identification level. Order $b_{ij}(0)=1/(1+n)$, in which $i=1,2,\cdots,n$; $j=1,2,\cdots,m$.

2) Input the external signal X, $X=\{x_i\}^n=(x_1,x_2,\cdots,x_n)$, $i=1,2,\cdots,n$, in which xi can take the value 0 or 1.

3) Calculate the maximum matching level net_{j*}:

$net_j=\sum_{i=1}^{n} b_{ij}x_i$, $j=1,2,\cdots,m$, $net_{j*}=\max\{net_j\}$, in which $j*$ is the winner node.

4)Detection of similarity: $N_0=\sum_{i=1}^{n} t_{ij}x_i$,

$N_1=\sum_{i=1}^{n} x_i$, $i=1,2,\cdots,n$, $j=1,2,\cdots,m$.

5) Comparison of similarity: Suppose ρ is the threshold value of similarity. If $N_0/N_1 \geqslant \rho$, turn to 7); If $N_0/N_1<\rho$ and all the nodes have been matched, turn to 6); If $N_0/N_1<\rho$ and all the nodes have not been matched, turn to 3) and search for the next node with the highest matching level.

6) Add a node, i.e. now there are m +1 output nodes. The weight of $t_{i,m+1}$ and the weight of $b_{i,m+1}$ are the initialization value, i=1,2,⋯,n.

7)Weight adjustment: $t_{ij*}(t+1)=t_{ij*}(t)x_i$,

$$b_{ij*}(t+1)=\frac{t_{ij*}(t+1)}{0.5+\sum_{i=1}^{n} t_{ij*}(t+1)},\quad i=1,2,\cdots,n\ 。$$

8) Accept the new input and turn to 2).

3. The problems and improvements of ART1 neural network

3.1 Inadequacy and improvement of the regular similarity judgment

The similarity of regular ART1 is to compare the outstar weight vector T_{j*} of the winner node and the number of "1" in the corresponding position of the input vector X. It can be expressed as follows:

$$\rho_1 = \frac{N_0}{N_1} = \frac{\sum_{i=1}^{n} t_{ij^*} x_i}{\sum_{i=0}^{n} x_i}, \quad i = 1,2,\cdots,n$$

This regular similarity judgment does not consider the role of "0" in the outstar weight vector and only takes into account the function of "1". On the occasions when only the state indicated by "1" needs to be considered, this method is feasible. In the actual applications, however, 0 and 1 represents two states respectively and they are both useful information for judgment. They are equally important. Therefore, this method has obvious inadequacies. The example of identifying English letters shall be used for comparison here.

Example 1: Fig. 1 is the model figure of English letters, among which, (a) is the standard "E" mode signal, (b), (c) and (d) are the noise signal, and (e) is the standard "F" mode signal. For the calculation of similarity, the blackened part is "1" and the blank part is "0". The coding of each signal is as follows:

*X*1＝111111000010000111101000010000111111，

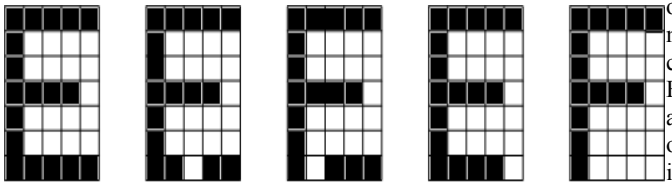
*X*2＝111111000010000111101000010000111011，

*X*3＝111111000010000111101000010000110111，

*X*4＝111111000010000111101000010000111110，

*X*5＝111111000010000111101000010000110000

When the signal (b) is input (X=X2) for the second time, calculate the matching similarity between X2 and each node of the competition layer and it is the maximum at the first node (higher than that at other unused nodes). According to formula (2), the value $\rho 1$ from its similarity formula is 1.000, indicating that (b) and (a) have the same signal. From the figure it can be clearly seen that there exists difference between (b) and (a). If the input sequence of (a) and (b)A is reversed, the similarity $\rho 1$ according to formula (2) is 17/18 ＝0.944. In terms of (a) and (b), different similarity is obtained under difference input sequence. This is because in the comparison of similarity, as long as t_{ij^*} is 0, x_i on its corresponding position is also 0. At this time, on the corresponding position where t_{ij^*} is 1, no matter the input xi is 1 or 0, the obtained similarity ρ1 is bound to be 1.



X1 X2 X3 X4 X5
(a) (b) (c) (d) (e)

Fig. 1. The mode figure of English letters

In order to compensate for this inadequacy, this paper designs a simple, prompt and accurate computing method for similarity. The guiding principles used here is to consider the state of 0 and 1 at the same time and find out the number of similar status on the corresponding position of t_{ij^*} and x_i, i.e. the specific value between the numbers of exclusive nor value and n the dimensions of the input vector. The computing method of this similarity is as formula (3) shows. This method can accurately compare the two vectors, free from the problem of which one is more important.

$$\rho_2 = \frac{\sum_{i=1}^{n} t_{ij^*} x_i + \sum_{i=1}^{n} (1 - t_{ij^*})(1 - x_i)}{n}, \quad \text{i=1,2, } \cdots \text{ ,n} \tag{3}$$

When using this method, the result of inputting signal (a) is the same as the regular ART1 network. When inputting signal (b) ($X=X_2$), its matching level is still the highest at the first node (higher than other unused nodes). According to formula (3), its similarity value ρ_2 is: ρ_2＝34/35=0.971, indicating that (b) is different from (a).

The similarity judgment of this improvement is carried out according to the input pattern and the existent matching mode. Since at the time of network initialization t_{ij} is 1, when the unused nodes are used for the first time, its outstar vector is directly adjusted as the input vector, i.e. $t_{ij^*} = x_i$, $i = 1,2,\cdots,n$.

3.2 Inadequacy and improvement of the regular outstar weight vector's judgment

The adjustment of the outstar weight vector of the regular ART1 neural network is indicated as formula (1). It is the "AND" of the outstar weight t_{ij^*} and the input x_i. This tends to give rise to gradual pattern drifting at when identifying the network pattern [6]. Namely, for a certain winner node, its outstar weight vector becomes increasingly away from the original pattern vector with the gradual change of the input pattern. On this occasion, the input pattern matches in similarity with a certain node, and there is very small change in the input vector and the original pattern. The weight of this node will be changed every time, resulting in unstable representation of pattern by each node. Sometimes they may drift into other patterns, causing mistakes for the network in pattern classification.

The obvious feature of the ART1 neural network is that it has only two statuses: 1 and 0.It is very difficult to fundamentally eliminate the gradual pattern drifting and the only practical thing is to reduce the occurrence of this. The regular ART1 neural network will adjust the weight of the outstar weight vector at this node once the input pattern matches with the existent nodes, no matter it is in conformity with our expectation or against our expectation. Faced with this situation, we made changes in the time for adjusting the outstar weight vector, i.e. the weight of the outstar is no longer adjusted once the input pattern matches in similarity with a certain node. The method we adopted is when there are quite a few input patterns at a certain node to match the similarity, store the information corresponding to these groups of patterns. When the information reaches a certain amount, integrate the knowledge of these outstar vectors and made adjustments to them. For example, take three groups of vectors to conduct the adjustment of comprehensive knowledge, i.e. adopt the 2/3 principles. Compare the three numbers of the corresponding position of each group of vectors and take the one which has the largest number of similarities as the adjusted value of the outstar weight vector. When all the three numbers are 1 or 0, the adjusted number is 1 or 0. When two are 1 and one is 0, the adjusted value is 1. When two are 0 and one is 1, the

adjusted value is 0.

3.3 Further Improvements

In order to further improve the identifying performance of the ART1 neural network, here the discussion shall focus on the identifying effect of the ART1 network by integrating the method of improving similarity and the computing method of adjusting the outstar weight vector. In the above-mentioned example of identifying English letters, the threshold value of similarity ρ_0 is 0.9. In fig. 1., according to formula (2), the similarity of (e) and (a) is $\rho_1 = 14/14 = 1 > \rho_0$. According to formula (3), the similarity is $\rho_2 = 31/35 = 0.886 < \rho_0$. These two patterns can be distinguished from each other through the improved ART1 similarity judgment. However, if the signal is input according to the sequence of (a) (b) (c) (d) and adjust the instar and outstar weight after matching the similarity each time, then input signal (e), and the system will transform signal (e) into the standard category E. This is because in the adjustment of weights, the outstar weight vector of expressing the pattern features moves away toward the unexpected directions. When this movement reaches a certain degree, it will drift to other patterns. If the method of improved similarity judgment is used, the drifting problem will also exist.

(1) The Adjusted Result of Weight of the Regular Outstar Weight Vector

As Table 1 shows, input the signal in the sequence of (a) (b) (c) (d) (e). The similarity threshold value $\rho 0$ is 0.9. When inputting signal (a) at the first time, the similarity matching with the "empty" pattern is 1. When inputting signal (b) at the second time, the similarity ρ_1 obtained through the regular method is bound to be 1, matching with the first node, because in the comparison of signal (b) and signal (a), just one becomes 0 at the position of (a), namely, X_1 is 1. At this time, adjust the weight of the outstar weight vector of the first node. Likewise, when inputting signal (c) at the third time, the similarity ρ1 obtained through the regular method is not 1 but 0.941 instead. Then Adjust the weight of the outstar weight vector. When inputting signal (d), the obtained regular similarity ρ1 is 0.880, lower than the similarity threshold value ρ0, thus being classifying into the other category. When inputting signal (e), it is classified into the standard category E and the similarity ρ_1 is 1, enjoying complete similarity. It is obvious that the result does not conform to the actual situation.

If the computing method of improving similarity is used, (d) can be classified into category E, but the pattern drifting still exists and signal (e) is still classified as the standard category E.

Table 1. A comparison between the adjusted result of weight of the regular outstar weight vector and the adjusted result of weight of the improved outstar weight vector

Similarity	The Adjusted Result of Weight of the Regular Outstar Weight Vector		The Adjusted Result of Weight of the Improved Outstar Weight Vector	
	ρ_1	ρ_2	ρ_1	ρ_2
(a)	1.000	1.000	1.000	1.000
(b)	17/17=1.000	34/35=0.971	17/17=1.000	34/35=0.971
(c)	16/17=0.941	33/35=0.943	17/17=1.000	34/35=0.971
(d)	15/17=0.880	32/35=0.914	17/17=1.000	34/35=0.971
(e)	14/14=1.000	35/35=1.000	14/14=1.000	31/35=0.886

2) The Adjusted Result of Weight of the Improved Outstar Weight Vector

The method of improving weight of the outstar weight vector adjustment is the principle of minority subordinate to majority. In the adjustment of weight of the regular outstar weight vector, t_{ij^*} can only be changed from 1 to 0 or unchanged, while after the adoption of the principle of minority subordinate to majority, t_{ij^*} can be changed from 1 to 0 or from 0 to 1. On the basis of adopting the computing method of the improved similarity, the drifting problem of the above-mentioned pattern is well resolved through the method of adjusting the weight of the improved outstar weight vector. As indicated by Table 1., the obtained value of ρ_2 has obviously classified signal (b) (c) (d) into Standard Category E, while signal (e) is classified as another category. For a certain established pattern, the method of adjusting the weight of the improved outstar weight vector is to learn the network with choice in the future learning process. Only those important information which occurs frequently will be learned by the network and those often-changing uncertain information will not will learned by the network. This method does not eliminate the gradual pattern drifting in essence, but it greatly reduces the occurrence of the drifting problem. Take the 2/3 principle as an example, although the pattern drifting will not occur when inputting the above-mentioned example, the gradual pattern drifting will occur only when each group of vector is input in sequence and input consecutively at least twice. If this situation occurs in the actual application, we can adopt the more strict principles, such as the 3/5 principles, the 4/7 principles, etc.

4. Conclusion

The improved ART1 neural network proposed by this paper is problem arises from the actual application. This paper puts forward the judging method for the exclusive nor similarity of two vectors, thus finding a good solution to the problem arising from the different inputting sequence. However, only making improvements to the similarity does not solve the problem of pattern drifting existed in the network. On the basis of this situation, this paper proposes the principle of minority subordinate to majority to improve the method of adjusting the weight. Although the problem of pattern drifting is not completely solved, the improvement method proves to be very effective, greatly reducing the occurrence of this drifting problem.

5. References

[1] G.A. Carpenter, S. Grossberg, "A massively parallel architecture for a self-organizing neural pattern recognition machine", Trans. IEEE on Computer Vision, Graphics and Image Processing, 1987, 37(1): 54-115.

[2] S. Grossberg, Competive learning, from interactive

activation to adaptive resonance. Cognitive Science, 1987, 11(1): 23-63.
[3] C.Y. Tong, J.Y. Shi, "A pattern classification-oriented modified ART1 neural network", Chinese J.Computer, 1995, 18(9): 671-677.
[4] D.Q. Zhu, "Theory and practice of fault diagnosis to electronic device", Publish house of electronics industry, 2004, 145-150.
[5] L.Q. Han, "Theory、devise and application to artificial neural networks", Publish house of chemical industry. 2002, 81-87.
[6] Laura I, Burke, "Clustering Characterization of Adaptive Resonance", Neural Networks, 1991, 16(4): 485-491.

Research on Forecast Model of Heavy Metal in Sludge based on BP Network

Ying Sun
College of Environmental and Chemical Engineering, Shanghai University
Shanghai, 200072, China
Email: sunyingsh@sina.com

Lin Yang
College of Economics and Management , Tongji University
Shanghai, 200092, China
Email: folinyang@126.com

ABSTRACT

Content forecast of heavy metal is emphasized during sludge utilization in agriculture. General forecast algorithms are not competent for the forecast since the relationship between content of heavy metal in sludge and its influences are multidimensional and nonlinear. With influences of both sequence and consequence upon content of heavy metal in sludge considered, a forecast model based on BP network was constructed and trained with sludge data sampled from a municipal wastewater treatment plant. The comparison of forecast performance was made between the constructed and linearity regression algorithms. Result shows that the constructed algorithm has better optimization searching mechanism and is more precise in forecast.

Keywords: Heavy Metal, Sludge, Forecast, BP Network.

1. INTRODUCTION

Sludge utilization in agriculture is one of sludge disposal methods. Yet, the content of heavy metal in sludge is so high that it is hundreds times of that in soil. When the sludge is utilized in agriculture, the heavy metal infiltrated into soil is hard to be cleaned away. It may bring harm to human's body by infiltrating into groundwater or via plant absorption.

The content of heavy metal in municipal domestic sludge changes with influences such as treatment methods, temperature and rain volume, also is influenced by seasonal changes, population and standard of living. Among these influences, some influences are certain, some uncertain. Meanwhile, the relationship between content of heavy metal in sludge and influences is multidimensional and nonlinear. However, current researches focus mainly on the measurement and analysis of existing heavy metal in sludge, fewer on its forecast or prediction. With the emphasis of environment Protection and rapid spread of sludge utilization in agriculture, mankind cares about not only the current status of the content, but also the rule in it and its trend, so as to make better decision on what the content or the heavy metal is to be and on better treatment strategy.

Heavy metals like Zn, Cu and Ni whose content exceeds utilization criteria of sludge in agriculture are chosen as research targets. With temperature selected as influence, the content of heavy metal in sludge is forecasted with forecast model based on BP network.

2. APPLICATION OF BP NETWORK ON FORECAST

As a research field developed rapidly in the mid 80s, Artificial Neural Networks has influenced a lot upon computer science and artificial intelligence.

Back Propagation network, BP in short, is the best algorithm used as forecasting among neural networks [1]. The main reason lies in that the nonlinear and the collateral feature keeps it successful to do forecast issues, simple or complex. The nonlinear guarantees ample flexibility to describe the real world while the collateral, by storing distributed information with interlink index, enables the network higher error handling capacity and higher robust to resolve the yawp disturb and partial loss of input, which are normal existing puzzles among pattern identification.

The forecast of heavy metal's content in sludge is to find out the mapping relationship between heavy metal content and influential factors. BP network is verified to approach any mapping program with given precision, which makes it superior to do forecast than traditional methods [2].

The strongpoint of BP network in forecast has been verified with practices. BP network has been used successfully to do forecast in different industries, such as manufacturing and business. Reference [3] resolved the forecast of customer behavior in power industry by building up a forecast model based on BP network and generating cluster and forecast rules based on fuzzy control. Reference [4, 5, 6, 7, 8, 9] made a research on bankruptcy forecast with BP network. There are some researches on environment forecast with BP network [10, 11]. Yet the application of BP network on forecast of heavy metal content in sludge is not found yet.

3. HEAVY METAL FORECAST BASED ON BP NETWORK

3.1 The Principle of BP Network Forecast Model

Fig.1 describes a topology structure of BP network. From left to right, they are named as input layer, hidden layer and output layer.

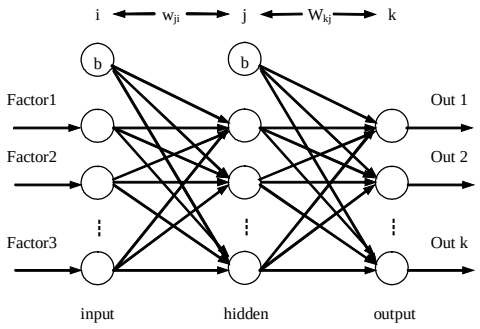


Fig. 1. A topology structure of tree layer BP network

The training process includes forward propagation and back propagation. During forward propagation, information input will be transferred into output layer after being processed in hidden layer. If an expected output can not be approached, both the output and the deviation are turned back to back propagation process via original route. By modifying the weight of each nerve cell, the deviation tends to become smaller. Then turn back to forward propagation and back and forth till the given deviation is reached.

Generally for each combined (x_i, y_i), the network output $\{o_k^t\}$ is not matched with expected output $\{y_k^t\}$ with the deviation as follows:

$$E_t = \sum_{k=1}^{K} (y_k^t - o_k^t)^2$$

In order to decrease the deviation and step the network convergence, weight index should be improved on the opposite direction of grads changed.

$$w_{ji}(n+1) = w_{ji}(n) + \eta\sigma_j o_i + \alpha[w_{ji}(n) - w_{ji}(n-1)]$$

Where:

$w_{ji}(n+1)$ stands for the trained weight of the No n+1 step;

α stands for the ratio constant to modify the changes, named as momentum index also;

η stands for training rate. Both α and η are training index which values between 0 and 1;

δ_j stands for deviation grads, calculating on the opposite direction of network convergence.

The average deviation is calculated as following if the total num of training sample is P:

$$\overline{E} = \frac{1}{P \times K} \sum E_t = \frac{1}{P \times K} \sum_t \sum_k (y_k^t - o_k^t)^2$$

The training process will be finished if $\overline{E} \le errorlevel$ is reached and a forecast model based on BP network is achieved. $errorlevel$ is a permitted error level. The forecast model is expressed as weights between layers.

3.2 Confirm the Influential Factors

Considering the association between sludge and influences such as season, standard of living and so on, the forecast model to be built up combines together the influences of sequence and consequence. In order to consider the influence of season, the content data of former 12 months is used to forecast that of the 13th month. For the reason that influences like population and rain volume keep relatively unchanged during a year, only rain volume is selected as key consequence factor in the model.

3.3 Build up Forecast Model

1). Confirm the num of input node, output node, hidden node and sample

The total num of sample is suggested to be 12, which means a data collection of a whole year is needed to organize the sample. The num of input node is set to be 9, among which the former 8 is attributed to sequence influence and the remaining 1 to consequence one. The num of output node is eight. Here, eight stands for num of heavy metals to be forecasted.

The hidden node can be chosen after the deviation trend of BP network is get. The process contains:

- The training sample is organized with p_i, p_j and K.
- The initial hidden node is set as $p_k = p_i + \text{int}(sqrt(p_j))$, where initial weight is generated with time generator.
- The relationship between deviation and hidden node is graphed after BP network training.
- Set up $p_k = p_k + 1$, and then repeat the 3rd step. The final hidden node is set up according to practice.
- At last, the node mapped to lowest point in the graph should be selected as optimized hidden node.

In Fig.2, this case selects 20 as optimized hidden node after training.

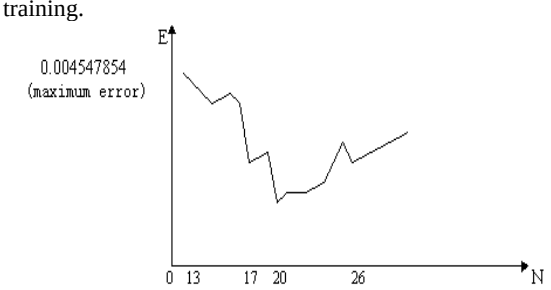


Fig. 2. The optimized hidden node

The advantage of optimizing hidden node is to approach an optimized BP network and to ensure the stability of optimized weight.

2). Set up training rate η and momentum index α

Set up $\eta = 0.15$ and $\alpha = 0.075$, which is also suitable to most of issues using ANN network.

3). Error level

The error lever named $errorlevel$ in the BP network is set up as 0.0002.

3.4 Collect Samples

Municipal domestic sludge sampled from Qu Yang wastewater treatment plant is used as sample in the experiment. Lasting for one and half a year from January 2004 to June 2005, each sludge sample is collected per the middle ten days of a month, among which one year's data is organized for training sample and others for testing sample and forecasting sample. The content and the corresponding influences are measured separately. The content is calculated as average of tree parallel values.

Table 1. Contents of heavy metals in sludge changed with corresponding influences (mg/kg dry sludge)

Month	year	Zn	Cu	Ni	Pb	Hg	Cd	Cr	As	Temperature (℃)
Jan	2004	1489	172	44.3	27.5	2.69	1.03	39.2	10.8	3.7
Feb	2004	1524	163	40.6	21.3	2.39	1.08	28.9	17.6	4.2
Mar	2004	1868	185	45.6	26.2	3.04	0.98	31.6	15.2	7.4
Apr	2004	2378	215	65	13.4	3.33	1.29	71.7	27.1	14.6
May	2004	2022	176	176	19.7	5.15	0.93	127	28.7	19.4
Jun	2004	2093	197	134	27.8	3.96	1.13	149	36.5	24.4
Jul	2004	2280	268	96.0	31.8	2.82	1.36	57.5	23.3	29
Aug	2004	2759	203	109	39.0	4.00	1.22	131	40.8	28.1
Sep	2004	2058	168	159	25.7	3.12	1.02	54.8	35.9	23.3
Oct	2004	1740	181	51.4	20.2	2.96	0.97	37.5	16.1	17.3
Nov	2004	1477	165	43.5	32.6	2.43	0.76	36.9	15.7	11.9
Dec	2004	1433	169	42.8	14.7	2.57	1.17	27.3	5.50	6.5
Jan	2005	1987	176	59.2	8.98	3.28	1.25	65.2	23.5	4
Feb	2005	1786	196	68.3	16.2	10.6	0.98	62.9	20.5	4.5
Mar	2005	1887	195	45.3	10.6	3.54	1.38	34.6	15.9	7
Apr	2005	2237	224	84	34.2	3.73	1.74	64.7	27.3	14.7
May	2005	2043	193	186	10.7	5.40	1.23	142	30.7	18.6
Jun	2005	2153	217	154	32.8	4.36	1.23	162	39.5	24

4. Results and Analysis

With the last sample in table 1 adopted as forecast sample, the content of next month (June, 2005) is forecasted. See the results and the analysis in table 2.

Meanwhile, comparison is made between BP network forecast model and linearity regression. Limited by algorithm itself, linearity regression does forecast on the data of the thirteenth month using the content data of last twelve months, without adopting other consequence influences.

Table 2. Results and comparison of two forecast models

	Zn	Cu	Ni	Pb	Hg	Cd	Cr	As	Absolute deviation
Actual content of Jun, 2005	2153	217	154	32.8	4.36	1.23	162	39.5	
BP forecast model	2064	188	133	34.3	3.39	1.14	172	32.5	0.02
Linearity regression	1968	185	126	38.2	3.04	0.98	137	43	0.167

With sample data organized as training sample, the training deviation calculated after the forecast model based on BP network is trained for 1000 iterative times is 0.000039, the checkout deviation 0.000463. Input the forecast sample into the trained forecast model.

Adopting the least squares error of deviation as target, the linearity regression is trained with the same sample data. The training deviation calculated is 0.00373, the checkout deviation 0.00328. Input the forecast sample into the trained forecast model, too.

See the comparison in table 2 to find out that:

(1) Both the training deviation and the checkout deviation of forecast model based on BP network are far less than those of linearity regression, which shows that the former has better optimization searching mechanism to be adapt to the research question better.

(2) The absolute deviation of BP network forecast model is 2%, which is far less than that of linearity regression, for the corresponding one is 16.7%. This shows that the forecast of former forecast model is more precise than the later.

5. REFERENCES

[1] Y. Dong, X.L. Liu, “generalized BP algorithms and research on its error handling and panlogistic capability”, Control and Decision, No.3, 1998, pp.120-124.

[2] K. Robert, “Jure Zupan, Application of a Feed-Forward Artificial Neural Network as a Mapping Device”, J.Chem.Inf.Comput.Sci, Vol.37, 1997, pp.985-989.

[3] Q. Guo, P. Zhang, “Neural networks for classification: a survey”, IEEE Transactions on system, man, and cybernetics-Part c: application and reviews, 30(4), November 2000, pp. 451-458.

[4] M. Sforna ,”Data mining in a power company customer database”, Electric Power Systems Research Vol.55, 2000, pp. 201–209.

[5] E.I. Altman,G. Marco, et al., “Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience)”, J.Bank Finnace, Vol.18, 1994, pp. 505-529.

[6] R. Clacher, P.K. Coats, et al., “A neural network of classifying the finacial health of a firm”, Eur.J.Oper. Res., Vol.85, 1995, pp.53-65.

[7] M. Leshno , Y. Spector, “Neural network prediction analysis: the bankruptcy case, Neurocomput”, Vol.10, 1996, pp.125-147.

[8] K.Y. Tam , M.Y. Kiang, “Managerial application of neural networks: the case of bank failure predictions”, Manage, SCI, 138(7), 1992, pp.926-947.

[9] R.L. Wilson , R. Sharda, “bankruptcy prediction using neural networks, decision support system”, Vol.11, 1994, pp.545-557.

[10] K. Cunevt, “A new approach to mathematical water quality modeling in Reservoirs : Neural Networks[J] , Proceedings of the 3rd International Conference on Reservoir Limnology and water quality”, Internaional Review of Hydrobiology, Vol.83 , 1998 , Special Issue : pp.689-696.

[11] F. Recknagel, “Modeling and prediction of phyto and zooplankton dynamics in Lake Kasumigaura by artificial neural networks[J] , Lakes & Reservoirs” : Research and Management ,3 (2), 1998, pp.123-133.

Neural Network Control of Underwater Vehicles Based On Robust Learning Algorithm

Yushan Sun, Xiao Liang, Lei Wan, Shaoji Fang
College of Shipbuilding Engineering, Harbin Engineering University
Harbin, Heilongjiang Province 150001/ China
Email: sunys2000@sohu.com

ABSTRACT

Liang Peng etc. successfully applied neural network technologies to motion control of underwater robots, but low response speed and sensitivity to external noises limit the development in practice. In order to improve the performance of neural network controller, a robust learning algorithm is proposed. Firstly, the stable robust learning algorithm is presented, which is based on variable structure control theory. Secondly, the global stability conditions are explained in detail. Finally, in order to evaluate the performance, the control system based on the proposed algorithms is applied to the simulation platform of General Detection Remotely Operated Vehicle (GDROV).The results show that it not only possesses the good robustness to changing of learning-ratio and external noises but also can keep learning of neural network fast and stable, which is feasible to be applied to real time control of underwater vehicle.

Keywords: Neural Network, Robust Learning Algorithm, Variable Structure, Learning-ratio, General Detection Remotely Operated Vehicle (GDROV).

1. INTRODUCTION

Owing to the nonlinearity of underwater vehicles, many uncertainties of the system and the big noise of sensors in the harbor, the precise control of underwater in some intelligent options has been puzzling us in the application of underwater vehicles. Generally there are three control methods, i.e. PID control, fuzzy control and neural control [1]. PID controllers are abroad applied because of the simplicity of their structure [2, 3]. However, the major drawback of PID controllers is difficulty in parameter design to multi-input and multi-output (MIMO) system, especially for strong coupling system. Fuzzy controllers are also simple to be designed for underwater vehicles, but their disadvantages are obvious: deficiency in knowledge acquisition and self-learning. Because of the ability of learning, non-linear mapping and self-adapting in recent years, neural network technologies have been applied in the field of underwater vehicles.

In 1995, Liang Peng etc. applied neural network technologies to motion control of underwater vehicles and have good results [4], but low response speed and sensitization to external noises limit the development in practice.

To solve the problem mentioned above, a novel robust learning algorithm is proposed in neutral network control, which relates to variable structure control theory [5,6,7] and EBP algorithm[8,9]. The algorithm is presented in detail in section II. And it can improve the robustness to external disturbance and changing of learning ratio greatly. In section III the global stability conditions are deduced and analyzed. Section IV introduces the open-frame underwater robot GDROV some simulation demonstrations, which shows the feasibility and superiority.

* Supported by the National High Technology and Development Program Foundation of China (No. 2002AA420090).

2. ROBUST LEARNING ALGORITHM

In EBP algorithm, regulation of modifying net weights in continuous time domain can be described as such a dynamic model:

$$\dot{\Delta\phi} = (\eta_\varphi N_\phi - \Delta\phi)/T_s \tag{1}$$

Where T_s is sampling period, $N_\phi = \sum_{j=1}^{\text{outputs}} e_j(\partial f_{j(\phi,x)}/\partial\phi)$ is variation of net weights based on gradient method, η_ϕ is scale coefficient.

Difference format of Eq.(1) is:

$$(\Delta\phi_{(k+1)} - \Delta\phi_{(k)})/T_s = (\eta_\phi N_{\phi(k)} - \Delta\phi_{(k)})/T_s \tag{2}$$

$$\Delta\phi_{(k+1)} = \eta_\phi N_{\phi(k)} \tag{3}$$

Basing on approach law in variable structure control theory and EBP algorithm described upward, parameter ballasting part of robust learning algorithm can be deduced as follows.

Where slip-form switching functions is given by:

$$s_\phi = \Delta\phi \tag{4}$$

We select exponent approach law,

$$\dot{s} = -[Q_\phi \tanh(s_\phi/\varepsilon) + K_\phi s_\phi]/T_s \tag{5}$$

Where Q_ϕ and K_ϕ are scale coefficients, ε is thickness of boundary layer.

On the assumption that (5) and (1) are equal, so

$$\Delta\phi = \eta_\phi N_\phi + A_\phi \tag{6}$$

$$A_\phi = Q_\phi \tanh(\Delta\phi/\varepsilon) + K_\phi\Delta\phi \tag{7}$$

If η_ϕ is selected properly, $\Delta\phi$ will run according to exponent approach law in Eq.(5) till approaching 0 at utmost, namely, adjustable parameters of the system will approach certain steady-state values. However, it does not always make object function minimum. How to select η_ϕ and learning algorithm that has complete global stability are going to be discussed then.

To the system described upward, we select positive definite Liapunov function as follows:

$$V_\phi = s_\phi^2/2 = (\Delta\phi)^2/2 \tag{8}$$

$$\dot{V}_\phi = \Delta\phi \cdot \dot{\Delta\phi} \tag{9}$$

Substituting Eq. (1) and (6) into (9), qualification of Liapunov stability can be obtained.

$$(\eta_\phi + A_\phi / N_\phi)(\eta_\phi - \Delta\phi / N_\phi) < 0 \tag{10}$$

Because A_ϕ and $\Delta\phi$ are of the same sign, two roots in (10) are of the opposite sign. η_ϕ must satisfy the condition described in Eq.(11):

$$0 < \eta_\phi < \min\{|\frac{1}{N_\phi}\Delta\phi|, |-\frac{1}{N_\phi}A_\phi|\} \tag{11}$$

Namely,

$$\eta_\phi = \beta \min\{|\frac{1}{N_\phi}\Delta\phi|, |-\frac{1}{N_\phi}A_\phi|\}, \quad 0 < \beta < 1 \tag{12}$$

Substituting Eq.(12) into Eq.(6), it can be deduced:

$$\Delta\phi_{VSS} = \beta \min\{|\Delta\phi|, |A_\phi|\}\mathrm{sgn}(N_\phi) + A_\phi \tag{13}$$

As can be seen, parameter-ballasting part upward can only assure the minimization of Liapunov function. To assure the minimization of object function, there should be regulation of modifying net weights in EBP algorithm likewise. It can be described as:

$$\Delta\phi_{EBP} = \xi_\phi N_\phi \tag{14}$$

Where ξ_ϕ is learning ratio.

Finally, robust learning algorithm is described as follows to assure both stability of parameters and the minimization of object function.

$$\Delta\phi = (\alpha_1 \Delta\phi_{VSS} + \alpha_2 \Delta\phi_{EBP})/(\alpha_1 + \alpha_2) \tag{15}$$

Where $\Delta\phi_{VSS}$ is parameter ballasting part, $\Delta\phi_{EBP}$ presents traditional EBP algorithm part. α_1 and α_2 are scale coefficients separately, they determine the balance between $\Delta\phi_{VSS}$ and $\Delta\phi_{EBP}$.

3. GLOBAL STABILOITY CONDITIONS

The stability of neural network is important enough, so analysis should be made on stability conditions of neural network. We select Liapunov function as follows,

$$V_\phi = [(\Delta\phi)^2 + \gamma_\phi (N_\phi)^2]/2 \tag{16}$$

γ_ϕ is a positive constant which is determined by experience.

The derivative of Eq.(16) is

$$\dot{V}_\phi = \dot{\Delta\phi}\Delta\phi + \gamma_\phi \dot{N}_\phi N_\phi \tag{17}$$

Then, we can get,

$$\dot{V}_\phi = [-(\Delta\phi)^2 + \min\{|\Delta\phi|, |A_\phi|\}\mathrm{sgn}(N_\phi)\Delta\phi \cdot \alpha_1\beta/(\alpha_1+\alpha_2) + \alpha_1/(\alpha_1+\alpha_2)A_\phi\Delta\phi + \alpha_2\xi_\phi N_\phi\Delta\phi/(\alpha_1+\alpha_2)]/T_s + \gamma_\phi \dot{N}_\phi N_\phi \tag{18}$$

There are two express forms because of the existence of minimization operator.

If $|\Delta\phi| < |A_\phi|$, $|\Delta\phi| = \Delta\phi\,\mathrm{sgn}(\Delta\phi)$, it can be also described as:

$$\dot{V}_\phi = [-(\Delta\phi)^2 + \mathrm{sgn}(\Delta\phi)\mathrm{sgn}(N_\phi)(\Delta\phi)^2 \cdot \alpha_1\beta/(\alpha_1+\alpha_2) + \alpha_1/(\alpha_1+\alpha_2)A_\phi\Delta\phi + \alpha_2\xi_\phi N_\phi\Delta\phi/(\alpha_1+\alpha_2)]/T_s + \gamma_\phi \dot{N}_\phi N_\phi \tag{19}$$

Because $A_\phi = Q_\phi \tanh(\Delta\phi/\varepsilon) + K_\phi\Delta\phi < (Q_\phi + K_\phi)\Delta\phi$, and $\mathrm{sgn}(x_1)\,\mathrm{sgn}(x_2) \le 1$, it must be satisfied,

$$\dot{V}_\phi < [-(\Delta\phi)^2 + \mathrm{sgn}(\Delta\phi)\mathrm{sgn}(N_\phi)(\Delta\phi)^2 \cdot \alpha_1\beta/(\alpha_1+\alpha_2) + \alpha_1/(\alpha_1+\alpha_2)Q_\phi \tanh(\Delta\phi/\varepsilon)\Delta\phi + \alpha_1/(\alpha_1+\alpha_2)K_\phi(\Delta\phi)^2 + \alpha_2\xi_\phi N_\phi\Delta\phi/(\alpha_1+\alpha_2)]/T_s + \gamma_\phi \dot{N}_\phi N_\phi \tag{20}$$

If $|\Delta\phi| > |A_\phi|$,

$$\dot{V}_\phi = [-(\Delta\phi)^2 + \mathrm{sgn}(N_\phi)|A_\phi|\Delta\phi \cdot \alpha_1\beta/(\alpha_1+\alpha_2) + \alpha_1/(\alpha_1+\alpha_2)A_\phi\Delta\phi + \alpha_2\xi_\phi N_\phi\Delta\phi/(\alpha_1+\alpha_2)]/T_s + \gamma_\phi \dot{N}_\phi N_\phi = [-(\Delta\phi)^2 + (\beta\,\mathrm{sgn}(A_\phi)\mathrm{sgn}(N_\phi) + 1)A_\phi\Delta\phi \cdot \alpha_1/(\alpha_1+\alpha_2) + \alpha_2\xi_\phi N_\phi\Delta\phi/(\alpha_1+\alpha_2)]/T_s + \gamma_\phi \dot{N}_\phi N \tag{21}$$

Because following conditions are satisfied,

a) $A_\phi\Delta\phi < (Q_\phi + K_\phi)\Delta\phi^2$, $A_\phi\Delta\phi > 0$

b) $\min(\beta\,\mathrm{sgn}(A_\phi)\mathrm{sgn}(N_\phi) + 1) = 1 - \beta$

$\max(\beta\,\mathrm{sgn}(A_\phi)\mathrm{sgn}(N_\phi) + 1) = 1 + \beta$; $0 < \beta < 1$

Eq. (21) can be also described as:

$$\dot{V}_\phi < [-(\Delta\phi)^2 + 2(Q_\phi + K_\phi + \beta)(\Delta\phi)^2 \cdot \alpha_1/(\alpha_1+\alpha_2) + \alpha_2\xi_\phi N_\phi\Delta\phi/(\alpha_1+\alpha_2)]/T_s + \gamma_\phi \dot{N}_\phi N \tag{22}$$

If the value of expression is minus on the right side of Eq.(22), it is satisfied automatically that the value of expression is minus on the right side of Eq.(20).Therefore, we can get

$$\dot{V}_\phi < -C_\phi(\Delta\phi)^2/T_s + \alpha_2\xi_\phi N_\phi\Delta\phi/[(\alpha_1+\alpha_2)T_s] + \gamma_\phi \dot{N}_\phi N_\phi \tag{23}$$

Where

$$C_\phi = 1 - 2(Q_\phi + K_\phi + \beta)\cdot\alpha_1/(\alpha_1+\alpha_2) \tag{24}$$

As can be seen, global stability of robust learning algorithm can be achieved if Eq. (23) is negative definite γ_ϕ is given by:

$$\gamma_\phi = \sigma_\phi^{\,2} / \sup_t |\dot{N}_\phi N_\phi| \tag{25}$$

Where $\sigma_\phi^{\,2}$ is nonzero minimum of $(\Delta\phi)^2$ in a learning course. Therefore,

$$\dot{V}_\phi < (T_s - C_\phi)(\Delta\phi)^2/T_s + \alpha_2\xi_\phi |N_\phi||\Delta\phi|/[(\alpha_1+\alpha_2)T_s] \tag{26}$$

$$\xi_\phi < (\alpha_1+\alpha_2)(C_\phi - T_s)|\Delta\phi|/(\alpha_2|N_\phi|) \tag{27}$$

Restrictive condition of ξ_ϕ is given upward, it keeps the first derivative of Liapunov function negative definite, which keeps global stability of robust learning algorithm.

4. APPLICATION TO GDROV

The robust learning algorithm mentioned above has been applied to the computer simulation system of GDROV, the simulation results demonstrate the feasibility and superiority.

4.1 General Detection Remotely Operated Vehicle

GDROV is a kind of open-frame remotely operated underwater vehicle designed for the inspection of cracks, crevices and other potential problems of the dams, which

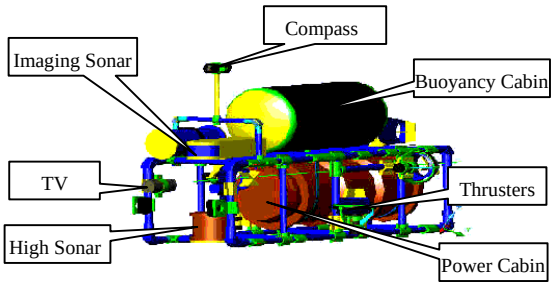


Fig. 1. Overall structure of GDROV

is developed by Harbin Engineering University. Overall structure and System configurations of GDROV are can be seen from Fig. 1. and Fig. 2. respectively. According to detection requirement of the dams GDROV is provided high frequency sonar or low frequency sonar, which results in the difficulty in acquiring the precise model of it. So, it is essential that the motion control system of GDROV be provided with the ability of self-learning and self-adaptation.

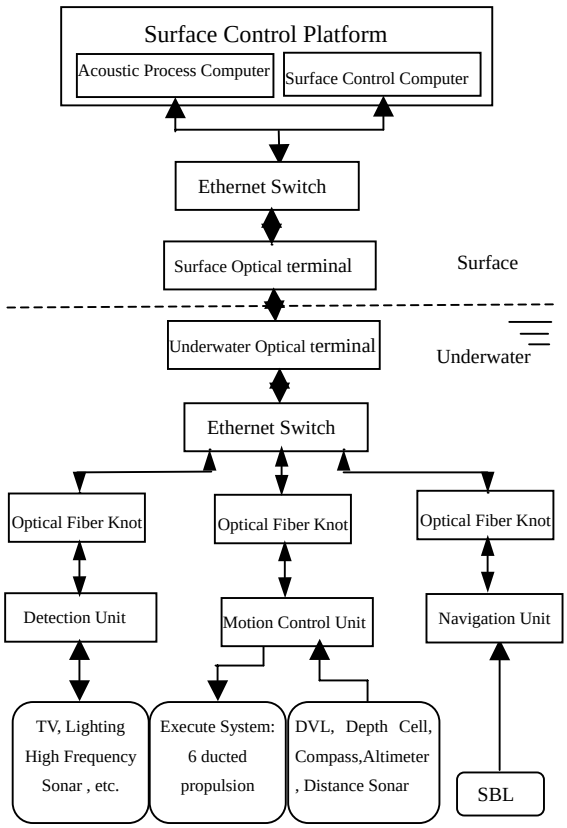


Fig. 2. System configuration of GDROV

4.2 Model of GDROV

GDROV has six thrusters: two longitudinal main thrusters, two vertical thrusters, and two lateral thrusters. Therefore, the rolling motion and the trim motion of GDROV are always quite small with these manipulative facilities. We ignore them and construct a four DOF hydrodynamic model (Surge, sway, heave and yaw actively controlled with thrusters, pitch and roll are left uncontrolled due to the passive stability of GDROV.).

We transpose the derivative term of the four DOF equation of underwater vehicles [10] to the left of the equation, and the rest terms to the right. After arrangement, the equation in vectors is expressed as follows:

$$\mathbf{E}\dot{\mathbf{X}} = \mathbf{F}_{vis} + \mathbf{F}_{\mathbf{t}} \tag{28}$$

Where

$$\mathbf{E} = \begin{bmatrix} m - X_{\dot{u}} & 0 & 0 & 0 \\ 0 & m - Y_{\dot{v}} & mx_G - Y_{\dot{r}} & 0 \\ 0 & 0 & m - Z_{\dot{w}} & 0 \\ 0 & mx_G - N_{\dot{v}} & 0 & I_z - N_{\dot{r}} \end{bmatrix} \tag{29}$$

Where $\mathbf{F}_{vis}$ is the non-inertial hydrodynamic force, which is expressed as:

$$\mathbf{F}_{vis} = [X_{vis} \;\; Y_{vis} \;\; Z_{vis} \;\; N_{vis}]^T \tag{30}$$

$$\mathbf{X} = [u \quad v \quad w \quad r]^T \tag{31}$$

Where

$$\begin{aligned} X_{vis} &= -m \cdot [(-vr + wq) - x_G \cdot (q^2 + r^2) + z_G pr] \\ &+ [X_{qq} q^2 + X_{rr} r^2 + X_{rp} rp] + [X_{vr} vr + X_{wq} wq] \\ &+ [X_{uu} u^2 + X_{vv} v^2 + X_{ww} w^2] \end{aligned} \tag{32}$$

$$\begin{aligned} Y_{vis} &= -m \cdot [(-wp + ur) + z_G qr + x_G qp] + [Y_{p|p|} p|p| \\ &+ Y_{pq} pq + Y_{qr} qr] + [Y_{vq} vq + Y_{wp} wp + Y_{wr} wr] \\ &+ [Y_r ur + Y_p up + Y_{v|r|} \frac{v}{|v|} |(v^2 + w^2)^{1/2} |r|] + [Y_0 u^2 + Y_v uv \\ &+ Y_{v|v|} v |(v^2 + w^2)^{1/2}|] + Y_{vw} vw \end{aligned} \tag{33}$$

$$\begin{aligned} Z_{vis} &= -m \cdot [(-uq + vp) - z_G (p^2 + q^2) + x_G rp] + [Z_{pp} p^2 \\ &+ Z_{rr} r^2 + Z_{rp} rp] + [Z_{vr} vr + Z_{vp} vp] + [Z_q uq + Z_{w|q|} \frac{w}{|w|} \\ &\cdot |(v^2 + w^2)^{1/2} |q|] + [Z_0 u^2 + Z_w uw + Z_{w|w|} w |(v^2 + w^2)^{1/2}|] \\ &+ [Z_{|w|} u|w| + Z_{ww} |w(v^2 + w^2)^{1/2}|] + Z_{vv} v^2 \end{aligned} \tag{34}$$

$$\begin{aligned} N_{vis} &= -(I_y - I_x) pq - m \cdot [x_G \cdot (ur - pw)] + [N_{pq} pq \\ &+ N_{qr} qr + N_{r|r|} r|r|] + [N_{wr} wr + N_{wp} wp + N_{vq} vq] \\ &+ [N_p up + N_r ur + N_{|v|r} |(v^2 + w^2)^{1/2}| r] \\ &+ [N_0 u^2 + N_v uv + N_{v|v|} v |(v^2 + w^2)^{1/2}|] + N_{vw} vw \end{aligned} \tag{35}$$

We assume that the gravity of the underwater robot is equal to the buoyancy and ignore the external influence, so $\mathbf{F}_t$ are composed of the forces produced by thrusters, namely, $\mathbf{F}_t$ are the outputs of the controller.

$$\mathbf{F}_t = [X \quad Y \quad Z \quad N]^T \tag{36}$$

Make the inverse matrix of the coefficient matrix $\mathbf{E}$ to multiply two sides of (29) simultaneity, we can get

$$\dot{\mathbf{X}} = \mathbf{E}^{-1} (\mathbf{F}_{vis} + \mathbf{F}_t) \tag{37}$$

4.3 Simulation experiments

In this paper, only vertical control results are presented. Reference input is 4m, the velocity of current is 0, and the voltage of thrusters is restricted by 2.5V. The figures of the control results and variation of thrust forces in different conditions are shown in Fig. 3. to Fig. 5.

Fig. 3. shows the comparison of the control results

between small learning ratio (ξ_ϕ =0.05) and large learning ratio (ξ_ϕ =0.5) by adopting EBP algorithm. As can be seen, if learning speed of neural network runs up by enlarging learning ratio, there is a quicker response and lower overshoot. However, there is also high frequency concussion to the values of the thrust forces around the balance point of the control system, so it is very difficult to achieve the control tasks. Furthermore, it causes excessive abrasion in mechanically driven system of the underwater vehicle.

Fig. 4 shows the control results by adopting robust learning algorithm (ξ_ϕ =0.5, α_1/α_2 =2.0), and the selection of parameters meets the global stability condition. As can be seen, although learning ratio is large enough, it can keep control outputs smooth and slow up the high frequency concussion to the values of the thrust forces, which reduces the abrasion of the mechanically driven system greatly.

Fig. 5 shows the control results that there are strong external noises around. The noises added are Gaussian white noises, whose average value are zero, and the error square is 0.01. As can be seen, there is good robustness to the external disturbance, which is in favor of the implementation in practical work.

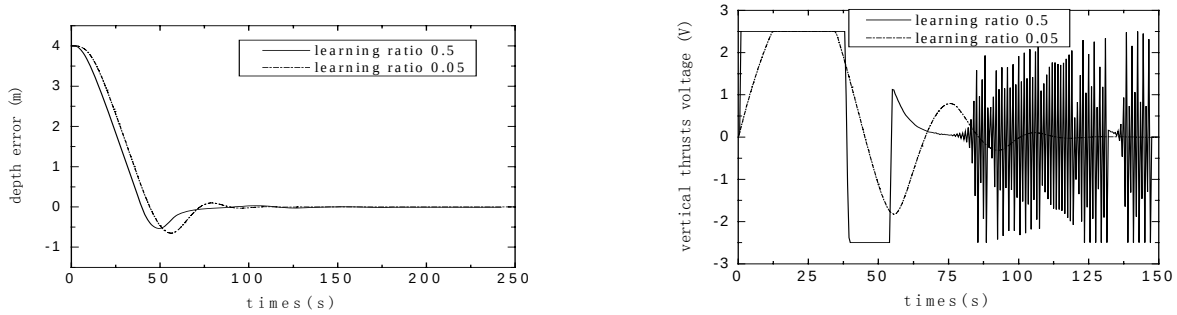


Fig. 3. Control performance by adopting EBP algorithm with different learning ratio

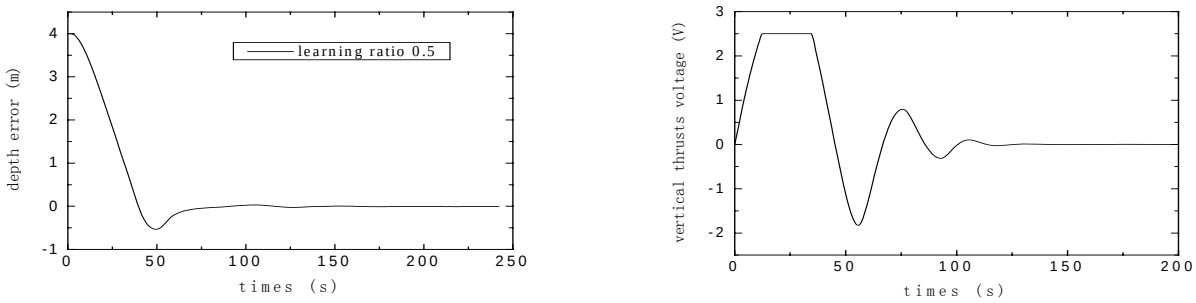
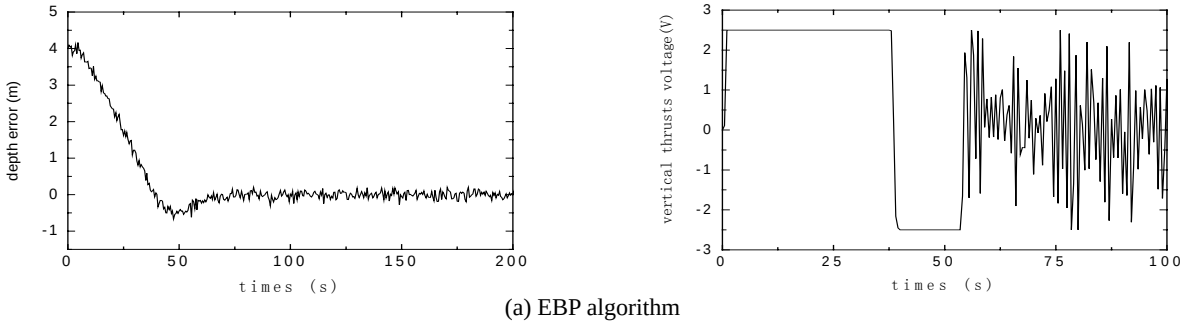
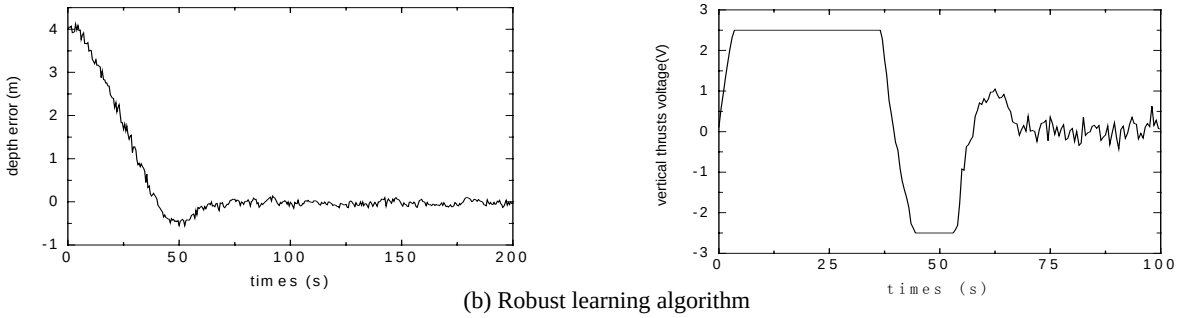


Fig. 4. Control performance by adopting robust learning algorithm with large learning ratio



(a) EBP algorithm



(b) Robust learning algorithm

Fig. 5. Comparison of control performance by adopting different learning algorithm in external noises

5. CONCLUSIONS

To make the control system of underwater vehicles fit to nonlinearity and unpredictable operating environment, a neural network control based on robust learning algorithm is presented, which possesses the ability of self-learning and self-adaptation. In papers variable structure control theory is also illuminated. In addition, the global stability conditions are deduced. And after control model of GDROV described, the simulation experiments based on the emulator of GDROV are introduced. The results show that it has on-line adapting ability for dynamic system of underwater vehicles. And it can be applied to controlling complicated, uncertain and nonlinear processes.

6. REFERENCES

[1] X.M. Liu, Y.R. Xu, "S control of automatic underwater vehicles," The Ocean Engineering, 19(3), 2001, 10-15.

[2] K.J. Astrom, T. Hagglund, et al, "Automatic tuning and adaptation for PID controllers-A Survey," Control Engineering Practice, 1993, 1(4): 699-714.

[3] H.X. Wu, S.P. Shen, "Basis of theory and applications on PID control," Control Engineering of China, 2003, 10(1): 37-42.

[4] L. Peng, Y.C. Lu, et al, "Neural network control of autonomous underwater vehicles," The Ocean Engineering, 1995, 13(2):38-46.

[5] F.R. Wang, Sliding mode variable structure control, Beijing: Mechanics Industry Press, 1998.

[6] L.R. Wang, J.C. Liu, et al, "Sliding mode control of an Autonomous Underwater Vehicle", Proceedings of 2002 International Conference on Machine Learning and Cybernetics, Beijing, P.R. China , 200,247-251.

[7] M.O. Efe, O. Kaynak, et al, "Sliding mode control of nonlinear systems using gaussian radial basis function neural networks", INNS-IEEE Int. Joint Conf. on Neural Networks, Washington, 2001.

[8] D.E. Rumelhart, G.E. Hinton, et al, "Learning internal representations by error propagation," Parallel Distributed Processing, Vol. 1, 1986,318–362.

[9] X.M. Liu, J.C. Liu, et al,"Motion control of underwater vehicle based least disturbance BP algorithm," Journal of Marine Science and Application, 2001,Vol. 1(1)16-20.

[10] Y. Li, J.C. Liu, et al, "Dynamics model of underwater robot motion control in 6 degrees of freedom," Journal of Harbin Institute of Technology, 2005, 12(4): 456-459.

Condition-Based Diagnosis of Thermal Facilities in Power Plant Based on Neural Network Fuzzy Evaluation

Daogang Peng[1,2], Hao Zhang[1,2], Ping Yang[1], Dajun Mao[1]
[1]Dept. of Information and Control Technology, Shanghai University of Electric Power
Shanghai, 200090, China;
E-mail: jypdg@163.com
[2]CIMS Research Center, Tongji University
Shanghai, 200092, China
E-mail: hzhangk@yahoo.com.cn

ABSTRACT

The condition-based diagnosis of facilities is a subject based on inter-disciplines and developed with the practical requirement of industrial process. Condition-based diagnosis is a whole working condition diagnosis strategy, and it should evaluate the conditions of the facilities, such as health condition, sub-health condition or morbidity condition. As to the thermal facilities in power plant such as sensor, electric actuator and so on, a diagnosis strategy based on neural network fuzzy evaluation is presented in this paper. Many experimental studies have proved its effectiveness.

Keywords: Condition-based Diagnosis, Neural Network, Fuzzy Evaluation, Thermal Facilities, Actuator.

1. INTRODUCTION

Condition-based diagnosis of facilities is a new subject, which has been developing in the recent 40 years. It is an inter-disciplinary subject due to the practical requirements in the industrial process. With the development of the management level of the facility maintenance and fault diagnosis technology, condition-based diagnosis turns into practical application gradually. The theoretical research has gone further whilst the production has been practised. Lots of achievements have been reached from the research and the production practice, which have won the attention from the world.

The condition-based diagnosis of facilities is a whole working condition diagnosis strategy presented in this paper. The condition-based diagnosis should evaluate the conditions of the facilities: whether they are in health condition, sub-health condition or morbidity condition, in other words, whether the facilities are in a natural condition, exceptional condition or fault condition. However fault diagnosis technology is just a judgement when the facility is in morbidity condition.

At present the technology of condition-based diagnosis of facilities has been used in some of the main large facilities in many modern power plants, such as the boiler, turbine and generator, etc, but not used in thermal facilities. With the continuous increase of the capacity of boiler-turbine unit, the thermal facilities are increased too, and they are becoming more huge and complex. This may result in security accident and bring out tremendous economic loss if any mistake took place. So the facility maintenance is carried out through its running condition. Combining the condition-based diagnosis technology and thermal facilities in power plant can obtain facilities conditions accurately and realize the condition-based diagnosis and maintenance of thermal facilities effectively. As to the thermal facilities in power plant such as electric actuator [1], this paper proposed a diagnosis strategy based on neural network fuzzy evaluation. Many experimental studies have proved that this strategy is feasible and effective.

2. CONDITION-BASED DIAGNOSIS BASED ON NEURAL NETWORK FUZZY EVALUATION

For the running condition of a facility, it's commonly classified as normal and non-normal condition. But more usually, it's classified as health condition, sub-health condition and morbidity condition. Here, an electric actuator in power plant was selected in order to illustrate the strategy that this paper proposed.

2.1. ELECTRIC ACTUATOR AND ITS WORKING PRINCIPLE

Electric actuator is a position-executing machine, which is driven by a two-phase servo electromotor. Its output axis has two run mode: angle displacement and beeline displacement. The working principle of them is basically the same. In this paper, the angle displacement electric actuator was selected and the structure of it is shown in figure 1.

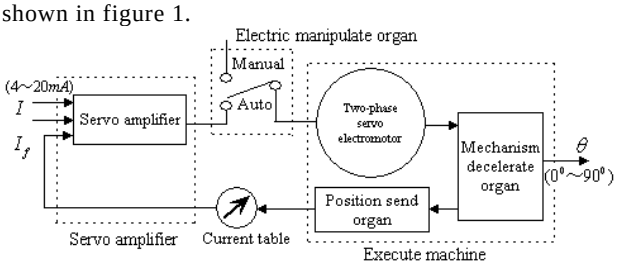


Fig. 1. The structure of electric actuator

The angle displacement electric actuator is a deep negative feedback closed system, which is composed by the independent servo amplifier and executing machine. The main characteristic parameters of electric actuator are pure delay, raising rate, dead zone, return error, and so on. All these parameters can be obtained by the data collection and analysis software, which is designed by the authors of this paper.

2.2. STRATEGY OF CONDITION-BASED DIAGNOSIS

In order to estimate effectively what working condition the electric actuator is, a strategy of condition-based diagnosis based on neural network fuzzy evaluation is proposed. This strategy is as follows. Firstly, fuzzy quantify the main characteristic parameters of electric actuator which are used as the input information of neural network. Secondly, the

neural network is trained to the stage where the desired target is satisfied by using the error of neural network output and the Heming degree or Euclid degree of the fuzzy quantified characteristic parameters [2,3]. Thirdly, the weights of the trained neural network are saved for diagnosing. Then, it can obtain the running conditions of electric actuator through the fuzzy evaluation rules [4].

2.3. FUZZY QUANTIFY FOR CHARACTERISTIC PARAMETERS OF ELECTRIC ACTUATOR

Table 1 shows the five samples of the main characteristic parameters of an electric actuator. In table 1, sample No.1 is the original characteristic parameters. No.2, No.3 and No.5 are the characteristic parameters of electric actuator when working on health condition. No. 4 is the characteristic parameters on morbidity condition.

However these samples can't be used as the inputs of neural network directly. By the use of fuzzy mathematics, these samples can be transformed to fuzzy information variables with close zone of [0，1] by using of a fuzzy subject functions, and these variables can be used as the input information of neural network. Here we suppose that every variable matches two quantified inputs of neural network. It also can match three or more quantified inputs of course, which is decided by the practical needs. In this paper, defining the fuzzy subject functions of the two quantified inputs for every variable as follows

$$\mu_S = \exp(-kx^2) \tag{1}$$

$$\mu_L = 1-\exp(-kx^2) \tag{2}$$

Where, μ_S is the fuzzy subject function of small deflection, μ_L is the fuzzy subject function of large deflection, x is the real variable of subject function, and k is a selected proper parameter.

Table 2 shows the data of each sample quantified by the fuzzy subject functions. $x_1 \sim x_8$ are the input information of neural network, y is the Heming degree, which is used as the output teacher signal of the neural network.

Table 1. Main characteristic parameters of electric actuator

Targets and samples		1	2	3	4	5
1	Pure delay τ	1.0	0.22	0.80	1.20	0.50
2	Raising rate k	0.55	0.54	0.60	0.50	0.58
3	Dead zone Δ	3.0	1.22	2.0	3.50	2.50
4	Return error θ	1.50	1.24	1.40	1.80	1.20

Table 2. Fuzzy subject function values of characteristic parameters

Fuzzy subclass and samples			1	2	3	4	5	Subject functions
Inputs	Pure delay τ	small deflection x_1	0.37	0.95	0.53	0.24	0.78	e^{-x^2}
		large deflection x_2	0.63	0.05	0.47	0.76	0.22	$1-e^{-x^2}$
	Raising rate k	small deflection x_3	0.74	0.75	0.70	0.78	0.71	e^{-x^2}
		large deflection x_4	0.26	0.25	0.30	0.22	0.29	$1-e^{-x^2}$
	Dead zone Δ	small deflection x_5	0.84	0.97	0.92	0.78	0.88	$e^{-0.02\times x^2}$
		large deflection x_6	0.16	0.03	0.08	0.22	0.12	$1-e^{-0.02\times x^2}$
	Return error θ	small deflection x_7	0.80	0.86	0.82	0.72	0.87	$e^{-0.1\times x^2}$
		large deflection x_8	0.20	0.14	0.18	0.28	0.13	$1-e^{-0.1\times x^2}$
Output	Heming degrees	y	0.68	0.88	0.74	0.63	0.81	--------

2.4. STRUCTURE OF NEURAL NETWORK FUZZY EVALUATION

As analyzed above and from the data shown in Table 2, we can confirm the neural network has eight inputs and one output. The hidden layer number of the neural network can be chosen randomly, but it also can ensure the best number by several experiments. In this paper, we choose five hidden layer numbers. Hence, the structure of neural network for fuzzy evaluation is shown as Fig. 2.

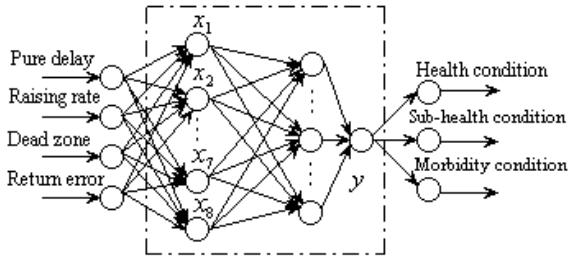


Fig. 2. Structure of neural network fuzzy evaluation

As it shown in figure 2, the structure is similarly to fuzzy BP neural network [5], and the algorithm in the dashed frame is improved BP neural network algorithm in this paper. The output y of the network has three conditions by the reverse fuzzy quantified rules, here, it named as fuzzy evaluation rules as follows

$$y_{Out} = \begin{cases} \text{Health condition} & y > y_1 \\ \text{Sub-health condition} & y_2 \le y \le y_1 \\ \text{Morbidity condition} & y < y_2 \end{cases} \quad (3)$$

In this paper, choosing $y_1 = 0.7$ and $y_2 = 0.65$. It need to be pointed out that the fuzzy evaluation rules are defined by lots of experiments and some experience of the local experts.

Figure 3 shows the error curve of neural network, which was trained with the inputs and outputs data in table 2. The weights of trained neural network are saved as diagnosis expert knowledge database for diagnosing online or offline.

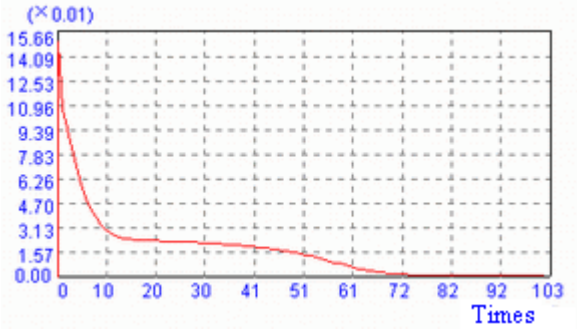


Fig. 3. Error curve of the trained neural network for fuzzy evaluation

2.5. RESULTS OF NEURAL NETWORK FUZZY EVALUATION DIAGNOSIS

Collecting the data of a health condition running electric actuator and obtained the sample parameters are { Pure delay、Raising rate、Dead zone、Return error }={0.60、0.58、1.80、1.20}. It can obtain the quantified input information of neural network as { x_1、…、x_8 }={0.70、0.30、0.71、0.29、0.94、0.06、0.86、0.14} according to the fuzzy quantify strategy shown in table 2. The calculated output of neural network is 0.79, which shows that the electric actuator is working on the health condition. The Heming degree of the sample parameters is 0.80, which equals to the output of neural network almost. That is to say, the result of the diagnosis based on neural network fuzzy evaluation is right.

On the other hand, collecting the data of a morbidity condition running electric actuator and obtained the sample parameters are { Pure delay、Raising rate、Dead zone、Return error }={1.15、0.62、3.58、1.86}. The calculated output of neural network is 0.62, which shows that the electric actuator is working on the morbidity condition.

3 CONCLUSIONS

Condition-based diagnosis of facilities is a new developing subject based on inter-disciplines due to the practical requirement of industrial process. With the development of management level of the facility maintenance and fault diagnosis technology, condition-based diagnosis turns into practicality gradually.

At present, condition-based diagnosis of facilities has been used in some of the main large facilities in many modern power plants. As to the thermal facilities of electric actuator in power plant, this paper proposed a diagnosis strategy based on neural network fuzzy evaluation. Many experimental studies have proved that this strategy is feasible and effective.

ACKNOWLEDGEMENTS

This work was supported by Shanghai Education Committee Key Project (Project Number: 05ZZ53) and supported by Shanghai Leading Academic Discipline Project (Project Number: P1303).

REFERENCES

[1] H. Wang, S. Daley, “Actuator Fault Diagnosis: An Adaptive Observer-Based Technique”, IEEE Transactions on Automatic control, July 1996, 41(7): 1073-1078.

[2] J.P. Wu, J.H. Xiao, Intelligent faulty diagnosis and expert system, Beijing: Science Press, 1997.

[3] D.H. Zhou, Y.Z. Ye, Moden faulty diagnosis and tolerance control, Beijing: Tsinghua Press, 2000.

[4] J.M. Kocielny, M. Barty, “Application of information theory for actuators diagnosis”, IFAC Symposium SAFERPROCESS, Budapest, Hungary, 2000,Vol.2: 949-1054.

[5] K. Patan, T. Parisini, “Dynamic neural networks for actuator fault diagnosis: application to the damadics benchmark problem”, Proceedings of IFAC Symposium SAFEPROCESS, Washington, USA, 2003, 1077-1082.

Supplier Relationship Management Based on Stock-outs Compensation Mechanism in E-commerce

Guang Shu Chang
Research Division for Quality Engineering, Zhengzhou Institute of Aeronautics
Zhengzhou 450015, P.R. China
Management School, Northwestern Polytechnical University
Xi'an 710072, P.R. China
Email: cgs30@163.com

ABSTRACT

The buyer-supplier relationship will affect the operation efficiency of the whole supply chain system. However, the ubiquitous phenomenon on stock-outs, especially in e-commerce, will greatly affect the buyer-supplier relationship because the lost is usually afforded by the buyer, such as the sales decrease, alteration cost, reputation damage, etc. In this paper, we proposed that supplier should afford some proportional compensation for buyer in case of stock-outs, so as to improve the collaboration and relationship between supplier and buyer. Then, a compensation model for stock-outs was established by exploiting the principle of non-linear programming, and an optimal solution for order quantity was provided. Finally, the validity of the model was proved through a case study that resulted that coordinated operation of supply chain and win-win situation is realized if supplier provides compensation for buyer for stock-outs.

Keywords: Supply Chain, E-Commerce, Supplier, Relationship Management, Stock-outs.

1. INTRODUCTION

Since there is more and more integration and collaboration among the members of supply chain, management of supply chain is essentially management of the relationships and activities among the member organizations [1]. Therefore, the relationship management of the correlative organizations has been paid more and more attention by the academic researches and practices of supply chain management [2]. Especially, the ability of establishing long-term partnership based on trust with suppliers has been a new competition advantage. Up to now, the academic researches about supplier relationship management (SRM) have acquired some results. For example, Krapfel et al. proposed a strategic relationship mapping model based on the relationship types they classified [3]; Cox developed a five-step model to illuminate the development of the range of supplier relationships [4]; Olsen & Ellram developed a three-step model to assist in managing different kinds of supplier relationships [5]; Jap & Mohr established a Relationship×Technology Matrix to analyze how web-enabled efficiencies facilitate critical outcomes [6]; Roberts & Mackay discussed how the e-commerce applications are used to support the portfolio of supplier relationships [7], and so on. These research results can be classified into two types: one is New Institutional Economics based on transaction cost analysis; the other one is organization theory based on network analysis.

In stochastic inventory management strategies of supply chain, the phenomenon on backorders is ubiquitous because suppliers have various problems. Stock-outs are inevitable because it is not necessary that all times of the demand of customer can be met even if there is safety stock, especially in e-commerce circumstance [8]. When stock-outs, more than half of customers of a manufacturer may switch to the rivals and a retailer may lose 14 percent of its customers [9]. Furthermore, not only the sales of present products out of stock are impaired, but also the subsequent sales and other products will be affected [10]. Therefore, stock-outs will affect inevitably the buyer-supplier relationship.

However, viewing from all the existing related literatures, it is almost assumed that the buyers should bear all kinds of loss which is resulted from stock-outs, such as the decrease of sales, alteration cost, the damage on fame, etc. Therefore, it will be feasible that initiator, i.e. the supplier, make appropriate compensation for buyer in case of stock-outs so as to avoid the disaggregation of the supply chain system and maintain cooperative relationship between buyer and supplier. From the point view of supplier, we establish a relationship management model by utilizing of stock-outs compensation mechanism (SOCM) to weaken the effect of stock-outs on relationship between buyer and supplier.

This paper is organized as follow. First, in section 1, the status quo and documents about inventory research is under discussion. Second, in section 2, the traditional EOQ model with planned stock-outs is reviewed. Third, in section 3, the compensation model for stock-outs is established and an optimal solution to order quantity is provided. Fourth, in section 4, the validity of our model is verified through a case study. Finally, in section 5, concluding remark is offered.

2. TRADITIONAL EOQ MODEL WITH PLANNED STOCK-OUTS

Under the traditional EOQ model, it is usually assumed that market demand and lead-time are known, replenishment ratio is infinite, all the orders are satisfied in one time, and stock-outs are not allowed. It is actually Utopian. In our true life, market demand and lead-time are all stochastic variables. Therefore, EOQ model with planned stock-outs is proposed to overcome the limitations in the traditional EOQ model. Just as its name implies, the most remarkable of the EOQ model with planned stock-outs is that planned stock-outs is allowed. There are several basic assumptions in it as follow:

The demand rate for inventory system from external market is known, even and constant;

Lead-time is also known;

Goods ordered for replenishing inventory can be delivered at any time when need;

Planned stock-outs is allowed and customer must wait.

Replenishment ratio is infinite and all the orders are satisfied in one time.

We define the following variables:

h: unit holding cost;
b: unit stock-outs cost;
A: setup cost;
S: stock-outs quantity;
Q: order quantity;
D: total demand of a year;
T: cycle time;
TC: total cost per cycle time.
C(*x*,*Q*): total variable cost per cycle time.

Then, the basic model is:

$$TC = A + \frac{(Q-S)^2}{2D}h + \frac{S^2}{2D}b \tag{1}$$

Then,

$$c(x,Q) = \frac{AD}{Q} + \frac{(Q-S)^2}{2Q}h + \frac{S^2}{2Q}b \tag{2}$$

On the right of Eq. (2), the first item is unit setup cost, the second item is unit-holding cost, and the third item is unit stock-outs cost. Thus, the optimal order quantity and maximal stock-outs quantity is:

$$Q^* = \sqrt{\frac{2AD(h+b)}{hb}} \tag{3}$$

$$S^* = \frac{h}{h+b}Q^* \tag{4}$$

The way of inventory changing along with time in the EOQ model with planned stock-outs is shown as Figure 1.

We denote *x* as critical level of inventory. Let $x = \frac{h}{h+b}$, then $S^* = xQ^*$. Thus, if stock-outs cost *b* is less and less, S^* will be less and less too, namely, stock-outs quantity should be small. In contrary, if *h* is bigger while *b* is rather small, then S^* will be close to Q^*, namely, stock-outs quantity can be augmented properly.

Substituting *x* into Eq. (2), we can get:

$$c(x,Q) = \frac{AD}{Q} + \frac{Q(1-x)^2}{2}h + \frac{x^2Q}{2}b \tag{5}$$

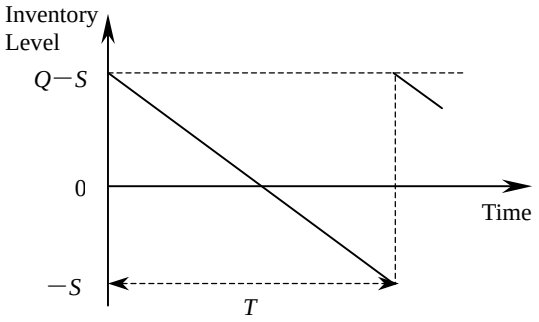


Fig. 1. EOQ model with planned stock-outs

3. RELATIONSHIP MANAGEMENT MODEL BASED ON SOCM

There is an inherent limitation in traditional EOQ model with planned stock-outs, namely it belongs to distributed SCM in which each member is independent rational market player and pursue maximal profit for himself without concerning the relationship coordination among all the members. This kind of distributed decision-making cannot ensure the realization of the holistic optimization of the supply chain system. The most outstanding exhibition is the need of setting safety cost and the payment of high holding cost. Furthermore, one of the most important assumptions in the model is that the demand rate for inventory system from external market is fixed whether suppliers have stock or not. However, as above-mentioned, this assumption is actually impractical. The limitation of the model makes it outdated not in conformity to supply chains in pursuit of zero-stock status. Therefore, a more reasonable assumption should be that the demand rate confronted by suppliers is different whether stock-outs or not.

In this paper, we propose that supplier should provide a certain proportion of compensation for buyer when stock-outs in order to promote the relationship between supplier and buyer and realize win-win situation. In contrast with traditional model, a new cost is introduced in our model, i.e. compensation cost. We denote *r* as unit compensation cost. Since the reactions of buyers are different with different factors, such as product, purchasing motive, exigent degree of needs, waiting time, risk of searching substitute supply, and so on. Therefore, the amount of compensation will be determined accordingly as per different factors, such as transaction amount, waiting time, important degree of buyer, etc. [9]. Demand brought by providing compensation, D_r, can be determined easily through market investigation or business experience.

The way of inventory changing along with time in our model is shown as Fig. 2. In contrast with Fig. 1., the inventory curve in Fig. 2. is no longer a beeline but a broken line because demand rate changes.

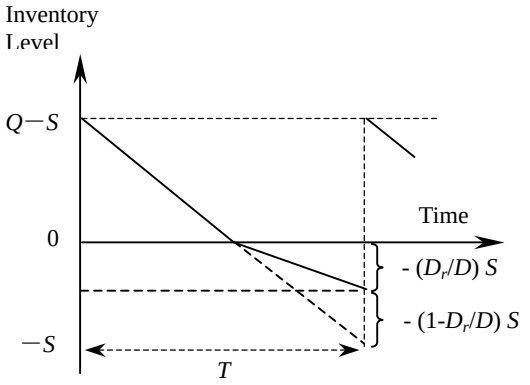


Fig. 2. Relationship Management Model Based on SOCM

From Fig. 2., we can establish the total variable cost model as follow:

$$TC = A + \frac{(Q-S)^2}{2D}h + \frac{D_r S^2}{2D^2}b + \frac{D_r}{2D^2}S^2 r \tag{6}$$

Then,

$$c(x,Q) = \frac{AD}{Q} + \frac{(Q-S)^2}{2Q}h + \frac{D_r S}{2D}\frac{hb}{h+b} + \frac{D_r S}{2D}\frac{h}{h+b}r$$

$$= \frac{AD}{Q} + \frac{Q(1-x)^2}{2}h + \frac{D_r Q x^2}{2D}b + \frac{D_r Q x^2}{2D}r \tag{7}$$

On the right of Eq. (7), the third item is new unit stock-outs cost, and the fourth item is unit compensation cost.

Unlike in the EOQ model with planned stock-outs, the optimal *x* is not independent of *Q* in our model. Therefore, we need to minimize a nonlinear programming problem over

x and Q. Here, we consider three cases for finding the optimal solution.

3.1 Case 1: No Stock-outs

The buyer has to have inventories always to minimize his costs, namely $x^*=0$.

$$TC = A + \frac{Q^2}{2D}h \tag{8}$$

Then,

$$c(x,Q) = \frac{AD}{Q} + \frac{Q}{2}h \tag{9}$$

Since the cost function is convex, it is easy to get the solution. Setting the first derivate of $c(x, Q)$ with respect to Q equal to zero we obtain:

$$Q^* = \sqrt{\frac{2AD}{h}} \tag{10}$$

Substituting x^* and Q^* into Eq. (9), we can get:

$$c(x^*, Q^*) = \sqrt{2ADh} \tag{11}$$

Here, supplier need not provide compensation because buyer has no stock-outs.

3.2 Case 2: No Inventory

The buyer has no inventory and stock-outs rate is percent, namely $x^*=1$. Then,

$$TC = A + \frac{D_r Q^2}{2D^2}b + \frac{D_r}{2D^2}Q^2 r \tag{12}$$

Thus,

$$c(x,Q) = \frac{AD}{Q} + \frac{D_r}{2D}Q(b+r) \tag{13}$$

Since the cost function is also convex, it is easy to get the solution. Setting the first derivate of $c(x, Q)$ with respect to Q equal to zero we obtain:

$$Q^* = \sqrt{\frac{2AD^2}{D_r(b+r)}} \tag{14}$$

Substituting x^* and Q^* into Eq. (13), we can get:

$$c(x^*, Q^*) = \frac{1}{2}\left(\frac{D_r}{D} + \sqrt{\frac{2AD_r}{b+r}}\right)(b+r) \tag{15}$$

3.3 Case 3: Some Inventory and Some Stock-outs

Since it is a nonlinear programming problem, we need to check whether the function is convex or not.

$$c(x,Q) = \frac{AD}{Q} + \frac{Q(1-x)^2}{2}h + \frac{D_r Q x^2}{2D}b + \frac{D_r Q x^2}{2D}r \tag{16}$$

The Hessian Matrix of Eq. (16) is:

$$H(x,Q) = \begin{pmatrix} \frac{\partial^2 c(x,Q)}{\partial x^2} & \frac{\partial^2 c(x,Q)}{\partial x \partial Q} \\ \frac{\partial^2 c(x,Q)}{\partial Q \partial x} & \frac{\partial^2 c(x,Q)}{\partial Q^2} \end{pmatrix}$$

$$= \begin{pmatrix} Qh + \frac{D_r}{D}Qb & (x-1)h + \frac{D_r}{D}xb \\ (x-1)h + \frac{D_r}{D}xb & \frac{2AD}{Q^3} \end{pmatrix} \tag{17}$$

The first-principal of Eq. (17) is nonnegative, but the second principal could be positive or negative. It means that the function is neither convex nor concave. Therefore, we cannot apply well-known *Sylvesters Theorem* to judge the function is positive definite or negative definite, semi-positive definite or semi-negative definite. However, we can get the solution through another method. Since $c(x, Q)$ has only one optimal solution, its local minimum is just its global minimum [11]. Therefore, we can determine firstly its local minimum and replace it as its global minimum. The calculation process is as follow:

$$\nabla c(x,Q) = \begin{pmatrix} \frac{\partial c(x,Q)}{\partial x} \\ \frac{\partial c(x,Q)}{\partial Q} \end{pmatrix}$$

$$= \begin{pmatrix} (x-1)Qh + \frac{D_r(b+r)}{D}Qx \\ -\frac{AD}{Q^2} + \frac{(1-x)^2}{2}h + \frac{D_r(b+r)}{2D}x^2 \end{pmatrix} = 0 \tag{18}$$

Let $b' = \frac{D_r}{D}b$, $r' = \frac{D_r}{D}r$ and substitute them into Eq. (18), we obtain:

$$x^* = \frac{h}{h+b'+r'} \tag{19}$$

Substituting x^* into Eq. (16), we can get:

$$Q^* = \sqrt{\frac{2AD(h+b'+r')}{h(b'+r')}} \tag{20}$$

If $D_r=D$, then supplier need not provide compensation because buyer has no lost, namely $r=0$. Therefore, Eq. (6) will be translated as:

$$TC = A + \frac{(Q-S)^2}{2D}h + \frac{D_r S^2}{2D^2}b + \frac{D_r}{2D^2}S^2 r$$

$$= A + \frac{(Q-S)^2}{2D}h + \frac{S^2}{2D}b \tag{21}$$

We can find that Eq. (6) and (21) are consistent with Eq. (1). It means that the traditional EOQ model with planned stock-outs is a special case of our model.

In contrast with Eq. (3), not only stock-outs cost but also compensation cost is considered in Eq. (20). Furthermore, the effect of the evocable demand brought by compensation on demand rate of external market is also considered in stock-outs cost and compensation cost of Eq. (20). Therefore, the optimal order quantity provided by Eq. (20) can reflect more accurately the demand rate confronted by the supply chain system and promote the coordination of the complete operation of the supply chain system.

4. CASE STUDY

Here, we consider a main-board supplier who supply synchronously for five computer manufacturers, namely M_1, ⋯, M_5. The order quantities of the manufacturers are

different because the market demand they confronted is different. Their total demand of a year, holding cost, stock-outs cost, setup cost, evocable demand, unit compensation provided by supplier, and the calculation results of optimal order quantities are listed in Table 1.

Table 1. The optimal order quantity of the computer manufacturers

	D	h	b	A	R	D_r	b'	r'	Q^*	Q_r^*
M_1	100	1.5	1.6	40	1.0	110	1.76	1.10	102	91
M_2	300	1.4	1.7	120	1.2	350	1.98	1.40	306	270
M_3	500	1.3	1.8	200	1.3	570	2.05	1.48	514	459
M_4	700	1.3	1.9	280	1.4	810	2.20	1.62	713	636
M_5	1000	1.2	2.0	400	1.5	1200	2.40	1.80	1032	926

From Table 1., we can find that the order quantities of the manufacturers without supplier's compensation are usually greater than demand to prevent stock-outs. While supplier provides compensation, the order quantities will be less than demand. It means that the holding cost and risk resulted by inventory of the manufactures are reduced while supplier maintain cooperative relationship with the manufactures and avoid the decrease of sales. Therefore, coordinated operation and win-win situation is realized.

5. CONCLUSION

In this paper, we study the problem of relationship management in supply chain management from the point view of supplier providing a certain proportion of compensation for buyers when stock-outs occur. Our findings indicate that the demand rate for inventory system of the supply chain from external market will change when there is lost resulted by stock-outs. Therefore, it will be feasible that initiator, i.e. the supplier, make appropriate compensation for buyer in case of stock-outs so as to maintain cooperative relationship with buyer and improve the operation efficiency of complete supply chain system.

ACKNOWLEDGEMENT

This research was partially supported by the National Natural Science Foundation of China under Grant 70125004, 70572050 and 70401011.

REFERENCE

[1] S.L. Golicic, D.F. Davis, et al., "The impact of e-commerce on supply chain relationships", International Journal of Physical Distribution & Logistics Management, 2002, Vol. 32(10): 851-871.

[2] F. Lemke, K. Goffin, et al., "Investigating the meaning of supplier-manufacturer partnerships: An exploratory study", International Journal of Physical Distribution & Logistics Management, 2003, Vol. 33(1): 12-35.

[3] R.E. Krapfel, D. Salmond, et al., "A strategic approach to managing buyer-seller relationships", European Journal of Marketing, 1991, Vol. 25(9): 22-37.

[4] A. Cox, "Relational competence and strategic procurement management", European Journal of Purchasing and Supply Management, 1996, Vol. 2(1): 57-70.

[5] R.F. Olsen, L.M. Ellram, "A portfolio approach to supplier relationships", Industrial Marketing Management, 1997, Vol. 26(2): 101-113.

[6] S.D. Jap, J.J. Mohr, "Levering Internet technologies in B2B relationships", California Management Review, 2002, Vol. 44(4): 24-38.

[7] B. Roberts, M. Mackay, "IT supporting supplier relationships: The role of electronic commerce", European Journal of Purchasing & Supply Management, 1998, Vol. 4(2/3): 175-184.

[8] R.Q. Chen, S.H. Ma, "Production and operation management", China Machine Press, Beijing, China, 2004. [In Chinese]

[9] M. Emmelhainz, L. Emmelhainz, et al., "Consumer response to retail stock-outs", Journal of Retailing, 1991, Vol. 67(2): 138-147.

[10] K. Campo, E. Gijsbrechts, et al., "Towards understanding consumer response to stock-outs", Journal of Retailing, 2000, Vol. 76(2): 219-242.

[11] D. Sun, "Inventory management in e-business", Working paper, Pennsylvania State University, USA, 2001.

Empirical Study on Stock Market Through ARCH Model Based on Particle Swarm Optimization

Jun Sun[1], C. -H. Lai[2], Wenbo Xu[1]
[1]School of Information Technology, Southern Yangtze University,
No.1800, Lihu Dadao, Wuxi Jiangsu 214122, China
Email:sunjun_wx@hotmail.com
Email:xwb@sytu.edu.cn
School of Computing and Mathematical Sciences,
University of Greenwich,
Greenwich, London SE10 9LS, UK.
Email:C.H.Lai@gre.ac.uk

ABSTRACT

In this paper, we discuss the properties of ARCH model and propose an approach for estimating the parameters in ARCH model, the method of maximum likelihood (ML) with particle swarm optimization (PSO). The purposes of using PSO algorithm is to obtain the global maximum value of likelihood function as well as to make ARCH model applicable to immediate disposition of a vast amount of data from financial market. The ARCH model based on PSO is applied to empirical study of real assets returns analysis. The sample data we use in this study is Dow Jones Industrial Average index return. Results show that the PSO-based ARCH model is effective in predicting the volatility in the US Stock Market

Keywords: Time Series, ARCH Model, PSO, Stock Market.

1 Introduction

Many Financial time series exhibit *volatility clustering*, which means that the series have periods where volatility is low and other periods where volatility is high. As econometricians, we will understand “volatility" to mean “conditional variance”. The conditioning set will always be the behavior of relevant variables up to “time t" --the time when we observe the series. There have recently been developed many different models for data that shows volatility clustering, and it is still a *very* active research area. The main differences between competing models will often be the choice of which variables to condition on, and (as usual) the choice of functional forms. The problems of modeling series with varying volatility is of course well known in econometrics under the heading of heteroskedasticity; but most of the interest in time-varying volatility models comes from Finance.

One of the main ideas of asset pricing is that the variability of an asset should be reflected in its price [8]. One would expect an asset with high variance (“risk") to give a higher return (for investors to want to hold it). With some polishing and generalizations this is the main point of the CAPM-, and APT-models that are common in Finance. But standard economic reasoning also says that risk should not be measured as the unconditional variance but rather as the conditional variance, and therefore Finance characters are interested in modeling conditional heteroskedasticity in its own right (e.g. for the purpose of pricing options)

The most popular approach for modeling heteroscedasticity is a class of Autoregressive Conditional Heteroskedasticity (ARCH) models originally introduced by Engle in 1982 [1]. ARCH models have a variety of characteristics, which make them attractive for applications in different fields. According to Mac Nees, the leading economist, “the inherent uncertainty or randomness associated with different forecast periods seems to vary over time and large and small errors tend to cluster together”. This certainly suggests application of ARCH models in econometrics data. In finance, portfolios of financial assets are held as functions of the expected mean and variance of the rate of return. Since any shift in asset demand must be associated with changes in expected mean and variance of rate of return, ARCH models are the best suitable models. In regression, ARCH models can be used to approximate the complex models. In this case, presence of ARCH disturbances can be interpreted as evidence of misspecification either by omitted variables or through structural change

In this paper, we use Particle Swarm Optimization (PSO), an evolutionary search technique, to estimate ARCH model, and therefore explore the applicability of the PSO-based ARCH model to the empirical study on US stock market. The paper organized as follows. In the next section, ARCH model is discussed. In Section 3, the methodology of PSO algorithm is introduced. Section 4 presents the empirical results on US stock market and the conclusion is made in Section 5

2 An Introduction to ARCH Models

The approach to modeling financial time series volatility has changed significantly since the development of the autoregressive conditional heteroscedasticity (ARCH) model by Engle [1]. Later on, this model has been generalized to provide richer modeling of volatility dynamics. Specifically the generalized autoregressive conditional heteroscedasticity (GARCH) model by Bollerslev [6] [7] has proved popularity in modeling financial data. In general, it seems that ARCH models are applicable to a wide range of financial data across stock, bond, foreign exchange and derivative markets. In this section, we review ARCH (p) model briefly and discuss the approach for estimating ARCH (1) model.

2.1 ARCH (p) Model

Consider data $X_1, \cdots, X_n$ arising as an observed stretch from a financial returns time series $\{X_t, t = 0, \pm 1, \pm 2, \cdots\}$ such as the percentage returns of

a stock price, stock index or foreign exchange rate. The returns series $\{X_t\}$ will be assumed strictly stationary with mean-- zero which--from a practical point of view--implies that trends and other nonstationarities have been successfully removed

The celebrated ARCH models of Engle [1] [2] were designed to capture the phenomenon of volatility clustering in the returns series. An ARCH (*p*) model can be described by the following equation:

$$X_t = \sigma_t \varepsilon_t \tag{1}$$

$$\sigma_t^2 = \alpha + \sum_{i=1}^{p} \alpha_i X^2_{t-i} \tag{2}$$

Where $\alpha_i (i = 1,2,\cdots,q)$ and α are scalar parameters to be estimated. ε_t is supposed to be an i.i.d (0,1) process. In order to obtain positive values for the estimate of the condition variance, we have to assume that α_i and α are positive, that is

$$\alpha_i > 0, \qquad i = 0,1,\cdots,p \tag{3}$$

And that

$$\alpha_1 + \cdots + \alpha_p < 1 \tag{4}$$

2.2 Properties of ARCH (1) Model

This paper focuses on ARCH (1) model, whose conditional variance of regression error has the form

$$X_t = \sigma_t \varepsilon_t \tag{5}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 \tag{6}$$

Consider the first mean of X_t

$$\begin{aligned} E(X_t) &= E[E(X_t \mid X_{t-1})] \\ &= E[E(X_t[\alpha_0 + \alpha_1 X^2_{t-1}]^{1/2}] = 0 \end{aligned} \tag{7}$$

Now consider the unconditional variance of X_t. Assume that $\alpha_0 > 0$ and $|\alpha_1| < 1$. Note that

$$X^2_t = \varepsilon^2_t [\alpha_0 + \alpha_1 X^2_{t-1}] \tag{8}$$

Assuming the stationarity of X_t gives that

$$\sigma^2 = \alpha_0 + \alpha_1 \sigma^2 \tag{9}$$

Where $\sigma^2 = E[X^2_t] = E[X^2_{t-1}]$. Thus $\sigma^2 = \alpha_0 / (1 - \alpha_1)$. (This expression explain why we assume that $\alpha_0 > 0$ and $|\alpha_1| < 1$.) It is worth noting that the conditional variance of X_t is homoskedastic. Finally, consider first-order autocovariance

$$E(X_t X_{t-1}) = E[X_{t-1} E(X_t \mid X_{t-1})] = 0 \tag{10}$$

Therefore, X_t's are serially uncorrelated (obviously, they are not independent)

2.3 Estimation of ARCH (1) Model

ARCH models are estimated by the method of maximum likelihood (ML). The resulting ML estimator is consistent and asymptotically normal so that the standard procedure can be used to carry out statistical inference. Considering the following simple model

$$Y_t = \mu + X_t \tag{11}$$

Where

$$X_t = \varepsilon_t \sigma_t = \varepsilon_t [\alpha_0 + \alpha_1 X^2_{t-1}]^{1/2} \tag{12}$$

Where $\theta(\mu, \alpha_0, \alpha_1)$ is unknown. This simple model can be used to describe the behavior of returns on financial assets. If ε_t ~i.i.d. N (0,1) with ε_t independent of Y_{t-1}, then the conditional distribution of Y_t given Y_{t-1} is normal. More specifically the conditional probability density function of Y_t given Y_{t-1} is

$$f(Y_t \mid Y_{t-1}; \theta) = \frac{1}{\sqrt{2\pi\sigma_t}} \exp\left(\frac{-(Y_t - \mu)^2}{2\sigma_t} \right) \tag{13}$$

Therefore, the ML estimator of θ can be obtained by maximizing the log likelihood

$$L(\theta) = \sum_{t=1}^{T} \log f(Y_t \mid Y_{t-1}; \theta). \tag{14}$$

That is, the Maximum Likelihood Estimator $\hat{\theta}_T$ is

$$\hat{\theta}_T = \arg\max_{\theta} L(\theta) = \arg\max_{\theta} \sum_{t=1}^{T} \log(f(Y_t \mid Y_{t-1}; \theta)) \tag{15}$$

The next section will review an effective evolutionary optimization algorithm, Particle Swarm Optimization, which can be employed to search the Maximum Likelihood Estimator.

3 Methodology of Particle Swarm Optimization

Particle Swarm Optimization (PSO) [3], originally proposed by J. Kennedy and R. Eberhart, has become a most fascinating branch of evolutionary computation. The underlying motivation for the development of PSO algorithm was social behavior of animals such as bird flocking, fish schooling, and swarm theory. Like genetic algorithm (GA), PSO is a population-based random search technique but that outperforms GA in many practical applications, particularly in nonlinear optimization problems. In the PSO model, each individual is treated as a volume-less particle in the D-dimensional space, with the position and velocity of *ith* particle represented as

$X_i = (x_{i1}, x_{i2}, \cdots, x_{iD})$ and $V_i = (v_{i1}, v_{i2}, \cdots, v_{iD})$. The particles move according to the following equation:

$$V_{id} = V_{id} + c_1 \cdot rand(\cdot) \cdot (P_{id} - X_{id}) + c_2 \cdot Rand(\cdot) \cdot (P_g - X_{id}) \quad (16)$$

$$X_{id} = X_{id} + V_{id} \quad (17)$$

Where c_1 and c_2 are positive constant and rand() and Rand() are two random functions in the range of [0,1]. Vector $P_i = (p_{i1}, p_{i2}, \cdots, p_{iD})$ is the best previous position (the position giving the best fitness value) of particle i called ***pbest***, and vector $P_g = (p_{g1}, p_{g2}, \cdots, p_{gD})$ is the position of the best particle among all the particles in the population and called ***gbest.***

The PSO with inertia weight is called Standard PSO (SPSO), which is described as follows:

1. Initialize an array of particles with random position and velocities inside the problem space.
2. Evaluate the desired objective function (for example minimization) for each particle and compare with the particle's previous best values: If the current value is less than the previous best value, the set the best value to the current value. That is, if $f(X_i) < f(P_i)$, then $X_i = P_i$
3. Determine the current global position minimum among the particle's best positions. That is: $g = \arg \min_{1 \le i \le M} (f(P_i))$ (M is the population size)
4. Compare the current global position to the previous global: if the current global position is less than the previous global position; then set the global position to the current global
5. For each dimension, change the velocity by using $V_{id} = V_{id} + c_1 * rand(\cdot) * (P_{id} - X_{id}) + c_2 * Rand(\cdot) * (P_g - X_{id})$
6. Move to the next position $X_{id} = X_{id} + V_{id}$
7. Repeat steps (2) –(6) until a stop criterion is satisfied **OR** a pre-specified number of iterations are completed.

Many revised versions of PSO algorithm are proposed to improve the performance since its origin in 1995. As far as the dynamics of PSO is concerned, two most influential and efficient approaches are to introduce Inertia Weight ***w*** [4] or Constriction Factor ***K*** [5] into evolution equation (1). Thus equation (1a) is substituted with

$$V_{id} = w * V_{id} + c_1 * rand(\cdot) * (P_{id} - X_{id}) + c_2 * Rand(\cdot) * (P_g - X_{id}) \quad (18)$$

in Inertia-Weight model and replaced by the following equation in Constriction Factor model

$$V_{id} = K \cdot [V_{id} + c_1 \cdot rand(\cdot) \cdot (P_{id} - X_{id}) + c_2 \cdot Rand(\cdot) \cdot (P_g - X_{id})] \quad (19)$$

$$K = \frac{2}{\left| 2 - \varphi - \sqrt{\varphi^2 - 4\varphi} \right|} \quad \varphi = c_1 + c_2, \varphi > 4 \quad (20)$$

4 Empirical Results

In this section, we estimate an ARCH (1) for Dow Jones Industrial Average (DJIA) using Particle Swarm Optimization algorithm. The experiment data is weekly closing value of DJIA index from February 2, 1996 to December 1, 2004 and the sample size is 461. We analyze the continuously compounded percentage rate of return that is taking the first differences of the logarithm of series:

$$X_t = \ln I_t - \ln I_{t-1} \quad (21)$$

where I_t is closing value for DJIA index on week *t*. The Summary statistics of returns is listed on Table 1. It is can be seen from Table 1 that the distribution of return series of DJIA index is approximately a normal distribution

Table 1. Summary statistics of returns

Statistics	X_t
Sample Mean	0.001497
Sample Variance	0.00062
Skewness	-0.677347
Kurtosis	3.475174
Standard Errors	0.001161
Maximum	0.0809
Minimum	-0.153881

In our experiment, Basic PSO (BPSO), Standard PSO (SPSO) and PSO with Constriction Factor (Co-PSO) are used to estimate the parameters in ARCH (1) model. When the algorithms are running, the population size is set to be 50 and the number of iteration is 500 in three PSOs; maximum value of velocity $V_{\max}$ is 0.6 in BPSO and SPSO. In SPSO, the inertia weight *w* is decreasing linearly from 0.9 to 0.4 as time elapsing and the value of c_1 and c_2 are both set to be 2. In Co-PSO, the value of c_1 and c_2 are both set to be 2.2. In all PSOs, the position vector of the particle corresponding to the parameter θ in ARCH (1) is initialized on [0,1]

Table 2. Maximum likelihood estimates of ARCH (1) Model with conditional Normal distribution

	Number of Iterations	*Parameter* α_0	*Parameter* α_1	*Maximum of L(*θ*)*
In Normal	-----	-----	-----	325.2428
Distribution	-----	-----	-----	
Basic PSO	500	1.0443689E-03	0.63528383	377.7598169
SPSO	500	1.0451092E-03	0 .63507891	377.7598174
Co-PSO	500	1.0447838E-03	0 .63504533	377.7598210

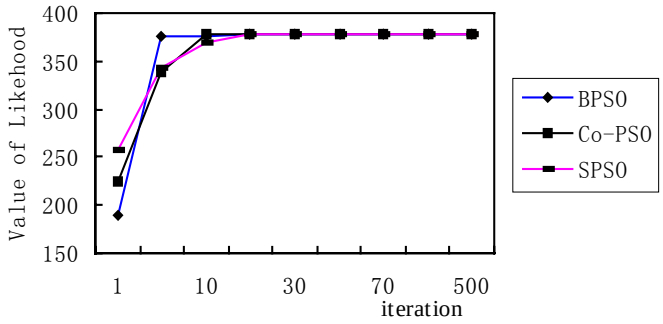


Fig. 1. The figure shows the dynamic changing of the best value of Likelihood as PSO algorithms is running.

The results recorded in Table 1 show that the estimations of ARCH (1) using PSO algorithms are better than the estimation given that the distribution of return series is normal distribution, which gives the facts that the residuals, say $\{\hat{\varepsilon}_t\}$, do not appear to be in accordance to the normality assumption as they are typically heavy-tailed.

It can be seen from figure 1 that the values of the likelihood function in the three PSO algorithms converge to same optimal solution after 70 iterations. Thus we can conclude that PSO algorithm has a very high convergence speed with high precision, which is of momentous significance for the immediate disposition of large amount of data from securities market.

As of the performance of the algorithms, BPSO has the highest convergence speed in early search stage among the three PSO algorithms because its global search ability is stronger than any other PSO algorithm, but its convergence precision is the lowest. Moreover, Co-PSO can get better value of likelihood function than BPSO and SPSO.

Table 3. Actual value (AV) of DJIA index and its predicted value (PV) by three PSO algorithms

	2004-11-8	*2004-11-15*	*2004-11-22*	*2004-11-29*	*2004-12-6*
AV of DJIA	**10539.01**	**10456.91**	**10522.23**	**10592.21**	**10543.22**
PV by BPSO	**10391.72**	**10393.63**	**10437.49**	**10438.39**	**10449.45**
PV by SPSO	**10396.93**	**10429.69**	**10430.58**	**10448.87**	**10461.12**
PV by Co-PSO	**10392.86**	**10401.69**	**10432.13**	**10462.02**	**10485.53**

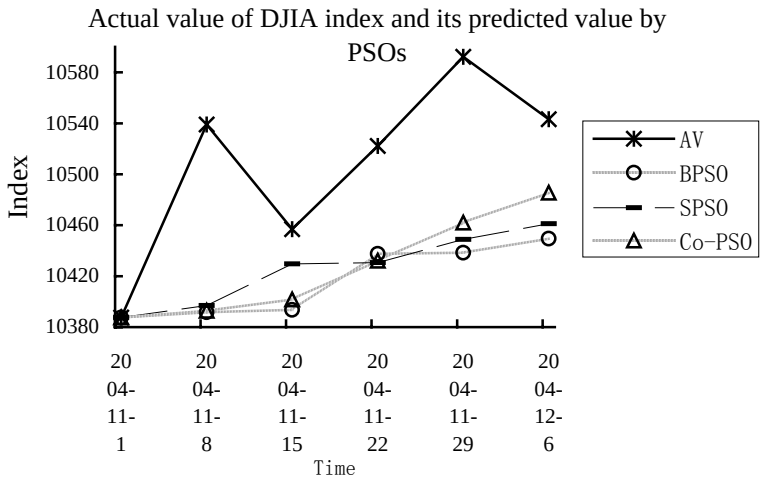


Fig. 2. The figure shows the same actual value and predicted value recorded in Table 2 and it display the prediction errors of ARCH (1) model with PSOs.

From Table 3 or Figure 2, we can see that the values of estimator $\hat{\theta}$ obtain by SPSO and Co-PSO are more accurate than that obtained by BPSO. Consequently, the values of index predicted by the former two PSO algorithms are more accurate than BPSO algorithm.

In ARCH (1) model, the conditional variance of return series at time t is expressed simply as the linear function of

squared error at time t-1, so we are not able to follow the dynamic change of the return very accurately. However, we can acquire some rough information about changing tendency of the index, the information that can help us to avoid financial risk.

5. Conclusions

In this study we investigate the applicability of PSO-based ARCH model on assets return of the US Stock Market. The reason for using PSO in estimating ARCH model is that PSO algorithm is a global search technique and has higher convergence speed than other evolutionary optimization algorithms (e.g. Genetic algorithm). Empirical results reveal that the model is an effective tool in smoothing the volatility in the US Stock Market return. Although we use ARCH (1) model as an example, the PSO-based approach can easily be generalized to other ARCH and its generalization (GARCH) models. Our future work is focusing on exploring the applicability of PSO algorithm to other ARCH model and the empirical study on real assets return.

6. References

[1] R.F. Engel, "Autoregressive conditional hetero skedasticity with estimates of the variance of U.K. Inflation", Econometrics, 1982; 50:987-1008

[2] E.K. Berndt, B.H. Hall, et al, "Estimation and inference in nonlinear structural models". Annals of Economic and Social Measurement, 1974 653-665

[3] J. Kennedy, R. Eberhart, "Particle Swarm Optimization". Proc. IEEE int.Conf. On Neural Network, 1995, 1942-1948

[4] Y. Shi, R.C. Eberhart, "A modified particle swarm optimizer",[R] IEEE Iternational Conference of Evolutionary Computation, Ancharage,Alaska,May 1998

[5] M. Clerc, "The swarm and the queen:towards a deterministic and adaptive particle swarm optimization," Proceedings of the Congress on Evolutionary Computation .Piscataway, NJ:IEEE Service Center,1999,1951-1957

[6] P. Bougerol, N. Picard, (1992) Stationarity of GARCH Processes and of some Nonnegative Time Series. Journal of Econometrics 52, 115-127

[7] I.V. Basawa, P.D. Feigin, et al, (1976) Asymptotic properties of maximum likelihood estimators for stochastic processes. Sankya Series A 38, 259-270

Knowledge-Based Profit Distribution and Bi-level Principal-Agent Supervisory Model of Virtual Enterprises *

Dongbo Liu [1,2]
[1] **Research Institute of Automation, East China University of Science and Technology**
Shanghai, 200237, China
[2] **College of Mechanical and Electronic Engineering, Shanghai Normal University**
Shanghai, 201418, China
Email: liudbyj@163.com
Dao Huang
Research Institute of Automation, East China University of Science and Technology
Shanghai, 200237, China
Email: dhuang@ecust.edu.cn
Yujuan Chen
College of Mechanical and Electronic Engineering, Shanghai Normal University
Shanghai, 201418, China
Email: chenyjfq@163.com

ABSTRACT

This paper mathematically described the process and function of the knowledge transfer between the Virtual Enterprises (VEs). The VE profit distribution model based on the knowledge transfer and knowledge innovation was presented according to the principal-agent theory. In order to quantify the actual value of the parameters in the profit distribution model, the Third Supervisory Organization (TSO) acting as level-2 client can be entrusted to perform dynamic supervision to the partner enterprises. The TSO can make comprehensive evaluation for the partners' quality; schedule, cost, and knowledge transfer level. The supervisory work of the TSO is proven to be effective and feasible with the proposed bi-level principal-agent model. The effective supervision and quantitative evaluation of TSO can make VE profit distribution more reasonable.

Keywords: Knowledge Transfer, Knowledge Innovation, Virtual Enterprises, Profit Distribution, Bi-level Principal-Agent Theory, Third Supervisory Organization.

1. INTRODUCTION

The intention of enterprises constructing Virtual Enterprises (VEs) is to optimize their own resources and core abilities, and to fully hold the rapid-changing market chances. Knowledge is the most important strategic embodiment of core abilities. With the further development of economic globalization, the competition between VEs would be the competition of knowledge and technical resources. Therefore, the problem about knowledge sharing and knowledge transfer has already attracted great attentions of many scholars [1,2,3,4]. Many studies about knowledge are demonstrated from empirical examples, and found that the ability to acquire or transfer knowledge, the motive to learn or teach knowledge is the most important factors to successfully transfer knowledge [4,5]. However, it is also quite meaningful to make the quantitative research about knowledge transfer and knowledge innovation.

This paper gives a mathematic description for the knowledge transfer process, and proposes the profit distribution model based on principal-agent theory [6,7,8,9]. The Third Supervisory Organization (TSO) acting as level 2 clients strengthens the dynamic supervision for the partner enterprises, periodically performing comprehensive check and evaluation about the quality, schedule, cost, and the extent of knowledge sharing. This measure can quantify the extent of knowledge transfer and knowledge innovation. The evaluation result would be fully embodied in the profit distribution mechanism. The profit distribution would become much more reasonable.

* Partially supported by the Science &Research Foundation of Shanghai Educational Committee (No. 05DZ33).

2. KNOWLEDGE TRANSFER AND INNOVATION

Knowledge transfer and knowledge innovation are associated with many factors, mainly including the actual knowledge ability and the willingness to share knowledge for the knowledge provider, as well as the learning ability and learning intent of the knowledge learner.

Knowledge teaching and learning are mutual. Suppose (S_i, S_j) and (h_i, h_j) are the knowledge transfer effort and transfer ability coefficient of knowledge provider i and j respectively. g_i and g_j are the knowledge-learning ability coefficient of knowledge learner i and j respectively. Therefore, the final acquired knowledge of enterprise j after learning from enterprise i is $g_j h_i S_i$ (depicted in Fig.1).

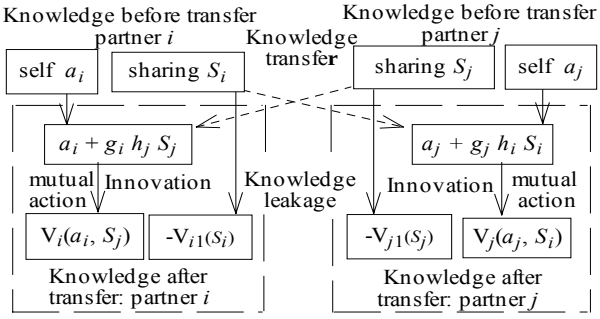


Fig. 1. VE knowledge transfer and knowledge innovation

The knowledge effort of enterprise j is composed of two parts, individual knowledge effort a_j used for its own production, and the knowledge transfer effort S_j trying to transfer knowledge to enterprise i. Therefore, the final knowledge value of enterprise i would become ($a_i+g_i h_j S_j$)

after knowledge transfer from enterprise *j*. This process can be illustrated with Fig. 1.

The acquired knowledge $g_i h_j S_j$ from knowledge transfer should be useful to the virtual enterprise, and would have mutual positive action on the enterprise *i*'s own knowledge a_i. This mutual positive action may generate new value $V_i(a_i,S_j)$ by the innovative effort of enterprise *i*. On the other hand, knowledge transfer may cause certain extent of core capacity leakage that may reduce the enterprise's competition ability, and extra profits may be taken by another enterprise. This kind of negative effect can be denoted with knowledge sharing cost $-V_{i1}(S_i)$.

3. VE PROFIT DISTRIBUTION MODEL

Suppose the virtual enterprise includes the core enterprise A and the partner enterprise B. The variables (a_1,a_2), (S_1,S_2) and (V_1,V_2) are individual knowledge effort, the knowledge transfer effort and knowledge innovation effort of enterprise A and B respectively. The parameters (h_1,h_2), (g_1,g_2), (e_1,e_2), (m_1,m_2), (b_1,b_2), (r_1,r_2) and (k_1,k_2), all within the scope between 0 and 1, are respectively knowledge transfer ability coefficient, knowledge learning ability coefficient, individual effort contribution coefficient, knowledge innovation contribution coefficient, individual effort cost coefficient, knowledge transfer cost coefficient, and knowledge innovation cost coefficient of enterprise A and B. Suppose $(0.5b_1a_1^2,\ 0.5b_2a_2^2)$ and $(0.5r_1S_1^2,\ 0.5r_2S_2^2)$ are the individual effort cost and knowledge transfer cost of enterprise A and B respectively.

Define d_1 and d_2 are the knowledge irrelevant coefficients between the individual knowledge $(a_1,\ a_2)$ and the acquired knowledge $(g_1h_2S_2,\ g_2h_1S_1)$ of enterprise A and B respectively. The bigger d_1 is, the smaller the knowledge relevance between a_1 and $g_1h_2S_2$ is, and the bigger the knowledge innovation cost $0.5d_1k_1V_1^2$ is.

3.1 Profit distribution without knowledge transfer

Suppose λ is the profit distribution coefficient of partner enterprise B without knowledge transfer and knowledge innovation. Enterprise A is risk-neutral. Partner enterprise B is risk-averse, and its risk cost is $0.5\rho_B\lambda^2\sigma_B^2$ under risk-averse coefficient ρ_B. T is the fixed transfer payment of enterprise B from enterprise A. T_0 is the minimal revenue of enterprise B. The VE profit distribution model without knowledge transfer is as follow.

$$\max_{\lambda,a_1}\pi_{A1}=(1-\lambda)(e_1a_1+e_2a_2)-\frac{1}{2}b_1a_1^2-T$$

s.t.

$$(IR):\pi_{B1}=\lambda(e_1a_1+e_2a_2)-\frac{1}{2}b_2a_2^2-\frac{1}{2}\lambda^2\rho_B\sigma_B^2+T\ge T_0$$

$$(IC):\max_{a_2}(\lambda(e_1a_1+e_2a_2)-\frac{1}{2}b_2a_2^2-\frac{1}{2}\lambda^2\rho_B\sigma_B^2+T). \tag{1}$$

According to the first order condition of a_1 and a_2 from Eq.(1), the optimal effort of enterprise A and B can be derived under the condition of maximizing their own individual profit.

$$a_1^*=\frac{e_1}{b_1}(1-\lambda),\ a_2^*=\frac{e_2}{b_2}\lambda \tag{2}$$

$$\frac{\partial a_2^*}{\partial e_2}=\frac{\lambda}{b_2}\ge 0,\ \frac{\partial a_2^*}{\partial b_2}=-\frac{e_2\lambda}{b_2^2}\le 0,\ \frac{\partial a_2^*}{\partial\lambda}=\frac{e_2}{b_2}\ge 0$$

Conclusion 1. The optimal individual effort a_2^* increases with the increase of individual knowledge contribution coefficient e_2, and decreases with the increase of individual effort cost b_2. The bigger the profit distribution coefficient λ is, the higher the individual effort extent of enterprise B is.

The VE total revenue is $\pi_{01}=\pi_{A1}+\pi_{B1}$.

$$\pi_{01}=e_1a_1+e_2a_2-\frac{1}{2}b_1a_1^2-\frac{1}{2}b_2a_2^2-\frac{1}{2}\lambda^2\rho_B\sigma_B^2. \tag{3}$$

The Eq. (2) is inputted into Eq. (3). Then the optimal profit distribution coefficient λ can be derived from the first order condition of λ by Eq.(3).

$$\lambda^*=(\frac{e_2^2}{b_2})/(\frac{e_1^2}{b_1}+\frac{e_2^2}{b_2}+\rho_B\sigma_B^2) \tag{4}$$

$$\frac{\partial\lambda^*}{\partial e_2}\ge 0,\ \frac{\partial\lambda^*}{\partial b_2}\le 0,\ \frac{\partial\lambda^*}{\partial\rho_B}\le 0,\ \frac{\partial\lambda^*}{\partial\sigma_B^2}\le 0$$

Conclusion 2. The optimal profit distribution coefficient λ^* increases with the increase of individual effort contribution coefficient e_2 of enterprise B, and decreases with the increase of individual effort cost b_2. The higher the risk-averse extent of partner enterprise B is, the smaller the profit distribution coefficient λ^* is.

The VE total revenue under the condition of maximizing individual profit can be derived from inputting Eq.(2) and Eq.(4) into Eq.(3).

$$\pi_{01}^*=\frac{e_1^2}{2b_1}-\frac{1}{2}(\frac{e_2^2}{b_2})^2/(\frac{e_1^2}{b_1}+\frac{e_2^2}{b_2}+\rho_B\sigma_B^2) \tag{5}$$

If enterprise A and B can cooperate with each other under the condition of maximizing collective profit, we can get the VE total profit π_{01}.

$$\pi_{01}=e_1a_1+e_2a_2-\frac{1}{2}b_1a_1^2-\frac{1}{2}b_2a_2^2 \tag{6}$$

The optimal individual effort can be obtained from the first order condition of a_1 and a_2 by Eq.(6).

$$a_1^{**}=\frac{e_1}{b_1},\quad a_2^{**}=\frac{e_2}{b_2}$$

After inputting a_1^{**} and a_2^{**} into Eq.(6), the VE total revenue under the condition of maximizing collective profit is as follow.

$$\pi_{01}^{**}=\frac{1}{2}(\frac{e_1^2}{b_1}+\frac{e_2^2}{b_2}) \tag{7}$$

The increased revenue between the revenue maximizing collective profit and the revenue maximizing individual profit is as follow.

$$\pi_{01}^{**}-\pi_{01}^*=\frac{e_2^2}{2b_2}(\frac{e_1^2}{b_1}+\rho_B\sigma_B^2)/(\frac{e_1^2}{b_1}+\frac{e_2^2}{b_2}+\rho_B\sigma_B^2). \tag{8}$$

3.2 Profit distribution based on knowledge transfer

Considering the factor of core ability leakage, both the core enterprise A and the partner enterprise B are risk-averse. The risk-averse coefficient and covariance of enterprise A and B are (ρ_1, $\sigma 12$) and (ρ_2, $\sigma 22$) respectively. β is the profit distribution coefficient of partner enterprise B only considering the factor of knowledge transfer and innovation. The VE profit distribution model based on knowledge transfer and innovation is as follow on the basis of Eq.(1).

$$\max_{\lambda,T} \pi_A = (1-\lambda)(e_1a_1 + e_2a_2) - \frac{1}{2}b_1a_1^2 - T + (1-\beta)(g_2h_1S_1 + g_1h_2S_2 + m_1V_1 + m_2V_2) - \frac{1}{2}r_1S_1^2 - \frac{1}{2}d_1k_1V_1^2 - \frac{1}{2}(1-\beta)^2\rho_1\sigma_1^2 \quad (9)$$

$$(IR)\ \pi_B = \lambda(e_1a_1 + e_2a_2) - \frac{1}{2}b_2a_2^2 - \frac{1}{2}\lambda^2\rho_B\sigma_B^2 + \beta(g_2h_1S_1 + g_1h_2S_2 + m_1V_1 + m_2V_2) - \frac{1}{2}r_2S_2^2 - \frac{1}{2}d_2k_2V_2^2 - \frac{1}{2}\beta^2\rho_2\sigma_2^2 + T \geq T_0. \quad (10)$$

The optimal efforts from the standpoint of maximizing the individual profit can be derived from the relevant first order condition from Eq.(9) and Eq.(10).

$$a_1^* = \frac{e_1}{b_1}(1-\lambda),\ S_1^* = \frac{h_1g_2}{r_1}(1-\beta),\ S_2^* = \frac{h_2g_1}{r_2}\beta,$$
$$a_2^* = \frac{e_2}{b_2}\lambda,\ V_1^* = \frac{m_1}{d_1k_1}(1-\beta),\ V_2^* = \frac{m_2}{d_2k_2}\beta. \quad (11)$$
$$\frac{\partial S_2^*}{\partial(h_2g_1)} \geq 0,\ \frac{\partial S_2^*}{\partial r_2} \leq 0,\ \frac{\partial S_2^*}{\partial \beta} \geq 0,\ \frac{\partial V_2^*}{\partial m_2} \geq 0,$$
$$\frac{\partial V_2^*}{\partial(d_2k_2)} \leq 0,\ \frac{\partial V_2^*}{\partial \beta} \geq 0,$$

Conclusion 3. Both the knowledge transfer effort S_2 and the knowledge innovation effort V_2 of enterprise B increase with the increase of the knowledge transfer-based profit distribution coefficient β. The knowledge transfer effort S_2 decreases with the increase of knowledge transfer cost r_2. The knowledge innovation effort V_2 decreases with the increase of knowledge innovation cost k_2. The bigger the knowledge irrelevant coefficient d_2 is, the smaller the enthusiasm of enterprise B engaging knowledge innovation is. The knowledge transfer effort S_2 of enterprise B is proportional to the knowledge transfer ability h_2 of enterprise B, and proportional to the knowledge learning ability g_1 of enterprise A.

The increased revenue generated from knowledge transfer and knowledge innovation is $\pi_{0\beta}$.

$$\pi_{0\beta} = [g_2h_1S_1 + g_1h_2S_2 + m_1V_1 + m_2V_2 - \frac{1}{2}r_1S_1^2 - \frac{1}{2}d_1k_1V_1^2 - \frac{1}{2}r_2S_2^2 - \frac{1}{2}d_2k_2V_2^2 - \frac{1}{2}(1-\beta)^2\rho_1\sigma_1^2 - \frac{1}{2}\beta^2\rho_2\sigma_2^2]$$

The VE total revenue based on knowledge transfer from the standpoint of maximizing the individual profit is the π_{01} in Eq.(3) plus the $\pi_{0\beta}$.

$$\pi_{02} = \pi_{01} + \pi_{0\beta} = [(e_1a_1 + e_2a_2) - \frac{1}{2}b_1a_1^2 - \frac{1}{2}b_2a_2^2 - \frac{1}{2}\lambda^2\rho_B\sigma_B^2] + [g_2h_1S_1 + g_1h_2S_2 + m_1V_1 + m_2V_2 - \frac{1}{2}r_1S_1^2 - \frac{1}{2}d_1k_1V_1^2 - \frac{1}{2}r_2S_2^2 - \frac{1}{2}d_2k_2V_2^2 - \frac{1}{2}(1-\beta)^2\rho_1\sigma_1^2 - \frac{1}{2}\beta^2\rho_2\sigma_2^2]. \quad (12)$$

After inputting Eq.(11) into Eq.(12), we can get the knowledge transfer based optimal profit distribution coefficient β^* under the condition of maximizing individual profit from the first order condition of β by Eq.(12).

$$\beta^* = \frac{\frac{(h_2g_1)^2}{r_2} + \frac{m_2^2}{d_2k_2} + \rho_1\sigma_1^2}{\frac{(h_1g_2)^2}{r_1} + \frac{(h_2g_1)^2}{r_2} + \frac{m_1^2}{d_1k_1} + \frac{m_2^2}{d_2k_2} + \rho_1\sigma_1^2 + \rho_2\sigma_2^2}$$
$$\frac{\partial\beta^*}{\partial(h_2g_1)} \geq 0,\ \frac{\partial\beta^*}{\partial m_2} \geq 0,\ \frac{\partial\beta^*}{\partial(d_2k_2)} \leq 0,\ \frac{\partial\beta^*}{\partial r_2} \leq 0,\ \frac{\partial\beta^*}{\partial\rho_2} \leq 0$$

Conclusion 4. The knowledge transfer-based profit distribution coefficient β^* increases with the increase of knowledge innovation contribution coefficient m_2, individual knowledge transfer ability h_2 of enterprise B, and knowledge learning ability g_1 of enterprise A, and decreases with the increase of knowledge innovation cost k_2. The bigger the knowledge irrelevant coefficient d_2 is, the bigger the knowledge innovation cost is, the smaller the profit distribution coefficient β^* of enterprise B is.

The increased revenue from the knowledge transfer under the condition of maximizing individual profit is $\pi_{0\beta}^*$.

$$\pi_{0\beta}^* = \frac{\frac{1}{2}(\frac{(h_1g_2)^2}{r_1} + \frac{m_1^2}{d_1k_1} - \rho_1\sigma_1^2) + \frac{1}{2}(\frac{(h_2g_1)^2}{r_2} + \frac{m_2^2}{d_2k_2} + \rho_1\sigma_1^2)^2}{\frac{(h_1g_2)^2}{r_1} + \frac{(h_2g_1)^2}{r_2} + \frac{m_1^2}{d_1k_1} + \frac{m_2^2}{d_2k_2} + \rho_1\sigma_1^2 + \rho_2\sigma_2^2} \quad (13)$$

The VE total revenue without knowledge transfer is π_{01}^* in Eq.(5). The VE total revenue π_{02}^* based on knowledge transfer under the condition of maximizing individual profit is π_{01}^* plus $\pi_{0\beta}^*$.

$$\pi_{02}^* = \pi_{01}^* + \pi_{0\beta}^* = \frac{e_1^2}{2b_1} - \frac{1}{2}(\frac{e_2^2}{b_2})^2/(\frac{e_1^2}{b_1} + \frac{e_2^2}{b_2} + \rho_B\sigma_B^2) + \frac{1}{2}(\frac{(h_1g_2)^2}{r_1} + \frac{m_1^2}{d_1k_1} - \rho_1\sigma_1^2) + \frac{\frac{1}{2}(\frac{(h_2g_1)^2}{r_2} + \frac{m_2^2}{d_2k_2} + \rho_1\sigma_1^2)^2}{\frac{(h_1g_2)^2}{r_1} + \frac{(h_2g_1)^2}{r_2} + \frac{m_1^2}{d_1k_1} + \frac{m_2^2}{d_2k_2} + \rho_1\sigma_1^2 + \rho_2\sigma_2^2} \quad (14)$$

Consider the condition of maximizing collective profit of VE, The VE total revenue π_{02} can be formulated as follow.

$$\pi_{02} = [(e_1a_1 + e_2a_2) - \frac{1}{2}b_1a_1^2 - \frac{1}{2}b_2a_2^2] + [g_2h_1S_1 + g_1h_2S_2 + m_1V_1 + m_2V_2 - \frac{1}{2}r_1S_1^2 - \frac{1}{2}r_2S_2^2 - \frac{1}{2}d_1k_1V_1^2 - \frac{1}{2}d_2k_2V_2^2]. \quad (15)$$

The relative optimal efforts under the condition of maximizing collective profit can be derived from the first order condition by Eq.(15).

$$a_1^* = \frac{e_1}{b_1},\ a_2^* = \frac{e_2}{b_2},\ S_1^* = \frac{h_1g_1}{r_1},\ S_2^* = \frac{h_2g_2}{r_2},$$
$$V_1^* = \frac{m_1}{d_1k_1},\ V_2^* = \frac{m_2}{d_2k_2}. \quad (16)$$

Under the condition of maximizing the collective profit, the increased revenue $\pi_{0\beta}^{**}$ from the knowledge transfer and knowledge innovation can be obtained by inputting Eq.(16) into the latter part of Eq.(15).

$$\pi_{0\beta}^{**} = \frac{1}{2}\left(\frac{(h_1 g_2)^2}{r_1} + \frac{(h_2 g_1)^2}{r_2} + \frac{m_1^2}{d_1 k_1} + \frac{m_2^2}{d_2 k_2}\right) \quad (17)$$

The VE total revenue π_{02}^{**} based on knowledge transfer from collective profit standpoint is π_{01}^{**} in Eq.(7) plus $\pi_{0\beta}^{**}$ in Eq.(17).

$$\pi_{02}^{**} = \pi_{01}^{**} + \pi_{0\beta}^{**} = \frac{1}{2}\left(\frac{e_1^2}{b_1} + \frac{e_2^2}{b_2}\right) + \frac{1}{2}\left(\frac{(h_1 g_2)^2}{r_1} + \frac{(h_2 g_1)^2}{r_2} + \frac{m_1^2}{d_1 k_1} + \frac{m_2^2}{d_2 k_2}\right). \quad (18)$$

Comparing the Eq.(13) and Eq.(17), the increased VE revenue generated from knowledge transfer and knowledge innovation between maximizing collective profit and maximizing individual profit is $\pi_{0\beta}^{**}$ minus $\pi_{0\beta}^{*}$.

$$\pi_{0\beta}^{**} - \pi_{0\beta}^{*} = \frac{1}{2}\frac{\left(\frac{(h_1 g_2)^2}{r_2} + \frac{m_2^2}{d_2 k_2} + \rho_1\sigma_1^2\right)\left(\frac{(h_1 g_2)^2}{r_1} + \frac{m_1^2}{d_1 k_1} + \rho_2\sigma_2^2\right)}{\frac{(h_1 g_2)^2}{r_1} + \frac{(h_2 g_1)^2}{r_2} + \frac{m_1^2}{d_1 k_1} + \frac{m_2^2}{d_2 k_2} + \rho_1\sigma_1^2 + \rho_2\sigma_2^2} \quad (19)$$

Comparing the Eq.(14) and Eq.(18), the VE total increased revenue based on the knowledge transfer and knowledge innovation from the individual standpoint to the collective standpoint is as follow.

$$\pi_{02}^{**} - \pi_{02}^{*} = (\pi_{01}^{**} + \pi_{0\beta}^{**}) - (\pi_{01}^{*} + \pi_{0\beta}^{*}) = (\pi_{01}^{**} - \pi_{01}^{*}) + (\pi_{0\beta}^{**} - \pi_{0\beta}^{*}) \geq 0.$$

Conclusion 5. The VE total revenue from the standpoint of maximizing the collective profit is always more than the VE total revenue from the standpoint of maximizing the individual profit. The Knowledge transfer and knowledge innovation can create more revenue for the virtual enterprise.

4. DYNAMIC SUPERVISION MECHANISM

The interests of core enterprise may be impaired due to the moral hazard problem of partner enterprises. In order to enhance the effectiveness of dynamic supervision, the core enterprise can entrust the Third Supervision Organization (TSO) with specialized knowledge and supervision experience to supervise the partner enterprises.

The core enterprise A need have a project coordination center based on the market chance or certain product (see Fig. 2). The project coordination center is in charge of the general coordination and control about quality, schedule, cost and contract for the TSO and all partner enterprises.

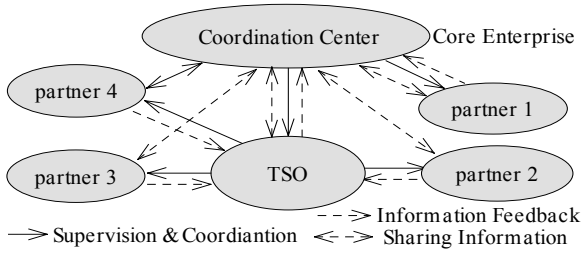


Fig. 2. Virtual enterprise organization based on TSO

The TSO accepts the entrust of core enterprise, and fully utilizes the sharing information and field acquired information to supervise and control the quality, schedule and cost of partial or all partner enterprises. The TSO also manages the revised contract about technology, quality, schedule and cost appearing during product design and manufacturing process.

5. SUPERVISION MECHANISM BASED ON THE BI-LEVEL PRINCIPAL-AGENT THEORY

Suppose the core enterprise A is level-1 client. The enterprise C acting as level-2 agent is the upstream Partner supplying materials for the core enterprise A. The third supervision organization B accepts the entrustment of core enterprise A to supervise the partner enterprise C. The TSO acts as both level-1 agent and level-2 client. All the participants A, B and C are risk-neutral.

The parameters (a_0, a_m, a_1), (e_0, e_m, e_1) and (b_0, b_m, b_1) are the individual efforts, individual effort contribution coefficients, and individual effort cost coefficients of level-1 client A, level-2 client B and level-2 agent C respectively. The parameter β is the profit distribution coefficient of level-2 agent C set by level-1 client A. λ is the profit-inspiriting coefficient of level-2 agent C set by level-2 client B. The level-2 client B makes dynamic supervision for level-2 agent C, and determines the profit inspiriting coefficient λ according to the comprehensive evaluation about the quality, schedule and cost of level-2 agent C. The bi-level principal-agent model based on the TSO is as follow.

$$\begin{aligned}
&\max_{\beta, a_0} \pi_A = (1-\beta)(e_0 a_0 + e_m a_m + e_1 a_1 - T_m) - T_1 - \frac{1}{2} b_0 a_0^2 \\
&\max_{\lambda, a_m} \pi_B = \beta(1-\lambda)(e_0 a_0 + e_m a_m + e_1 a_1 - T_m) + T_m - \frac{1}{2} b_m a_m^2 \\
&s.t.(IR): \pi_C = \beta\lambda(e_0 a_0 + e_m a_m + e_1 a_1 - T_m) - \frac{1}{2} b_1 a_1^2 + T_1 \geq T_0 \\
&(IC): \max_{a1} \pi_C = \beta\lambda(e_0 a_0 + e_m a_m + e_1 a_1 - T_m) - \frac{1}{2} b_1 a_1^2 + T.
\end{aligned} \quad (20)$$

Where, T is the fixed transfer payment of level-2 agent C from level-1 client A. T_0 is minimal profit of level-2 agent C. Level-2 client B doesn't directly participate the production activities. It tries to increase the VE total revenue by strengthening the supervision for level-2 agent C and improving the total efficiency of VE. The fixed minimal profit of level-2 client B would be proportionally apportioned by the level-1 client A and level-2 agent C.

The optimal efforts of all participants can be derived from the first order condition of a_0, a_m and a_1 by π_A, π_B and π_C in Eq.(20).

$$a_0^* = \frac{e_0}{b_0}(1-\beta),\ a_m^* = \frac{e_m}{b_m}\beta(1-\lambda),\ a_1^* = \frac{e_1}{b_1}\beta\lambda \quad (21)$$

$$\partial a_1^* / \partial e_1 \geq 0, \partial a_1^* / \partial \beta \geq 0,\ \partial a_1^* / \partial \lambda \geq 0,\ \partial a_1^* / \partial b_1 < 0$$

Conclusion 6. The optimal individual effort a_1^* of level-2 agent C would increase with the increase of its own contribution coefficient e_1, profit distribution coefficient β set by level-1 client A and the profit inspiriting coefficient λ set by level-2 client B, and decrease with the increase of its own individual effort cost coefficient b_1.

The equivalent revenue π_{BC} of level-2 client B and level-2 agent C is as follow.

$$\pi_{BC} = \pi_B + \pi_C = \beta(e_0 a_0 + e_m a_m + e_1 a_1 - T_m)$$

$$+T_m-\frac{1}{2}b_m a_m^2-\frac{1}{2}b_1 a_1^2+T_1 \qquad (22)$$

After inputting Eq.(21) into Eq.(22), we can get the optimal profit inspiriting coefficient λ^* from the first order condition of λ by Eq.(22).

$$\lambda^*=\frac{e_1^2/b_1}{(e_1^2/b_1)+(e_m^2/b_m)} \qquad (23)$$

$$\partial\lambda^*/\partial e_1\geq 0,\ \partial\lambda^*/\partial b_1\leq 0,\ \partial\lambda^*/\partial e_m\leq 0$$

Conclusion 7. The profit inspiriting coefficient λ^* set by level-2 client B increases with the increase of individual effort e_1 of level-2 agent C. The λ^* decreases with the increase of effort cost b_1 of level-2 agent C and with the increase of the supervision extent e_m of level-2 client B. it is determined by level-2 client B, and has nothing with level-1 client A.

The equivalent revenue of level-1 client A, level-2 client B and level-2 agent C is as follow.

$\pi_{ABC}=\pi_A+\pi_B+\pi_C=$

$$(e_0a_0+e_ma_m+e_1a_1)-\frac{1}{2}b_0a_0^2-\frac{1}{2}b_ma_m^2-\frac{1}{2}b_1a_1^2. \qquad (24)$$

After inputting Eq.(21) into Eq.(24), we can get the optimal profit distribution coefficient β^* from the first order condition of β by Eq.(24).

$$\beta^*=\frac{(\frac{e_1^2}{b_1})^2+(\frac{e_m^2}{b_m})^2(\frac{e_1^2}{b_1}+\frac{e_m^2}{b_m})}{\frac{e_0^2}{b_0}(\frac{e_1^2}{b_1}+\frac{e_m^2}{b_m})^2+(\frac{e_1^2}{b_1})^3+(\frac{e_m^2}{b_m})^3} \qquad (25)$$

We can find from Eq.(25) that the profit distribution coefficient β^* mainly decreases with the increase of the individual effort e_0 of level-1 client A, and increases with the increase of individual effort cost b_0 of level-1 client A.

Comparing the Eq.(23) and Eq.(25), the sensitivity of β^* changing with the individual effort contribution coefficients (e_m, e_1) and effort costs (b_m, b_1) of level-2 client B and level-2 agent C is distinctively smaller than the sensitivity of λ^* changing with these factors.

Conclusion 8. The profit distribution coefficient β^* of level-2 agent C is mainly determined by level-1 client A. level-2 agent C is much more sensitive to the profit inspiriting coefficient λ^* set by level-2 client B. The inspiriting intensity of λ^* to level-2 agent C is stronger than the inspiriting intensity of β^*.

From the above conclusions, we know the mechanism that the core enterprise A takes the third supervision organization B as the level-2 client to dynamically supervise the partner enterprise C should be feasible and effective.

The TSO acting as level-2 client proceeds multi-stage dynamic supervision for the partner enterprises, and makes comprehensive evaluation about the quality, cost, schedule, technology innovation, knowledge transfer extent, and cooperative attitude of the partner enterprises. The evaluation result is the important evidence to determine the profit inspiriting coefficient λ^*. This measure can effectively urge the partner enterprises to increase the internal self-supervision extent and the external cooperative enthusiasm.

6. CONCLUSIONS

Core capacity and knowledge are the most basic and most important factors to construct the virtual enterprise. This paper mathematically describes knowledge transfer process between the different enterprises. The VE increased revenue from knowledge transfer and knowledge innovation is quantitatively analyzed based on the principal-agent theory. The proposed dynamic supervision mechanism based on bi-level principal-agent theory can effectively supervise the work of partner enterprises, and makes the determination of profit distribution coefficient much more concrete, quantitative and reasonable. The reasonable profit distribution mechanism can drive the virtual enterprise running healthily and profitably.

7. REFERENCES

[1] C.U. Ciborra, A. Rafael, “Sharing Knowledge Across Boundaries”, Journal of Information Technology, 2001, 16(2): 73-81.

[2] P.H. Gray, D.B. Meister, “Knowledge Sourcing Effectiveness”, Management Science, 2004, 50(6): 821-834.

[3] L. Li. “Information Sharing in a Supply Chain with Horizontal Competition”, Management Science, 2002, 48(9): 1196-1212.

[4] P. Wang, T.W. Tong, et al “An Integrated Model of Knowledge Transfer from MNC Parent to China Subsidiary”, Journal of World Business, 2004, 39(2): 168-172.

[5] R.S. Marshall, T.V. Nguyen, et al “A Dynamic Model of Trust Development and Knowledge Sharing in Strategic Alliances”, Journal of General Management, 2005, 31(1): 41-57.

[6] G.M. Hou, C.J. Li, Modern Management Inspiriting and Constraint Mechanism, Beijing: Higher Education Press, 2002.

[7] J.J. Laffont, D. Martiort, The Theory of Incentives Ⅰ: The Principal-Agent Model, Beijing: Renmin University Press of China, 2002.

[8] I. Giannoccaro, P. Pontrandolfo, “Supply Chain Coordination by Revenue Sharing Contracts”, International Journal of Production Economics, 2004, 89(2): 131-139.

[9] Y.Z. Wang, L. Jiang, et al , “Channel Performance Under Consignment Contract with Revenue Sharing”, Management Science, 2004, 50(1): 34-47.

Agile Supply Chain Management Base on Web Services

Guangchao Wu, Shu Yu
School of Mathematical Sciences, South China University of Technology
Guangzhou 510640, China
Email: Brisk_su@hotmail.com Yushu_scut@tom.com

ABSTRACT

By analyzing of the weakness of present supply chain technologies and the strength of Web services, this paper discusses how to make full use of those legacy systems and proposed the architecture of inner supply chain management system and a agile supply chain cooperation model by incorporating the SOA (Serviced Oriented Architecture) model of Web services. We also used a purchase process to detailedly demonstrate the interoperation among different enterprises in supply chain based on Web services. The main objective is to improve the agility of supply chain and enterprises' competitiveness.

Keywords: Supply Chain Management, Agile Supply Chain, Web Services, Service Oriented Architecture

1. INTRODUCTION

Agile supply chain (ASC), composed of suppliers, demander (autonomous, semi-autonomous or subordinate), and is a dynamic supply/demand network aiming to quickly response to the demand changes in the market. It comes from the competitive, cooperative and dynamic market environment and can meet the requirements of the increasing intense competition in global market [1, 2].

By combining enterprise agility with supply chain management (SCM), ASC differs from normal SCM. It can quickly be reconstructed and integrated according to the formation or break-up of dynamic alliance and use SCM to promote the association among enterprises, which can finally improve the agility of enterprises. However, how to use present computer technologies to build an integrated, automated and agile application system for ASC becomes an urgent problem.

The development of Web services provides new support tool and technology for agile supply chain management (ASCM). It enables different application systems in each enterprise to interoperate and share their information and resources easily. So, the major objective of this paper is to discuss how to combine Web services technology with ASC to build a technology support platform for networked, collaborative and ASCM.

2. RELATED WORK

The dynamic characteristics of ASCM cause the relationship among all relevant entities to change frequently and require that they should be able to response to the constantly changing environment. However, most of the applications running in the business area are made up of tight coupling modules, which can hardly meet the requirements of ASC.

Traditionally, major technologies used to support SCM systems include EDI (Electronic Data Interchange), DCOM (Distributed Component Object Modular) and CORBA (Common Object Request Broker Architecture) /IIOP (Internet Inter ORB Protocol), etc.

EDI systems are built on the closed VAN (Value Added Network) with the restriction that both parts of the communication should be EDI users. Additionally, EDI messages use fixed business sets and must be processed by special software, which make them very complex and lack of flexibility. As a result, most of the small and medium-size enterprises can't afford the expensive cost of building VAN.

Many distributed object technologies (like CORBA, DCOM) rely on binary communication protocols, which often encounter the firewall problems later during the communication process. The interoperation among them also faces the problems of low efficiency and incompatibility.

At present, a lot of research papers in China mainly focus on using CORBA as the technology to support supply chain integration and discuss how to share the information of ASC and integrate original applications [3]. They use methodologies such as united enterprise modeling or united information modeling centering on the leader of alliance.

Recently, some scholars begin to discuss the application of Web services in ASC. The author in [4] proposes an ASC integration framework based on Web services, but he didn't mention how to make full use of current legacy systems.

3. WEB SERVICES

Web services are one of the most important technologies for distributed applications. It include a series of open standard protocol such as XML (extensible Markup Language), SOAP [5] (Simple Object Access Protocol), WSDL [6] (Web Services Description Language) and UDDI [7] (Universal Description, Discovery and Integration), etc.

Web services are built upon SOA (Service Oriented Architecture) model, which is made up of three actors and three basic operations. The three actors are Service Provider, Service Requester and Service Registry, and the three operations are Publish, Find and Bind. The SOA model is shown in Fig. 1.

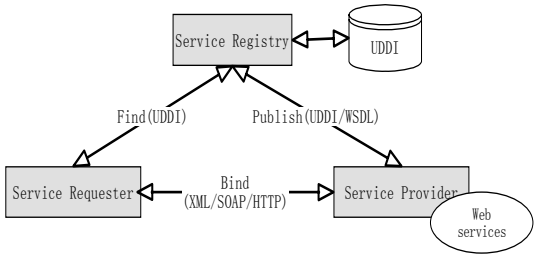


Fig. 1. SOA Model

In the SOA model enterprises can develop their own businesses in the form of Web services and share them as

component. After the core business processes are implemented as Web services, information can be easily shared among different enterprises so that cross-platform ASC will be set up.

Nowadays, Web services are gradually applied in many application areas. Companies such as Imperial Sugar, Nordstrom.com and Hewitt &Associates even used it in SCM and got satisfied results.

4. ASCM BASED ON WEB SERVICES

4.1. Architecture of Inner Supply Chain Based on Web Services

Due to the heated competition, organization models and business processes inside enterprises need to be adjusted frequently to adapt to the quickly changing market requirements. Every department of each separate enterprise in supply chain such as sale, plan, production, purchase or stock also exists supply chain relationship, which is called inner supply chain. It is the foundation for global ASC.

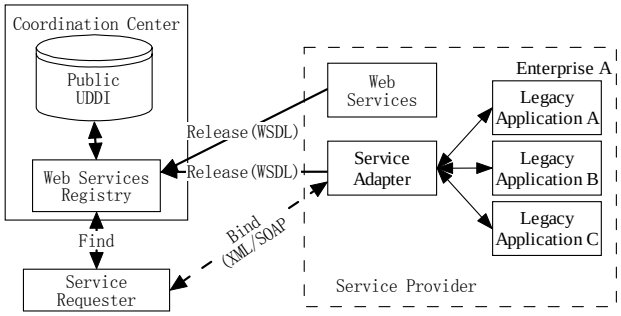


Fig. 2. Architecture of Inner Supply Chain Based on Web Services

In Fig. 2, we propose the architecture of inner supply chain based on Web services. Coordination center is the most important part of supply chains, whose main functions are as follows:

1. Deciding the services of each member enterprise in supply chain.

For example, it can use UML (United Model Language) to define and describe the management activities of different enterprises and decide the contents of Web services provided by them.

2. Recording and querying Web services.

It contains a public UDDI, allowing users to publish and query Web services.

3. Providing security service and access control.

Security is of great importance to supply chain. There are many security technologies such as PKI (Public Key Infrastructure), SAML (Security Assertion Markup Language, EACM (Extensible Access Control Markup) for Web services. How to integrate them to build a secure and dependent system for supply chain still remains open.

4. Evaluating and publishing the reputation of its members.

The reputation of an enterprise becomes crucial in modern society and coordination center should provide such QoS (Quality of Service) information about those Web services.

For most of the enterprises, legacy applications still play a very important role and it is impossible to discard them immediately. In order to make full of them and enable them to be accessed and shared in the form of Web services, some Web services adapters should be developed. They are usually the programs that link to applications in the backend server. For each SOAP service request, adapters invoke a backend application and the outputs are constructed as a SOAP response and sent back to the requester. By using these adapters, the complexity of calling backend functions can be hidden form users and the requesters just provide the necessary parameter sets needed for executing the services.

Web services adapters make it feasible for those applications such as CRM (Client Relationship Management) and ERP (Enterprise Resource Plan) to be organized flexibly and shared in the Internet. The concrete operation processes are as follows:

1、 Enterprises use Web services adapters to wrap applications such as CRM and ERP as Web services and registry them in public UDDI.
2、 Service requester browses the public UDDI registry and acquires the needed information about those business entities.
3、 After choosing a service, the service requester will get the service's location and its WSDL binding information.
4、 Then, the requester creates the SOAP request according to the WSDL description and sends it to the service provider to bind the service.
5、 After receiving the request, Web service adapters transfer the parameters to those corresponding CRM or ERP applications in the backend servers.
6、 Finally, the adapters build the output into a SOAP response and send it back to the requester.

4.2. Cooperation Model Based on Web Services

ASC is a large dynamic network made up of clients, producers and suppliers. Its business processes often involve many enterprises, each of which takes charge of different parts, and need to cross the boundary of a single enterprise. To finish these processes efficiently, information exchanges and cooperation step by step according to the flow are of great importance. The characteristics of cross-enterprise and loose coupling relationships of business processes demand that the SCM system should be distributed, and heterogeneous. So, an effective ASC system should avoid the traditional tight coupling pattern and have the ability of high flexibility and customization.

As the major business process in supply chain, purchase and sale process is cross-enterprise and its participants contain buyers and sellers. To detailedly demonstrate the interoperation among enterprises in supply chain, we take the order creation request as an example. In a simple purchase process, as what is shown in Fig. 3, the client first uses product lookup service to find the needed services and then offers the necessary information to create an order. When creating the order, the purchase service will also invoke payment service and delivery service.

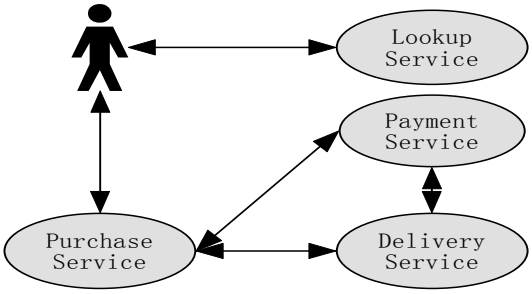


Fig. 3. A Simple Purchase Process

Of course, practical purchase process won't be so simple as what is mentioned above, but Fig. 3 can basically demonstrate the cooperation model among those business Web services.

Supposed the client, broker and producer constitute a supply chain. Firstly, the client sends an order to the broker. After receiving the order, the broker will invoke an order dealing process. To finish it, the broker also needs the services provided by producers. So in the order dealing process, the services from producer are also called, which finally constitutes an integrate business process. It is similar with a chain workflow interoperation model, in which the interactions among enterprises become the cooperation among different Web services inside the enterprises. (A cooperation model based on Web services is shown in Fig. 4.)

In this way, an enterprise can choose the best partners according to its own benefits and join different supply chains to achieve the greatest business opportunity and find the most suitable services.

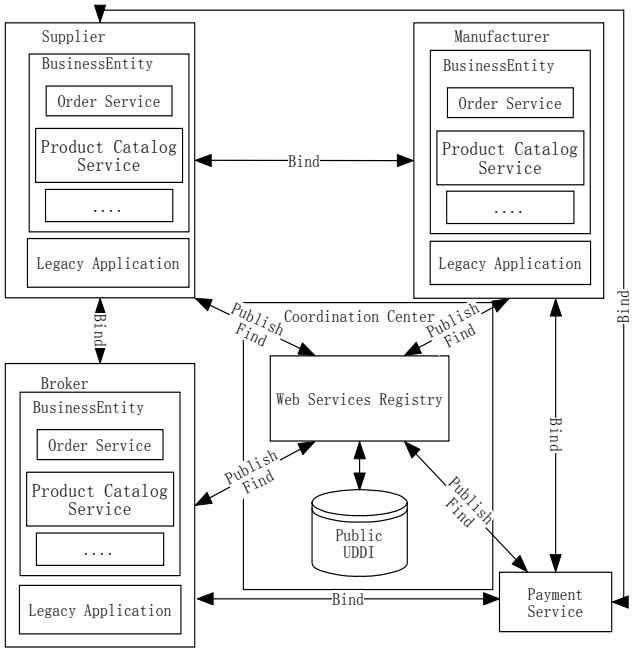


Fig. 4. Web services-based Cooperation Model

5. CONCLUSIONS

As a developing technology, Web services still exists some problems like workflow design and quality of service, but it is the most suitable technology for building ASC. So, in this paper, we discuss how to use it to improve the agility of supply chain and propose the architecture of inner supply chain and a cooperation model based on Web services.

In the future, we will focus on the following directions: First, we will use Web services to develop workflow management system in supply chain. Second, we will try to discuss how to add the QoS information about Web services. Third, we will further discuss the security and dependability of Web services in agile supply chain

6. REFERENCES

[1] C. Yueting, L. Yi, "Agile *supply chain management* [M]," Beijing: Tsinghua University Press, 2002.

[2] D.N. Chorafas, "Integrating ERP, CRM, supply chain management, and smart material [M]", Boston, MA, USA: AU ERBACH Publication Inc. , 2001.

[3] Z.H. Wang, Y.T. Chai, et al, "The application of CORBA in agile supply chain system [J]", Computer Engineering and Application, 2002, (1): 220-222.

[4] T. Weining, X. Fuyuan, "Integration technology research of SCM based on Web services [J]", Computer Applications and Software, 2005, 22(9): 139-141.

[5] D. Box, D. Ehnebuske, et al, Simple Object Access Protocol (SOAP) 1.2 Technical report, http://www.w3c.org/TR/SOAP12, 2003.

[6] UDDI. UDDI technical white paper. Technical report, http://www.uddi.org/, 2000.

[7] W3C. Web Services Description Language (WSDL) version 1.2. http://www.w3c.org/TR/wsdl20/, 2006.

The ECommerce Information model driven Semantic Searching algorithm

Yun Ling, Ouyang Yi, Biwei Li
College of Computer and Information Engineering, Zhejiang Gongshang University, Hangzhou 310035

Abstract

To make ECommerce information searching across Internet more efficient, ECommerce information searching becomes more and more important. In this paper, ECommerce Information Model (EIM) and a novel EIM-based semantic similarity algorithm are presented. This semantic similarity algorithm utilizes ECommerce-based information content and edge-based distance in calculating conceptual similarity. According to EIM, a semantic eigenvector, which consists of the semantic similarity values of a given document, is used to represent the semantic content of the document. The semantic eigenvectors and EIM-based similarity function could be applied to ECommerce information retrieval. Experimental results show that the performance of the proposed method is much improved when compared with that of the traditional Information retrieval techniques.

Keywords: Semantic Search, Similarity, ECommerce, Information content.

1. INTRODUCTION

The rapid development of Internet application has greatly increased the availability of the information on the Internet, making it difficult to carry out the first step towards ECommerce information retrieval. The problem how to efficiently retrieve valuable information from huge mounts of information remains to be solved. Traditional information retrieval depends on techniques such as contents, indexes, and keywords, with advantages of being simple and speedy, while the primary drawback is that those techniques could not mine the inner relationships among information and the retrieval results could not precisely and comprehensively meet user demands. Since there lacks a single semantic description for the entire internet information resources, users could not properly locate relevant contents and services [1]. The problem to realize intelligent business information retrieval relies on making the meaning of the internet information resources understandable by the retrieving system and facilitating the semantic integration of information resources.

Based on WordNet, Ray Richardson proposed the knowledge-based information retrieval method [2] by calculating the semantic distance between concepts, and the distance could be used to compute the similarity between the inquiring conditions and web pages. Though the precision and recall rates on information retrieval are not so idealistic, this method has provided a direction for semantic information retrieval. To make information retrieving more efficient and maximally realize metadata sharing and reusing, this paper proposes using EIM and EIM-based semantic similarity function to process semantic information comparison among words, making the system with a relatively better semantic information expression capability for the business domain. This similarity algorithm could be implemented in business information retrieving system, and experimental results show that this algorithm has better precision and recall rates for business information retrieving and effectuates semantic information service of business domain.

2. THE EIM DEFINITION

Of the conceptual hierarchy of the words in WordNet[3], the relationship between business information is not so specific, making it hard to reflect some related business concepts. For example, the word ‘camera’ and ‘photograph’ are positioned in the classification structures with far distance, namely low correlation from the aspect of semantic analysis; but from the aspect of business domain, those two words have strong ties. To realize business-oriented semantic information searching, this paper adopts EIM to process business information semantic retrieval. The primary difference between EIM and WordNet is that EIM enhances the description of relations among various business concepts and the definition of EIM focuses mainly on the definitions of conceptual entities and relevant relations.

Definition An EIM is a triple O=<C, R, S>, where C represents the set of ECommerce concepts; R represents the relations among concepts; S is the constraint, namely a two-tuple (C_S, R_S), where Cs represents the constraints on concepts and Rs represents the constraints on relations.

The structure of EIM contains concepts of ECommerce related to business information (totally 5821 concept nodes, in which 62 nodes are of First Layer Entity Nodes (FLEN), 212 nodes are of Second Layer Entity Nodes (SLEN) and 5489 are of Third Layer Entity Node (TLEN), five types of fundamental relations and three types of extended relations of ECommerce. The fundamental relations are defined as: is-a relation, part-of relation, associative relation, Synonymy relation and Meronym relation, whose definitions are the same as that of WordNet. To deal with the special features of ECommerce, three extended relations are added into business concept entities: Geo-location relation, Apply-range relation and Apply-object relation, where each concept object is defined by a synonym set.

To reduce the number of semantic eigenvectors in the retrieving process, light-weighted EIM structure is adopted, dividing the business information into three hierarchies. Hierarchy 0 is the common business information entities, such as Textile, Garment, Chemical, Mineral, Footwear, Headgear, Cases, Bag, Tools, Hardware, Vehicles, Household, etc. Hierarchy 1 is the entity set of detailed application of ECommerce, which connects each FLEN with several SLEN. Hierarchy 2 is the special cases of applications of SLEN, where each SLEN concept is divided into several hierarchy 2 TLEN according to detailed composition or application. The structure of EIM is showed in Fig. 1.

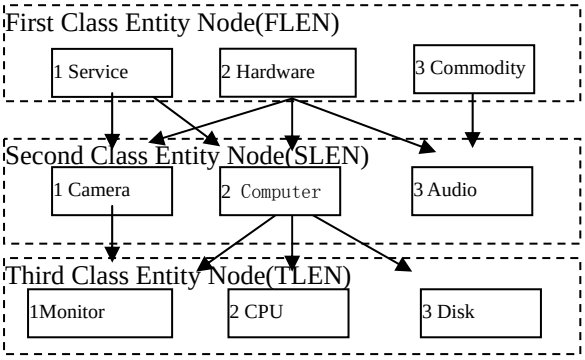


Fig. 1. The structure of EIM

3. COMPUTATION OF EIM-BASED SEMANTIC SIMILARITY

Semantic similarity [4] is proposed and applied to the research on detailed semantic relations. To compute concept similarity, there are many methods available. Similarity-based feature model proposed by Tversky [5] is the most widely used one. Since the introduction of WordNet, many researchers have conducted researches on word-concept relationship of the classification structure [6], [2]. Jiang has proposed a new semantic similarity measure method based on collected works and lexicon classification structure, which has integrated the edge-based similarity measure method and content based node similarity measure method. Dagan [7] has proposed a more complicated probability model to compute distance between words. Currently, the research on semantic similarity could be categorized into two types: information content-based and edge-based distance measure methods.

3.1. Information content-based similarity measure method

The method to determine similarity between concepts on the basis of nodes is called information content-based method, which tries to determine the information content of the concept by computing the frequency of the concept occurred in the documents. In hierarchical structure of concept, the appearance of current concept node in the documents means the definite appearance of its sub-concepts.

The method proposed by Resnidk [6] mainly bases on the IS-A relation from WordNet, which could only process nouns. The method has adopted the concept of information content, in which concept node Ni represents a concept entity composed by a certain amounts of information, edge represents the directed relation between tow concept entities and the semantic similarity between two concept entities is the degree of information sharing. Richardson and Smeaton proposed an amended formula by integrating basic concept factor. In 1997, Jiang and Conrath proposed another measure method [2] to compute semantic similarity based on information content. This method determines the semantic similarity through computing the semantic distance between concepts.

$$Distance = IC(c_1) + IC(c_2) - 2 * IC(lcs(c_1, c_2)) \quad \text{Formula 1}$$

$$sim(c_1, c_2) = \frac{1}{Distance} \quad \text{Formula 2}$$

3.2. Edge-based method

The edge-based method to compute semantic similarity is commonly adopted, relatively natural and objective. This method tries to obtain semantic similarity by measuring the distance of the edge between two concept nodes. In multiple-dimension space, the distance between concepts could be determined by computing the geometrical distance between two concept nodes within the conceptualized hierarchical structure. The shorter the distance between nodes, the more similar they are. To automatically determine the weight of edges, other relevant factors should also be considered in computing similarity, which are the features closely correlated with conceptualized hierarchical structure. For example, local domain density (namely the number of edges stretched from a father node to son nodes), the depth of the node in the conceptualized hierarchical structure, link type and link strength. In 1998，Leacock and Chodorow proposed normalized path length method [8].

3.3 Comparison of similarity measure methods

Those two methods compute semantic similarity from different aspects. The edge-based method is more direct and obvious, while the node information content-based method has stronger theoretical supports. So these two methods both have advantages and disadvantages.

The edge-based method has been applied to medical domain [4], and the method has found that distance function could well evaluate the distance between concepts without human interferences. In document [2], the researchers have pointed out that when the method was applied to more domains, its accuracy would decrease. They have also found that irregular link density between concepts is the direct cause of erroneous conceptual distance.

Information content-based method does not require detailed content of conceptualized hierarchical structure, thus it would not have the fluctuation problem of conceptual similarity resulted from various link types. But since it ignores some information of the conceptualized hierarchical structure, it also has some inevitable drawbacks. Since the initial designing intent of WordNet is not for similarity computation, therefore some local hierarchical structures are not suitable for direct distance operations. To address the features of business information domain, this paper proposed a semantic similarity computation, EIM-based and information content-based and edge-based methods integrated.

3.4. EIM-based conceptual semantic similarity computation

To compute the similarity between pages and inquires, each page with its corresponding semantic eigenvector is used to represent the association degree between page and semantic concept. Each page is related to several concepts, and each concept is related to several pages, resulting in a many-to-many relationship. In order to reduce the number of eigenvectors, only when the value of a component vector of the eigenvector is bigger than the threshold could appear the eigenvector in the next hierarchy. Dividing the semantic features of page into three hierarchies and using semantic eigenvectors to represent those hierarchies respectively, a page could be defined by semantic eigenvectors as follows:

$$D(L) = (< c_1, w_1 >, < c_2, w_2 >, \ldots < c_n, w_n >) \quad \text{Formula 3}$$

$$D(M_i) = (< c_{i1}, w_{i1} >, < c_{i2}, w_{i2} >, \ldots < c_{im}, w_{im} >) \quad \text{Formula 4}$$

$$D(S_i) = (< c_{i1}, w_{i1} >, < c_{i2}, w_{i2} >, \ldots < c_{iv}, w_{iv} >) \quad \text{Formula 5}$$

[1] This work was supported by the grand foundation of Education department of Zhejiang Province under Grant No (20050677)

Where c_{ik} represents the *k*th concept related to the *i*th page, w_{ik} is the weight, which could be obtained by using the learning method to be described in section 4. L, several Mi and Si compose the semantic eigenvector of each page.

In the research of semantic similarity algorithm of concepts, the information statistical method of collected works ignores the information of fundamental relation types, making the computation results different from the EIM actuality. This paper combines EIM with information statistical method of collected works and proposes a method to compute semantic similarity between words. This method also takes four factors (link density, link strength, and link type and node depth) into consideration. When the concept in computation does not appear in the EIM structure, we simply use the WordNet structure to describe the semantic of the word.

First consider the link strength factor. The ratio of the sub-link strength to the probability of son concepts' appearance under father concept

$$P(c_i \mid p) = \frac{P(c_i \cap p)}{P(p)} = \frac{P(c_i)}{P(p)}$$ Formula 6

According to the view of information theory, EIM link strength (ELS) is defined as:

$$ELS(c_i, p) = \frac{1}{P(c_i \mid p)} = IC(c_i) - IC(p)$$ Formula 7

L(c, p) and T(c, p) represent the link type factor. In computing semantic similarity among concepts, the relative importance between conceptual entities is not identical, so the weight of paths between entities should vary to represent actual situations. L(c, p) and T(c, p) use the form of experience factor, and the list of different relations and factors are showed in Table 1.

Table 1. the value of type factor of links

Fundamental relation L（c,p）				
IS-A	Part-of	Associative	Synonymy	Mero nym
τ=0.3	τ=0.2	τ=0.2	τ=0.2	τ=0.1

Extended relation T(c,p)		
Geo-location	Apply-range	Apply-object
η=0.3	η=0.3	η=0.4

Taking comprehensive considerations of EIM node depth, link type and link strength, the link weight between son concept c and father concept p is defined as follows:

$$sim(c, p) = a\frac{\overline{S}}{S(p)} + \beta(\frac{d(p)+1}{d(p)}) + \varepsilon ELS(c, p) + \eta T(c, p) + \tau L(c, p)$$ Formula 8

Where $\alpha + \beta + \varepsilon + \eta + \tau = 1$, S(p) is the strength of link density between two concept nodes, namely the number of links of the son concept node from the father concept node; d(p) is the length of father concept node p in EIM. When $\alpha = \beta = \eta = \tau = 0$ and $\varepsilon = 1$, formula 16 is degenerated to the semantic distance formula proposed by Jiang-Conrath. The semantic similarity between words is computed as follows:

$$sim(w_1, w_2) = \sum_{c \in Set(c1,c2)} sim(c, p) - sim(lcs(c, p), p)$$ Formula 9

where c1=sen(w1), c2=sen(w2) and Set(c1,c2) are the set of all the concept nodes on the shortest path from c1 to c2. With this formula, it is easy to compute the similarity between page semantic eigenvector and searching semantic eigenvector, and the procedures are listed below:

Input: the searching keywords $Q_{Q_i \in Q}$ ={t1,…tn} and

the semantic eigenvector of page $S(D_i)$

Output: the similarity between the *i*th page D and Q

For $Q_i \in Q$ (Q is a set of keywords)

For $L(t_k) \in D$ (D is a set of the FLEN)

if t $\notin$ EIM Entity Node

SIM_Value$_i$ e=WordNet-similarity(Qi,Li);

else

SIM_Value$_i$ =

$$sim(c, p) = a\frac{\overline{S}}{S(p)} + \beta(\frac{d(p)+1}{d(p)}) + \varepsilon ELS(c, p) + \eta T(c, p) + \tau L(c, p)$$

compute the maximal similarity FSim of component vector of FLEN

EndFor

locate LSim's corresponding component vector $L(t_k)'$;

For every none zero component vectors $M(t_j) \in L'_k$ (L'_k is a set of the SLEN)

if t $\notin$ EIM Entity Node

SIM_Value$_i$ =WordNet-similarity (Qi,Li);

else

SIM_Value$_i$ =

$$sim(c, p) = a\frac{\overline{S}}{S(p)} + \beta(\frac{d(p)+1}{d(p)}) + \varepsilon ELS(c, p) + \eta T(c, p) + \tau L(c, p)$$

compute the maximal similarity SSim of component vector of SLEN

endfor

locate LSim's corresponding component vector $M(t_j)'$;

For every none zero component vector $S(t_v) \in M(t_j)'$ ($M(t_j)'$ is a set of the TLEN)

if t $\notin$ EIM Entity Node

SIM_Value$_i$ =WordNet-similarity(Qi,Li);

else

SIM_Value$_i$ =

$$sim(c, p) = a\frac{\overline{S}}{S(p)} + \beta(\frac{d(p)+1}{d(p)}) + \varepsilon ELS(c, p) + \eta T(c, p) + \tau L(c, p)$$

compute the maximal similarity TSim of component vector of TLEN

endfor

endfor

The Simarity between Document and Query

DTQSim=($\phi * \mathrm{FSim}_i + \varphi * \mathrm{SSim}_i + \omega * \mathrm{TSim}_i$)/3

The explanation for the above algorithm: for every keyword of the searching, locate the concept C of the FLEN concept vectors which has the maximal similarity with t, then locate the maximal concept similarity of SLEN and TLEN respectively.

Use the formula Sim=($\phi * \mathrm{FSim} + \varphi * \mathrm{SSim} + \omega * \mathrm{TSim}$)/3 to compute the semantic similarity between the page and the words inquired. ϕ, φ, ω are all experience factors, so define $\phi = 0.2, \varphi = 0.3, \omega = 0.5$. Since the information content of FLEN is relatively less, let $\phi = 0.2$; while the TLEN has more information content, let $\omega = 0.5$ with a larger weight.

4. EIM-BASED BUSINESS SEMANTIC INFORMATION RETRIEVAL

For Web page information searching, searching web page content is the initial step to process semantic searching.

Currently, spider is used to search information on the Internet. This paper proposed an automatically learning "page-concept" system, which could interactive with user searching and feedback. When user initiates a searching, the searching pages are manually marked as relevant or irrelevant, so as to adjust the weights of the semantic eigenvector of those pages. To eliminate the influence of subjective errors from specific users, we present a weighted algorithm to learn and update semantic net, which is an automatically learning "page-concept" system that could interactive with user searching and feedback. The detailed learning algorithm is listed as follows:

Step1: record users inquire and mark the feedback pages as relevant and irrelevant pages

Step2: for every keyword K_i of the search, use formula 1 to obtain its corresponding concept C(Ki).

Step3: define $C = C(K_1) \cup C(K_i) \cup C(K_N)$, where N represents the number of keywords of the searching

Step4: for every relevant page Di, check whether there is any concept Cj that is not related to Di. If there exists such a concept Cj, then establish the associative link between Cj and Dj and set the initial weight as 1.

If not, increase the weight w_{ij} of the link between Di and Cj by 1.

Step5: for every irrelevant page Di, check whether there exists any concepts $c_j \in C$ that are related to Di. If there exists such a concept c_j, then divide the weight of the link between c_j and Di by 6.

If the weight is smaller than 1, then the link would be deleted.

This algorithm could realize the EIM-based information feedback retrieval. In the searching process, this algorithm could interactively improve the breadth and depth of the "page-concept" associations. According to the algorithm, for certain pages, the relevant concepts approved by most users would get a larger weight. Even if there are subjective errors from specific users, the wrong association would be corrected right away by correct interactions, so this algorithm has excellent robustness.

5. EXPERIMENTAL RESULTS AND ANALYSIS

5.1. Experimental analysis of EIM-based semantic similarity

1200 words were used for the test. All the data was categorized into two classes, one class was for general word pairs, and another class was for the word pairs that appear relatively more often in product information. The tests were conducted on currently available algorithms (lcn[8],lin[9], lesk[10]) and proposed EIM-based semantic similarity algorithm.

Define $\alpha = 0.1, \beta = 0.2, \varepsilon = 0.3, \eta = 0.2, \tau = 0.2$. We compared the proposed algorithm with those three similarity algorithms. The comparison of the results of those three algorithms and EIM-based semantic similarity algorithm is listed in Table 2.

Table 2. The comparison of similarity of four methods

Word1	Word2	lin	lesk	lch	EIM
Gem	Jewel	1	228	3.583519	3.583519
Journey	Voyage	0.823042	219	2.890372	3.683321
Boy	Lad	0.816054	150	2.890372	2.89037
Coast	Shore	0.966516	324	2.890372	2.89037
Asylum	Madhouse	0.982675	104	2.890372	2.89037
Magician	Wizard	1	907	3.583519	3.58351
Midday	Noon	1	152	3.583519	3.58351
Furnace	Stove	0.220048	185	1.504077	2.47647
Computer	Monitor	0.423037	439	1.791759	2.532846
Computer	Keyboard	0.428604	766	2.197224	2.532846
Computer	Mouse	0.133027	96	1.974081	2.532846
Computer	Disk	0.506633	635	1.974081	2.532846
Computer	Usb				1.289542
Computer	Internet	0.342019	77	1.504077	1.973774
Camera	Photo	0.301498	201	1.386294	1.631356
Camera	File	0.347691	96	1.504077	1.542982
Camera	Camcorder	0	564	2.890371	3.437023
Camera	Printer	0.347691	103	1.974081	2.934832
Camera	Lens	0.379291	238	2.197225	2.934832

As seen from table 2, The WordNet-based semantic similarity between the Ecommerce Information Nodes is not too high. But for EIM-based similarity algorithm, since the extended relations and link factors have influenced and adjusted the value, the result is much improved. For example, the value between journey and voyage is obviously higher than other algorithms. For concepts not existed in EIM, the algorithm would compute the same way as lch on the basis of WordNet, so the result would be identical to the result from lch..

The semantic relation between computer and monitor is not too close, but both of them play important roles in real business activities and have intimate relations. For example, the information interactions of computer mainly depend on monitor as the output device, while monitor and keyboard are the output and input devices respectively with similar functions. EIM-based algorithm has greatly increased their semantic similarity to computer, which suits the actual important relations between those concepts in business activities. Since the concept USB does not exist in WordNet, these three WordNet-based algorithms could not compute the similarity between computer and USB. But through EIM, we could get the similarity between those two concepts, while in actuality they do have relations. In the computation of the semantic similarity between camera/photo and file, the results from these three similarity algorithms do not coincide. We regard the result of Lesk algorithm more explanatory for the direct relation between camera and photo. EIM-based similarity algorithm increases the semantic similarity between camera and photo greatly, by considering the link weight and information content. Though the structures of camera and file and camera and photo are basically the same, different extension relations will cause different semantic similarities. From those data, we can conclude that EIM-based similarity algorithm from the aspect of business application could more properly reflect the semantic similarity between concepts.

5.2. Experimental analysis of ECommerce-oriented semantic information searching

We have chosen about 4000 documents relevant to business information as the training set from TREC-11 and selected about 9000 documents of various topics as testing set.

[1] This work was supported by the grand foundation of Education department of Zhejiang Province under Grant No (20050677)

Collected works of every class was again categorized into 40 related topics, and all the documents were chosen from news, journals, application files of patents, articles of computer sciences, all marked by SGML. The precision rate and recall rate were used to evaluate the results, where

$$\text{Precision} = \frac{|\text{Relevant} \cap \text{Retrieved}|}{|\text{Retrieved}|}, \quad \text{Recall} = \frac{|\text{Relevant} \cap \text{Retrieved}|}{|\text{Relevant}|},$$

Relevant represents related documents, Retrieved represents the documents retrieved and |x| is the number of document set x.

The comparison testing were conducted on TF*IDF method, which is commonly used in full-text retrieval, and EIM-based feedback method for 250 business information searching. Fig. 2 shows the Precision rate at the same Recall rate.

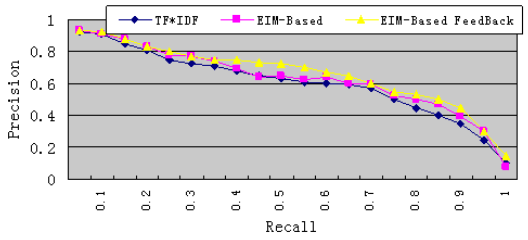


Fig. 2. With 50 searching, the average Recall-Precision Curve of the first 50 documents

From Fig 2, we can see that the precision rate of EIM-based method is much better than TF*IDF method for the first 50 documents, while the EIM-based feedback method could achieve even higher performance. But when the return set is relatively large, then the results of both EIM-based algorithms are not so realistic. Because EIM is a structure of ECommerce-oriented, when the ratio of ECommerce information to the entire collected works is low, the performance could not be properly reflected.

6. CONCLUSIONS

Based on EIM structure to process semantic feature extraction and the information content-based and edge-based semantic similarity methods, this paper presents EIM-based concept similarity method and an automatic learning "page-concept" system which could interactive with user searching and feedback. With proposed method, when user inputs the keyword for searching, the system would return a page set. User could mark the pages of the page set as relevant or irrelevant. Comparison tests were conducted on this method and text vector-space method. The experimental results show that EIM-based information retrieving method is much better than traditional text vector-space method. Since ECommerce ontology is quite extensive to construct manually, we have only established EIM for part of the product information of ECommerce domain. Further research would focus on the automatic learning method for establishing ontology structure, which could process semantic analysis of information and automatically extend the ontology database.

7. REFERENCES

[1] T. Berners-Lee, Hendler, The semantic web, Scientific American, 2001, 279

[2] Richardson, A.F. Smeaton, "Using WordNet in a Knowledge-Based Approach to Information Retrieval", Working Pape, PCA-0395, 1995.

[3] G. Miller, Nouns in WordNet: A Lexical Inheritance System, International Journal of Lexicography1990, 3(4):245-264.

[4] R. Rada, H. Miliet, "Development and Application of a Metricon Semantic Nets", IEEE Transactions on Systems, 1989, 19(1):17-30.

[5] A.Tversky, "Features of Similarity", Psychological Review, 1977, (4):327 - 352.

[6]P. Resnik, "Using Information Content to Evaluate Semantic Similarity in Taxonomy", Proceedings of the 14th International Joint Conference on Artificial Intelligence, 1995, 1:448-453.

[7] D.G. Ido, L.I. Lee, et al, "Similarity-based models of word cooccurrence probabilities", Machine Learning , 1999,34(1-3).

[8] C. Leacock, M. Chodorow, "Combining local context and WordNet similarity for word sense identification", In Fellbaum,C., ed., WordNet: An electronic lexical database. MIT Press, 1998, 265-283.

[9] D. Lin, "An information-theoretic definition of similarity",In Proceedings of the International Conference on Machine Learning, 1998,296–304.

[10] S. Banerjee , T. Pedersen, "An Adapted Lesk Algorithm for Word Sense Disambiguation Using WordNet",In Proceeding of the Fourth International Conference on Computational Linguistics and Intelligent Text Processing (CICLING-02) ,2002,136-145.

Simulation of Stock Prices Behavior Based on Cellular Automata

Ying Chen , Yun-jie Zhang, Yan-hui Bi
Department of Mathematics, University of Dalian Maritime
Dalian, Liaoning Province, China
Email: chenying@newmail.dlmu.edu.cn , yunjiezhangcn@sohu.com, biyh2004@newmail.dlmu.edu.cn

ABSTRACT

The modeling theory and method of cellular automata are applied to the study on the price behavior in stock market. Based on cellular automata, an evolutionary model for stock prices behavior is formulated. The paper classifies different kinds of investors in the stock market into three groups, each group has its own strategy. The random movement in stock prices reflects people's strategy and random changes going on in the outside environment. According to the study of stock prices movements under different kinds of investor combination and influencing degree of the macro factors, the simulation results show that technical investors and the macro factors are the key factors to bring the stock price move up and down.

Keywords: Cellular Automata, Price Behavior, Investors, Stability, Macro Factors.

1. INTRODUCTION

In the stock market, investors have their own relative independence, and they do investment analyses and make decisions according to the information they have. For each stock investor, there are three kinds of investment behavior: selling, holding, and buying. While in the process of decision-making, the difference in each stock investor's strategy leads to different investment behaviors. Meanwhile, the macro factor is another factor that influences investment behavior. The investor's investment behavior directly determines stock prices behavior, and the stock price is a reflection of investment behavior. Influenced by investment behavior the stock price goes up and down.

There are many investors as components in the stock market, each obeying simple rules, and the large number of investor make the stock market bring very complex behavior. In order to investigate the complexity in the stock market, some basic research approaches are developed: Artificial Life, Adaptive Agent Simulation, Cellular Automata, "CA", for short, System Dynamics, Neural Networks and so on. Sometimes the complex behavior in stock market may be simulated numerically with just a few components. But at the most of times the simulation requires too many components, many methods fail. But there is one method can solve this problem, that is CA.

Cellular automata were originally introduced by Von Neumann and his collaborator Ulam (under the name of "cellular spaces") as a possible idealization of biological systems in the 1950's, with the particular purpose of modeling biological self-reproduction. CA is discrete dynamical systems, of simple construction but complex and varied behavior [1]. It is used as mathematics model for systems in which many simple components act together to produce complicated behavior and can capture the essence of the process during evolution. For the characters of CA it has as above, it is applied in various areas to model the processes in physical [2], biological systems [3], geography [4], and economic [5].

Today, more and more people concentrate on the study of the stock prices behavior, but there are few papers are found in the area of the application of the CA to the simulation of the stock market, yet. Our motivation is to develop the potential of CA, and try to model prices behavior in stock market with cellular automata.

2. CELLULAR AUTOMATA

Cellular automata are mathematical idealizations of physical systems in which space and time are discrete and physical quantities take on a finite set of discrete values. A cellular automaton consists of a regular uniform lattice (or "array"), in some extent, usually infinite, with a discrete variable at each site ("cell"). All of the cells constitute the space of cellular automata, which is a discrete space lattice; the lattice could be one, two or three-dimensional. The state of a cellular automaton is completely specified by the values of the variables at each site. A cellular automaton evolves in discrete time steps, with the value of the variable at one site being affected by the values of variables at its neighborhood sites on the previous time step. For one-dimensional cellular automata, the neighbors of a site are typically taken to be the site itself and all immediately adjacent sites. For two-dimensional cellular automata, there are many forms of the neighbors, usually "five-neighbor square", and "nine-neighbor square" illustrated in Fig. 1, we call it as von Neumann and Moore neighbors, respectively. Triangular and hexagonal neighbors are also possible, but are not used in the examples given here. Note that the five-neighbor square, triangular, and hexagonal cellular automata may all be considered as special cases of nine-neighbor square. The variables at each site are updated simultaneously ("synchronously"), based on the values of their neighbors' variables at the preceding time step, and according to a definite set of "local rules". So a cellular automaton consists of cellular, cellular spaces, neighbors, cellular states and local rules.

(a) (b)

Fig. 1. The typical neighbor squares: (a) von Neumann; (b) Moore

3. CELLULAR AUTOMATA IN STOCK PRICE BEHAVIOR MODELLING

In the real financial market, the stock prices are determined by some complexity environment in which different investors apply their own strategy. The investors make decisions according to the information they have, all investors' decisions determine the stock prices; meanwhile, the present decisions they make may influence investors' next decision. Thus, investment behavior influences stock prices movement and makes it fall and rise. In our country's stock market, there are two investment entities: organization investor and personal investor. In this paper, for simplicity, we take personal investor for considered only. The types of the personal investor could be defined as follows:

Technical investors: The technical investor studies the movement of stock price and trading volume over time in an attempt to make a mathematics analysis. Generally, the technical investor relies on analysis of relation between stock price and trading volume to determine his investment strategy. This is achieved through a combination of complex graphs and charts, specially information and a strong knowledge of the markets.

Fundamentalists: The fundamentalists' studies the macro factors about political, economic to analyze the relation between macro factors and stock market, then he tries to determine where a stock is going and then capitalizes on that knowledge.

Interacting investors: He has no idea with himself, just obverse all neighbors around him. The passed behaviors of people around him influence his decision. He observes every one around him, if most of them buying, so he will buy too. Similarly, only time he will sell (hold) is when most of his neighbors selling (holding).

To formulate the cellular automata model for the simulation of the stock prices behavior in the stock market, we assume that the amount of money and the number of shares held are infinite; each investor does the stock exchange one time at the most in a day. The cellular automata model for the stock prices behavior evolution could be defined as follows:

A cellular automata represents a stock market. And the neighbors' types adopt Moore neighbors.

The state space of cellular automata S: based on investment behavior and investor, we built a two-dimensional state space of cellular automata. The first dimension denotes investment behavior (S_1), it can choose the following three different values: s_{11} (representing buy), s_{12} (representing hold), s_{13} (representing sell), namely: $S_1 = \{s_1, s_2, s_3\}$; the other dimension denotes investor S_2, it includes technical investor s_{21}, fundamentalist s_{22}, and interacting investor s_{23}, namely: $S_2 = \{s_{21}, s_{22}, s_{23}\}$. Thus, a state space of cellular automata can be represented as: $s = \{s_1, s_2\}$, where $s_1 \in S_1$, $s_2 \in S_2$.

Evolution rules: because each type of investor has its own strategy, so its evolution rules different.

Technical investors: Assume that all technical investors take the same technical analysis method.

When the stock price rises sharply, investors will sell the stock at the next time;

When the stock price rises at a moderate rate, investors will buy the stock at the next time;

When the stock price remains unchanged more or less, investors will hold the stock at the next time;

When there is a big fall in the stock price, investors decide to buy at the next time;

When there is a moderate fall in the stock price, investors will take the decision of selling the stock at the next time.

The fluctuation of price is defined as

$$f_p = \frac{p_C - p_O}{p_0} . \quad (1)$$

Where P_c represents present stock price, P_o represents previous stock price.

We use λ, θ to describe the daily volatility of price, the value of it is determined by the characters of investors. According to the comparison f_p with λ and θ, the technical investor's analysis also could be defined as:

When $f_p > 0, f_p > \theta$ investors will sell at the next time;

When $f_p > 0, \lambda < f_p < \theta$, investors will buy the stock at the next time;

When $-\lambda < f_p < \lambda$, investors decide to hold the stock at the next time;

When $f_p < 0, -\theta < f_p < -\lambda$, investors will sell at the next time;

When $f_p < 0, f_p < -\theta$, investors will take the decision of buying at the next time.

Table 1. Evolution rule of fundamentalists

Macro factors	The cellular states at the next time (the possible investment behavior of the investors and the corresponding transfer probability)		
	Buying	Holding	Selling
Positive information	P	$(1-P)/2$	$(1-P)/2$
Negative information	$(1-P)/2$	$(1-P)/2$	P

Fundamentalists: Analyze the political and economic aspects, we introduce the macro-factor coefficient P to decide what decisions should be made. The coefficient P describes the influencing degree of the macro factors. When the macro situation is positive information, fundamentalists

will buy with the probability P, where $P \in [0, 1]$; and while the macro circumstance is negative information, fundamentalists will sell with the probability P, where $P \in [-1, 0]$, so $P \in [-1, 1]$. Therefore, fundamentalists make investment decision in accordance with the coefficient P, the specific evolution rules are listed in Table 1.

Interacting investors: It is supposed that the given cell state in the stock market at the next time step can be affected only by the investment behavior of the neighbors, not include itself. The cellular state here means the investment behavior. The investment behavior of the neighbor cells is determined as the most investors among them at the previous time step according to the majority. The state of the considered cell will evolve into the state determined according to the majority principle from the neighboring cells at the previous time step. For example, the current state of its neighbors is buying, and then the cell itself will buy at the next time.

4. THE ANALYSIS OF SIMULATION EXPERIMENT

In our simulations, we introduce some input parameters of the simulation experiments, that is, the sizes of the grid spaces, the shapes of the initial states, the initial stock price, the steps of evolution, and the coefficient of the macro factors. For the sizes of the grid spaces, if they are too big, the speed of simulation will become slow, but if too small, the evident results of the simulation experiments cannot be identified. In this paper, the grid 50×50 is adopted. The shapes of the initial states have two types. One is identic, and the other is stochastic. We adopt the latter in this paper. The evolving steps are 30, just a month. There are two kinds of macro coefficients here: one that takes the same value through the space-time, and another which coefficient of each fundamentalist belongs to the set [0, 1].

The capacity of market (C), trading volume (V_b), and price (P_b) constitute the output parameters. The definition of capacity of market is the sum of the number of behavior investment at a given time. The trading volume could be expressed as in

$$V_b = \min\left(B_{toaal}, O_{total}\right) \tag{2}$$

where B_{total} represents the number of buying in the stock market and O_{total} represents the number of selling in the stock market at a given time. The dynamics of price could be expressed according to

$$P_b(t) = P_b(t-1) \times \left(1 + eta \times \left(B_{total} - O_{total}\right)\right). \tag{3}$$

where $P_b(t)$ represents the current price, $P_b(t-1)$ represents the previous price, *eta* is a adjusting coefficient, we take 0.1 here.

Table 2.
Statistics of stock prices under different kinds of investors combinations

Proportion of technical investors (%)	The highest price (P_h)	The lowest Price (P_l)	The opening price (P_p)	The closing price (P_c)
30	12.1838	6.2183	7.5	6.3489
40	14.0638	3.6539	7.5	4.9654
50	15.1654	3.6565	7.5	10.4089
60	25.6641	5.5561	7.5	15.7834
70	12.1838	0.0470	7.5	0..470

There are a lot of indexes of weighing the Chinese stock market price and fluctuating. In this study, we choose two, which contribute greatly to influence the analysis of the stock price fluctuation, that is, stock price index and amplitude. The former completely reflects the level of the stock prices and how prices change; the later reflects the range of stock prices fluctuate. Then we can simulate the stock market from two aspects: the different kinds of investors' combinations and influencing extent of the macro factors.

Table 3.
Statistics of monthly up/down and amplitude of stock prices under different kinds of investors combinations

Proportion of technical investors (%)	UP/DOWN of monthly stock prices (%)	Amplitude of monthly stock prices (%)
30	–15.35	99.86
40	–33.80	284.90
50	38.79	314.76
60	110.45	362.41
70	–99.37	25799.77

4.1 The simulation of the stock prices behavior under various kinds of investors' combinations

According to Shanghai Security Exchange's analysis report in 2003, the proportion of the high school education and under high school education among individual investors has up to 65.92%. They usually lack essential knowledge to analyze the market information, and they know their limitations, often try to ask about information and decisions around them to determine what they should do at the next moment. Based on the classification of investors in this test, this kind of investor is interacting investors. And we fetch 60% for convenience sake, technical investors and fundamentalists each account for 20%. In this paper, at first , in order to investigate the influence of different kinds of investors' combinations on the stock prices, we fix the proportion of fundamentalists 20%, while tune the proportion of technical investor during the simulation. The coefficient of macro-factors follows random distribution, that is different values of the macro coefficient P can be chosen from the set [0, 1].

Table 2 is the simulation results of the model. It manifests the statistics of stock prices under different kinds of investor combinations.

In order to make the movement of stock prices more

visually oriented, we select 30%, 50%, 60% these three kinds of situations from the experimental data to make a graphic representation of the tendency of stock prices (shown in Fig. 2). As can be seen from the Fig. 2, we can find out that there is a big rise and fall of the stock prices, take on the characteristic of a large margin and frequently volatility.

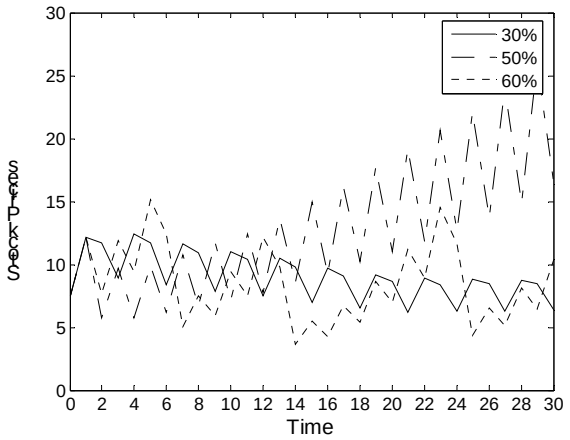


Fig. 2. The time series of the stock prices during the processing

Meanwhile, according to the data in the Table 2 can draw monthly UP/DOWN (P_{ud}) and monthly amplitude (P_a) of the stock price through calculating illustrated in Table 3. The monthly UP/DOWN of the stock prices was obtained by

$$P_{ud} = \frac{P_c - P_p}{P_p} \tag{4}$$

and monthly amplitude of the stock prices then was obtained in analogy with Eq. (4) according to

$$P_a = \frac{P_h - P_l}{P_l} \tag{5}$$

Table 3 indicates that with the increase of the technical investors' proportion, the amplitude rises from 99.86% to 25799.77%. It is apparent that the larger technical investors' proportion is, the larger the fluctuation of the stock price is.

4.2 The simulation of the stock prices behavior based on the influencing extent of the macro factors

In order to find the influencing degree of the macro factors, it is supposed that the proportion of the technical investors and fundamentalists is 20% respectively, and the coefficient of macro-factors takes a constant during the simulation. The macro factors are of two kinds: the positive and negative information. Since these two factors are symmetric, we can choose one of them arbitrarily. In this paper, we choose the positive macro information. According to the value of the macro coefficient, we simulate three situations of which the input parameters are: 0.2, 0.5, and 0.8. The previous two simulations represent common positive information. The last simulation means the positive information is uncommon. Fig. 3 and 4 describe the changing process of the numbers of the stock prices in these three situations.

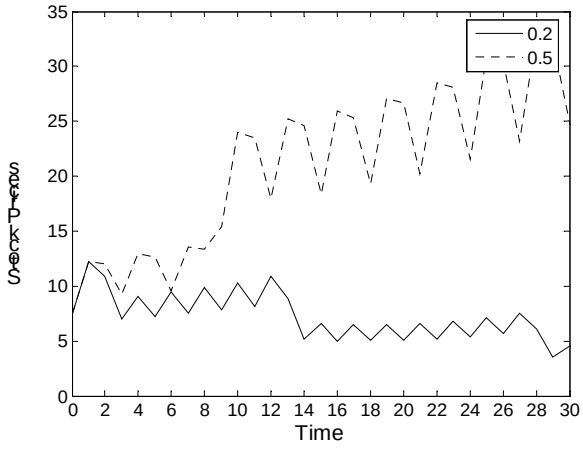


Fig. 3. The time series of the stock prices of common positive information

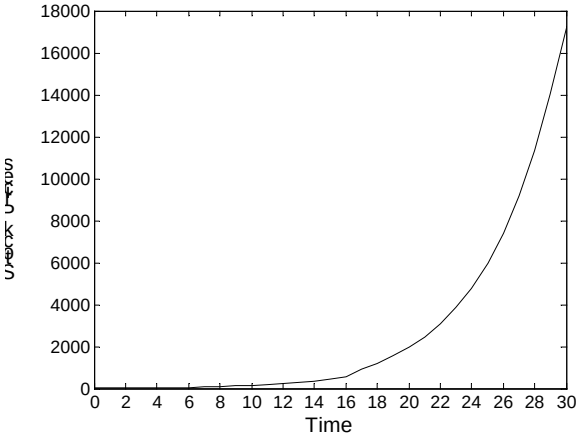


Fig. 4. The time series of stock prices of uncommon information

The Fig. 3 and 4 give information that the different positive degree will result in the different evolution in the stock market. While the common positive information appeared, the prices are subject to frequent change; the range of the prices became great with the increase of coefficient. While the uncommon positive information appeared, the investors did their best to buy, which caused the buying behavior of the investors take up the whole simulation space and made prices seemed out of control, climbed at the swift and violent speed. It embodies that the bigger the value of the macro coefficient, the more violent fluctuation, and the smaller, the less violent.

5. CONCLUSIONS

By the analysis of the simulation results of the prices behavior in the stock market, it could be concluded as follows:

Under the different combinations of technical investors and fundamentalists, the fluctuation of the stock price is more violent with the proportion of technical investors

increasing. It manifests that the technical investor is the key factor affecting the movement of the stock prices. and vice versa.

This phenomena mainly attributes to a large number of technical investors pour into the stock market, make the stock prices realize to enhance the mechanism oneself. When some technical indexes show that is “expected to rise” or “expected to fall”, the buyer and seller roll into the stock market in crowd and strengthen the present trend of stock prices. This leads more investors to adopt the same trade decision-making, thus there is a big rise or a big fall in the stock prices. But too much technical investors in the stock market will cause the collapse of the price.

Furthermore, the macro factor is anther important factors influencing the movement of the stock prices. The bigger the value of the macro coefficient, the more large fluctuation is the stock prices, and vice versa. While only the common macro information appears, the stock prices take on the steady shock, and while the macro information is uncommon, the stock market will finally collapse. The change of policy on the stock market lead the stock prices move up or move down, that proved the tendency of stock price of our country is closely linked with the tendency of macroeconomic.

In this paper, by means of Cellular Automata, we try to explain the possibility of the stock prices behavior in terms of investment behavior, provide a new idea for understanding the stock prices behavior.

6. ACKNOWLEDGEMENTS

We thank the anonymous referees for very helpful suggestions.

7. REFERENCES

[1] S. Wolfram, Theory and Applications of Cellular Automata, World Scientific, Singapore, 1986.

[2] M. E. Epstein, R. Axtell, Growing Artificial Societies, The MIT Press, 1996.

[3] N. Zimic, M. Mraz, “Towards the Bottom-up Concept: Extended Quantum-dot Cellular Automata”, Microelectronic Engineering, vol. 83, September 2006 ,. 1826-1829.

[4] D.G. Mallet, L.G. De, “A Cellular Automata Model of Tumor–immune System Interactions”, Journal of Theoretical Biology, April 2006, 293(3): 334-350.

[5] M. Bartolozzi, A.W. Thomas, “Stochastic cellular automata model for stock market dynamics”, Physics Review. E, April 2004, 69(42). 046112/1-046112/7.

Game Theory Study Of Currency Speculation In Financial Market

Ying Liu
School of Economics and Management, Beijing Jiao Tong University
Beijing China
Email: liuyingdolphin@hotmail.com
ShuJun Ye
School of Economics and Management, Beijing Jiao Tong University
Beijing China
Email: shjye@bjtu.edu.cn

ABSTRACT:

To guard from the risk of currency speculation，the course of currency speculation in the financial market was studied, in which the speculators' target was the countries with strict exchange rate regulation. The Incomplete Information Dynamic Game Theory was applied to study the game between the government and the speculators. The Complete Information Static Game, Boxed Pigs Game and Evolutionary Game Theory were applied to study the game among speculators. The Nash Equilibrium solution was taken. The reason was analyzed that why there were different results among speculators who had attacked Pound and Hong Kong dollar. Conclusion is that the speculators succeed in the countries whose economy goes unsteady and the nominal currency value does not equal the actual currency value. China controls the exchange rate strictly as well but will open its capital market and then RMB can exchange freely. After this, the risk of currency speculation will increase steeply .The aim of above analysis is to enlighten the Chinese exchange rate system and economy operation.

Keywords: Currency Speculation, Game Theory, Nash Equilibrium, Sheep Effect, Exchange Rate System.

1. INTRODUCTION

In the international financial market, along with massively flows of the international capital, the phenomenon of the currency speculation increase day by day. The speculators' target is the country with strict exchange rate regulation, in which the currency market does not reflect its economical status truly, and the nominal currency value does not equal the actual currency value. The speculators attack a country's currency to profit from the change of exchange rate. There are games between the speculators and the government and games among the speculators when the international speculators attack a country's currency. In this paper, the game between the government and the speculators would be introduced at first and then, we will focus on the games among the speculators. In the financial market, the speculators can be divided into heavy speculators and small and medium - sized speculators, the different strategies that they selected will affect the result of game.

2. THE INCOMPLETE INFORMATION DYNAMIC GAMES BETWEEN THE GOVERNMENT AND SPECULATORS

2.1Incomplete Information Dynamic Game

The term incomplete information indicates that the information referred in the process of decision-making is partial and uncompleted. Since our energy, intelligence and aptitude are limited, our ability is limited to obtain and discern the information and the man-made information barriers and the deceptive information existing in the competition as well. So people cannot obtain perfect information needed in the process of decision-making. In a dynamic game the players take action successively and the player who acts later can observe the action of the player goes in advance [1]. In the game between the government and the speculators, each side considers the other side to be their antagonist and both sides would hide their own information. On the other hand the revenue function of both sides includes each other. Both sides continually correct their decisions and actions on basis of the observation of their antagonist and the correction they made would also influence the action of their antagonist. So the game between the government and the speculators can be considered to be an incomplete information dynamic game.

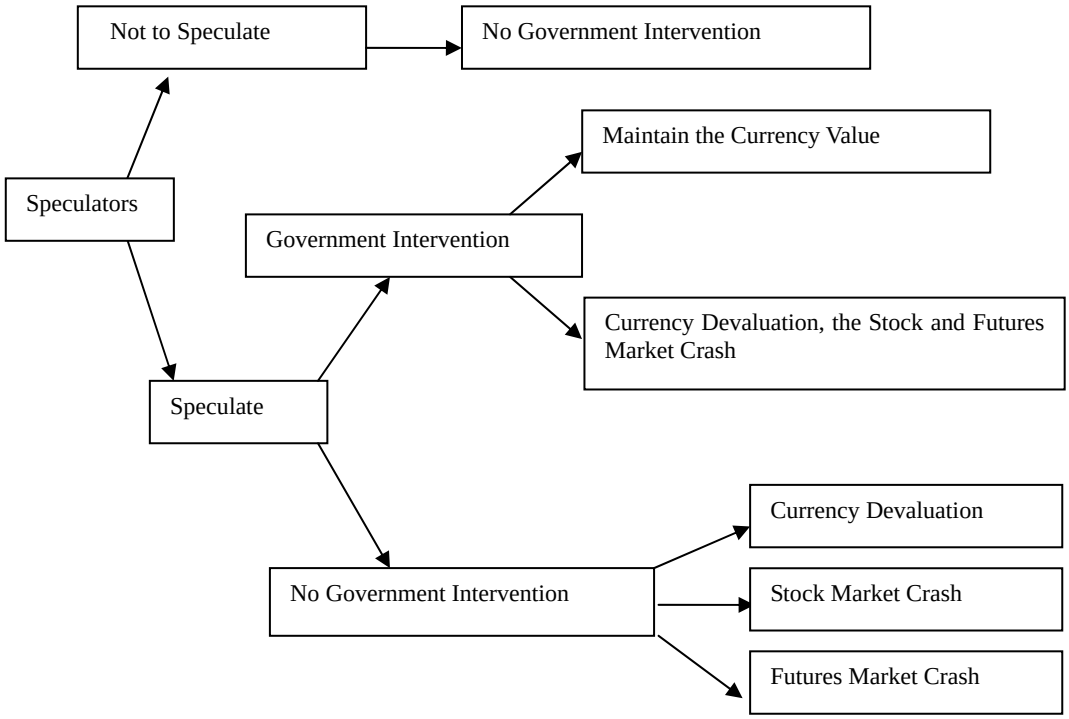
2.2The Mastery of the Information of the Government and Speculators and Their Countermeasure

First of all, in the international financial market, the speculators should estimate the strength of the government selected to be target before they attack the country's currency. If the speculators consider the government to be powerful and their attack cannot be resulted in the devaluation of the target currency but only to lead to loss of money, the speculators would not launch their attack. And if speculators consider the government to be weak and lack the ability to maintain a fixed exchange rate, their attack will bring to a huge profit, they will launch their attack. Secondly, the speculators will take the cost of their entrance into consideration. Since the speculators have to borrow a sum of target country's currency from bank in advance, the entrance cost depends on the lending rate of the target country. If the profit made from the stock and futures market can cover the cost of loan interest after the devaluation of the attacked currency induced by their speculation, the speculators will enter. In addition, the speculators also can use a part of their own capital. And this part of money can be used without paying loan interest. The cost of speculation can be reduced greatly, leaving the commission charge only. Generally speaking, the cost of the speculation depends on the amount of own capital and the degree of lending rate.

The information of the speculators that the government has is incomplete as well. The entrance cost of speculators can hardly be estimated. The only thing the government can do is to make decision of the intervention on the basis of estimation on the probability of the entrance cost of the

speculators. Generally speaking, the way to attack a currency is to sell a great lot of target country's currency to force the currency to devalue. And once the trend of the devaluation is formed, many people of the target country and common investors would lose confidence in their currency and would sell their currency and exchange it to foreign currency. This makes the pressure of devaluation increase. To maintain currency value, the government has to prepare enough foreign exchange reserve to buy currency which people sold back. Possibly, by reducing the monetary base the central bank can make the interest rate increase to such a high level that the speculators would find that the entrance cost is too high to enter [2]. But considering the other aspects of the economy and politics, a government can hardly maintain such a high interest rate in the long term because this can damage the whole bank system and the domestic economy operation critically. In the condition of incomplete information, the speculators judge the strength of the government by observing their performance. And the government does his best to show he is powerful to prevent speculators from attacking. Even after the speculators enter the market, the government and the speculators have not stopped making their continuous judgments. They are playing an incomplete information dynamic game.

2.3The Game Tree of the Game between the Government and Speculators：



3. THE COMPLETE INFORMATION STATIC GAME AMONG THE SPECULATORS

The complete information means that all players have the basic structure of the game, including: the payment function of the players and other information about the characteristics and policy spaces of players. The static game theory means, if the players meet each other only once, they take action at the same time. Or if they do not take action at the same time, the players that take actions later do not know what actions the previous players took [3]. In this paper's game among the speculators, a common goal that all speculators pursuing is to cause a country's currency devalues. Among the speculators, their information is relatively complete. The majority of the speculators are independent to each other. Before making decisions, one did not know what strategies the other players made. Further more, the massive players have only one opportunity to meet each other. Therefore in this paper we can describe the game among speculators as a complete information static game.

4. BOXED PIGS GAME

4.1The Boxed Pigs Game

Boxed Pigs is a game featuring a large pig and a small pig in a box. If one pig pushes a button, then some food appears in a slightly far away place. If a small pig pushes the button, the large pig gets most of the food. If a large pig pushes the button, then a small pig can get about a half [4]. The normal form of the game is listed in Tab.1

Table 1. The Model of Boxed Pigs Game

		Small pig	
		PRESS	WAIT
Large pig	PRESS	(5, 1)	(4, 4)
	WAIT	(9, -1)	(0, 0)

We can conclude from Tab.1 that whatever choice has the large pig chose; the optimum choice of the small pig is WAIT. Suppose the small pig has chosen WAIT, then the optimum choice of large pig is PRESS. What is the Nash Equilibrium strategy in this game? The Nash Equilibrium for the original Boxed Pigs Game is large pig PRESS, small pig WAIT.

4.2The Analysis of Financial Market on the Basis of the Theory of Boxed Pigs Game

The speculators in financial market can be divided into the heavy speculators and the small and medium-sized

speculators. The heavy speculators are similar to large pig. The small and medium-sized speculators are similar to small pig. The aim of all speculators is to devaluate this country's currency. PRESS strategy denotes attacking currency. The heavy speculators will decide whether to attack this country's currency according to the information of this country's economic condition and the degree that currency value is overrated or underrated. The optimum choice of small and medium-sized speculators is WAIT, that is, to follow the heavy speculators' choice. This behavior brings about the Sheep Effect [5]. But there is one difference in financial market, small and medium-sized speculators must invest capital to earn money, while small pig in the Boxed Pigs Game could get food without paying anything. What's more, in a certain degree, the small and medium-sized speculators' capital could contribute to the devaluate currency. If speculators make a right judgment and achieve speculation, they can profit from changing of exchange rate. But not all the currency speculations were successful, a large part of them failed. The fundamental factors, which determine the result of speculation, are the real economic condition and foreign exchange reserve. With both of the above factors staying, what is the key to the result of speculation? The key factor could be the ratio of initial participants. When the ratio of initial participants reaches a break point it will give rise to Sheep Effect，which will change the choice of speculators who did not participate speculation at first and they would start to speculate . As the number of speculators grows gradually, the government has to give up the initial policy of exchange rate and let it float freely. Finally speculators profit from changing of exchange rate. While the ratio of initial participants does not reach the break point, the population of speculators will decrease gradually, until all speculators withdraw from market. That means a speculation failed. The basic theory of the above conclusion is to be introduced Evolutionary Game Theory in the following passage.

5. EVOLUTIONARY GAME THEORY

Evolutionary Game Theory is originated as an application of the mathematical theory of games in biological contexts; arising from the realization that frequency dependent fitness introduces a strategic aspect to evolution. Recently, however, Evolutionary Game Theory has become of increased interest to economists. It is as readily applied to the field of economic behavior as that for which it was originally designed.

5.1Gene Duplication Dynamic Process Model

The so-called Gene Duplication Dynamic Process means each of the different communities of animal chooses a strategy of game in a moment. We assume that only the community with the best ability of adaptability is able to survive, it is usually indicated by the Malthus dynamic system in biological game theory, namely RD. According to RD model, if one community's adaptability has surpassed this population's average level, then the individual quantity of this community can continue to increase, while if one community's adaptability is lower than the average level, then the individual quantity of this community will decreased gradually.

We assume that there are massive participants in the game. The pure strategies of participants by the definition are: x and x'. Let n_t and n_t' respectively denote the quantity of the participants who choose strategies x and x' according to the certain procedure plan in the t moment. Let N_t denote the quantity of all the participants. Let $s_t(x)$ denote the proportion of participants who choose strategy x in the t moment:

$$s_t(x) = \frac{n_t}{N_t} \quad \text{Eq. (1)}$$

The expectation payment of the participants who prepare to choose the strategy x is:

$$u_t(x) = s_t(x)u_t(x,x) + s_t(x')u\ (x,x') \quad \text{Eq. (2)}$$

Thus, all participants' average payment is:

$$\bar{u}_t(x) = s_t(x)u_t(x) + s_t(x')u_t(x') \quad \text{Eq. (3)}$$

Finally we can obtain a RD model, namely:

$$F(s) = s(x)[u(x) - \bar{u}] \quad \text{Eq. (4)}$$

This equality has reflected the RD basic thought: If the strategy x result surpasses the average level, then these communities' proportion will increase in the animal, otherwise, the proportion will decrease [6].

5.2The Application of Evolutionary Game Theory

According to above analysis and The Boxed Pigs Game, the large pig and small pig all have two choices: PRESS and WAIT. Let s denote the proportion of the pigs, which is preparing to choose the strategy PRESS (that is, the ratio of initial participant of attacking a country's currency). When the ratio of initial participants reaches a break point, the population of speculators will grow gradually until all participants chose this strategy, while if this ratio does not reach the break point, the population of speculators will decrease gradually, until all speculators withdraw from market.

The pig, which initially chose PRESS, can obtain:

$$U_1 = s \times 5 + (1-s) \times 4 = s + 4 \quad \text{Eq. (5)}$$

Similarly:

$$U_2 = s \times 9 + (1-S) \times 0 = 9s \quad \text{Eq. (6)}$$

The average payment level is:

$$\bar{U} = s(s+4) + (1-s)9s = -8s^2 + 13s \quad \text{Eq. (7)}$$

The gene duplication equality $F(s)$ is:

$$F(s) = s[s + 4 - (-8s^2 + 13s)] \quad \text{Eq. (8)}$$

Finally it is:

$$F(s) = 4s(2s-1)(s-1) \quad \text{Eq. (9)}$$

The gene duplication equality is shown in Fig. 1.

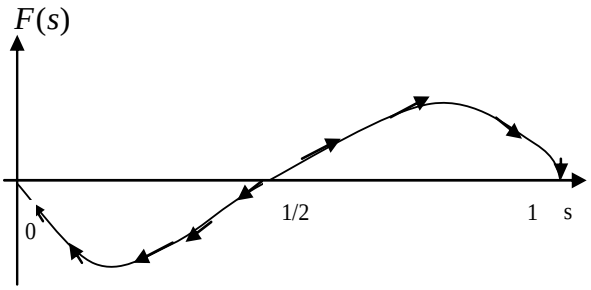


Fig. 1. The RD model

We can conclude from the above analysis that, in this game, if the proportion of pigs which have chosen the strategy PRESS (that is, the ratio of initial participant of attacking a country's currency) exceeds 1/2, then RD can cause this ratio continuing to increase until all participants have chosen to attack this country's currency. Otherwise, if the ratio of initial participants of attacking a country's currency does not exceed 1/2, RD can cause this ratio gradually decrease to 0.The model can help us to explain why the speculators' attacking the Pound had succeeded but speculators' attacking the Hong Kong dollar had failed.

6. GAME ANALYSIS OF CURRENCY SPECULATION

6.1Game Analysis of Speculation Process to Pound

In the international financial market, thousands of free capital flows from the places of low interest rate to the places of high interest rate, and flow from the inflatable countries to the states of strong currency. When Europe undergone the financial crisis, England was experiencing the long-term economic recession in which enterprise and the foreign trade were in depression, thus the Pound devaluation was unavoidable. The UK currency authority persisted in and kept from adjusting the monetary policy to maintain the value of Pound. But it did not prevent the occurrence of European financial crisis, because the financial crisis condition had already formed. During the game though The UK government had used about 24 billion US dollars of foreign exchange reserves and had raised the interest rate for two times within one day, this had been not able to maintain the Pound value [7]. After this disturbance, the Pound originally exchanging 2.95 German Mark converted into 2.2 German Mark in the spring of 1995 when Pound value was in its minimum point. It devaluated about (2.95-2.2) /2.95=25.4%. Although the UK economy recovered afterwards, meanwhile the Pound had revalued, the UK economy had been hurt critically. It was a miserable result that the UK government lost several billion Pounds, furthermore, UK withdrew from the mechanism of Europe unified exchange rate.

UK uses the foreign exchange reserve as much as 24 billion US dollars, on the assumption that all heavy speculators hold 10 billion US dollars, the small and medium-sized speculators hold 30 billion US dollars. Let the transaction cost 5%, the Pound devaluated about 25.4% after the speculation. The normal form of the game is listed in Tab.2:

Tab.2 the Model of Speculators' Attacking the Pound

		Small and medium-sized speculators	
		SELL	HOLD
Heavy speculators	SELL	(10.24,30.72)	(-5, 0)
	HOLD	No this condition	(0, 0)

10.24=240*25.4%*100/400-100*5%
30.72=240*25.4%*300/400-300*5%

Analysis: 1) transaction cost mainly consists of two parts, one is the handling charge of transactions; the other is lending rate. Generally speaking, the speculators use a little part of personal capital (do not have to pay interest) and borrow massive capital from the bank (need to repay principal and interest in the end), at that time UK lending rate was 12%, considering synthetically, we selected the transaction cost being 5% in this paper.

2) Because the small and medium-sized speculators always follow the heavy speculators' decision, so the decision-making power is in the hands of heavy speculators at all times, though the small and medium-sized speculators hold more capital than the heavy speculators. Therefore, the small and medium-sized speculators will choose strategy HOLD if the heavy speculators choose strategy HOLD. It is shown in the fourth frame of Tab.2, their income is (0, 0); It is impossible that the heavy speculators choose HOLD, while the small and medium-sized speculators choose SELL; The situation only happens in the extremely short time or even can not be displayed in practice, in witch the small and medium-sized speculators do not follow the heavy speculators to choose SELL. It is shown in the second frame of Tab.2, their income is (-5, 0). The common situation can be shown in the first frame of Tab.2, their income is (10.24, 30.72), this denotes that all speculators sell the Pound and earn the profit.

3) A Nash Equilibrium, named after John Nash, is a set of strategies, one for each player; such that no player has incentive to unilaterally change her action. Players are in equilibrium if a change in strategies by any one of them would lead that player to earn less than if she remained with her current strategy [8]. There are two Nash Equilibrium mentioned in this model: one is that all of the heavy speculators and small and medium-sized speculators anticipate that their rivals would not make the speculation of devaluation. On this condition, the optimum choice of the speculators is not to speculate. Although this is Nash Equilibrium, it is obviously untenable in the practical game. In the game, the speculators choose the Nash Equilibrium SELL and SELL; this is the best strategies for all speculators because they can achieve aim by cooperation. A remarkable character of Nash Equilibrium is: A player assumes that the behavior of his rivals is the optimum strategy corresponding to his own strategy. The anticipation becomes self-realization. Thus, the players can achieve their anticipation, such as the situation shown in this example.

4)The explanation of result: From the successful speculation process to Pound, we can conclude that the speculation could not succeed if the Sheep Effect did not work, that is, only the heavy speculators launch an attack without the small and medium-sized speculators following. The result can be explained by the Gene Duplication Dynamic Process Model, which means if the proportion of speculators who have chosen the strategy SELL exceeds break point (1/2), then RD can cause this ratio continue to increase until all participants have chosen strategy SELL. Finally speculators have achieved the optimum strategy Nash Equilibrium.

6.2Game Analysis of Speculation Process to Hong Kong Dollar

In the financial crisis of Southeast Asia, the speculators failed to attack Hong Kong dollar .The fundamental reason is that Hong Kong has solid economic basis, abundant foreign exchange reserve and stable political situation. At that time, Hong Kong's system of exchange rate adapted to the economy development perfectly, moreover there is formidable backing from China mainland [9]. The foreign exchange reserve of Hong Kong and China Mainland adds up to more than 200 billion US dollars, it is sufficient to resist the international speculators' attack [10].

The result can also be explained by the Gene Duplication Dynamic Process Model, due to the above macroeconomic, the speculators were lack of confidence in Hong Kong dollar speculation, the proportion of speculators who had chosen the strategy SELL did not exceeds break point (1/2), finally the RD would reduce this proportion to "0", namely all speculators would not choose the strategy SELL. This situation means the speculation failed and the speculators that chose strategy SELL at first had to pay for the loss of transaction cost.

7. CONCLUSION

In this paper we introduced the target of currency speculators, analyzed the game between the speculators and the government and the game among speculators, summarized the reasons which caused different results in speculators' attacking Pound and Hong Kong dollar, and all of these are to inspire the Chinese currency management and running of exchange rate system. China implements the managed floating exchange rate system which is controlled strictly now. We faced smaller risk of currency speculation previously because our capital market did not open completely. Due to the disability of RMB to exchange freely, the international currency speculators were unable to speculate in RMB. But now, our country is at a crucial time of opening the capital market. After opening the capital market, the risk of currency speculation will increase steeply. Therefore, we must embark on the relative research and prepare fully to prevent from the risk of currency speculation.

We can know from this paper that the radical reason causing the risk of currency speculation is the problem of economical operation. So keeping economy steady is the essential thing we should do. The insufficient foreign exchange reserve of a country as well as the unconformity between a country's nominal currency value and the actual currency value are the central reasons for the success of the speculation. Since the exchange rate system is managed floating in China, we should maintain large amount of foreign exchange reserve to guard against the currency speculators. Moreover, it is indispensable for our country to maintain RMB nominal value conformity with the actual value.

8. REFERENCES

[1] T. Huang, Game Theory, Capital Economy & Trade University Publishing House, Beijing, 2004.
[2] F.C. Li, "Interests Weighing and Game Theory in Exchange Rate System", Journal of International Business Study, Feb, 2003, 32-37.
[3] G. I, Gerhard, Theory of Monetary Policy, Social Science Document Publishing House, Beijing, 2002.
[4] W.J. Wang, Game Theory Application and Economics Development, Capital Economy & Trade University Publishing House, Beijing, 2003.
[5] Y.X. Zhang, "Fixed Exchange Rate System and Finance Safe", Journal of Haerbin College, Jan, 2003, 46-51.
[6] C. Montet, D Serra, Game Theory & Economics, Economy & Management Publishing House, Beijing, 2005.
[7] F. Gu, The Financial Big Shot Magnate Soros, China Economy Publishing House, Beijing, 1998.
[8] H.J. Yao, Game Theory Analysis of Economic Policy, Economy & Management Publishing House, Beijing, 2001.
[9] Y.G. Zhu, "Try Analyzing the Choice of the System of Exchange Rate of Our Country", Journal of Chongqing Technology and Business University, Jan 2003, 72-75.
[10] H.J. Zhang, "Restrain in the Route of China's Exchange Rate Reform in Transformation Period", Journal of Contemporary Finance and Economics, No.228, Nov 2003, 44-49.

Applied Research on Factor Analysis in Arbitrage Pricing Theory

Jianmin Sun, Pin Wang
Economics and Management School, Wuhan University
Wuhan, Hubei 430072, P.R.China
Email: alexsun2000@163.com **Email:** wp_whu@126.com

ABSTRACT

The biggest challenge of Arbitrage Pricing Theory（APT）is how to screen out the factors that need too much computation. Factor analysis abstracts a few factors from a large number of original variables. Through the factor analysis, this article identifies four factors out of eleven original variables containing most information about those eleven variables. Using the four factors, this article builds a pricing model that makes a better effect in pricing through empirical analyses.

Keywords: Arbitrage Pricing Theory (APT), Factor Analysis, Portfolio, Bartlett Ball Test, KMO Test.

1. INTRODUCTION

Stephen Ross proposed Arbitrage Pricing Theory (APT hereafter) in 1976, which is built on the assumption that the yield of any risky asset is influenced by *K* factors. We use a K-factor linear model as follows:

$$\tilde{r}_i = a_i + \sum_{k=1}^{K} b_{ik}\tilde{f}_k + \tilde{\varepsilon}_i, \quad i = 1,2,\cdots,n \qquad (1)$$

Here $E(\tilde{\varepsilon}_i) = E(\tilde{f}_k) = E(\tilde{\varepsilon}_i\tilde{\varepsilon}_j) = E(\tilde{\varepsilon}_i\tilde{f}_k) = 0$; $E(\tilde{\varepsilon}_i^2) = s_i^2 < S^2$.

$\tilde{r}_i$ is the returns of risky asset i; a_i is the average returns of risky asset when all the factors that may influence the returns of risky assets are zero. b_{ik} is the sensitivity of risky asset i about factor k, $\tilde{\varepsilon}_i$ is disturbance.

When there is no asymptotical arbitrage opportunity, we can receive the following pricing model -- APT approximately by K-factor linear model:

$$E[\tilde{r}_i] = a_i \approx \lambda_0 + \sum_{k=1}^{K} b_{ik}\lambda_k \qquad (2)$$

λ_k is the risk premium of risky asset i about factor k. If we take the error term as $v_i \equiv a_i - \lambda_0 - \sum_{k=1}^{K} b_{ik}\lambda_k$, we have $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} v_i^2 = 0$, when there is no asymptotical arbitrage opportunity.

The essence of APT lies in the selection of the factor. However, the return of a risky asset is under the influence of other factors and we do not clearly know how many factors we should choose [1]. Suppose the return of one risky asset is under the influence of *n* factors. There are $\sum_{m=1}^{n} C_n^m$ choices. Selecting the superior factor combination will be a huge effort.

This article employs the factorization method to select APT factors. We can reduce the number of the factors while preserving the information in original index variables and greatly reducing the complexity of the problem.

2. ASCERTAIN FACTOR PORTFOLIO

There are at least three types of different factors of APT: the index, which reflects the total economic activities, the inflation rate, and the interest rate factor of some types. We take GNP, gross industry product, gross second industry output, gross tertiary industry output, national consumption level, inflation rate, growth rate of the social fixed assets investment, gross social consuming sales, total currency supply, total net export per year, and interest rate structure as original index variables. The analysis referenced the data of "Year Book" from 1980 to 2000 in China.

2.1 Correlation analysis of the original variables

Factor analysis constructs a few representative common variables, which have strong correlation with the original index variables. We use KMO (Kaiser-Meyer-Olkin) test and Bartlett ball test. The numerical value of KMO Statistic is 0.771>0.6, original index variables adapt to factor analysis according to the statistician Kaiser's standard. The Statistic value of Bartlett ball test is approximately 780.924 with CHAID -χ2 degree of freedom 55; furthermore, concomitant probability is lower than significant level 0.05, which is deemed to be suitable for factor analysis. All the absolute values of coefficients of partial correlation are under 0.05, which means that all the original index variables adapt to factor analysis in the anti-reflection related matrixes.

2.2 Constructing the common factor variables

Here we use the principal component analysis.

1) Ascertain the quantity of common factor variables

According to the scatter chart of common factor variables and their eigenvalues, the first four common factor variables vary largely from 8.744 to 0.126. But from the fifth, the eigenvalues are within a small range, which indicates that the first four common factor variables work well on the description of the information. In order to obtain a more accurate APT, this paper makes sure to reserves 4 common factor variables [2,3].

2) Results of factor analysis (table omitted)

Factor analysis explains 99.5% variance of each common factor variable. The common degree of each common factor variable is above 98%.The information included in the original index variables cannot be explained by common factor variables below 2%. It shows that though some information is dropped, the common factor variables can reflect most of the information in the original index variables.

3) The result of factor distilling and revolving (table omitted)

According to the scatter diagram, this article distills four public factor variables to describe all the original target variables. The variance contribution of 4 public factors is 8.744, 1.348, 0.729, and 0.126. We know that the first four public factor variables explain 99.522% of the population

variance of target variable, especially the first public factor variable, which explains 79.490% of all population variance.

After revolving the factors, the eigenvalues of these four public factor variables are 8.444, 1.231, 1.102, and 0.171, which can explain 76.192%, 11.194%, 10.014% and 1.552% of the original target variable, and explain 99.522% of the whole population variance. These four public factor variables obviously have almost reflected all the information of the original target variable, and can replace the 11 original target variables in constructing the multi-factors linear model [4].

2.3 Naming explanation of the public factor variable

All the original target variables vary with concrete economic meaning. After analyzing the principal components, we make linear substitution on these variables, and obtain the new public factor variable, which has synthesized and simplified the original target variable. So it is necessary to explain the primary factor and the system characteristic that affect the constitution of the original target variable system.

1) The factor load matrix of pre-revolving and post-revolving

After analyzing these factors, we find out that before factor revolving, except these three original target variables, inflation rate, growth rate of the social fixed asset investment and interest rate structure, the other eight that reflect a national overall economic situation have high load with respect to the first public factor variable, with absolute value above 0.9; gross net import per year, inflation rate, growth rate of the social fixed asset investment and interest rate structure have high load with respect to the second public factor variable [5]; inflation rate, growth rate of the social fixed asset investment have high load with respect to the third public factor variable; inflation rate, gross net import per year and the interest rate structure have high load with respect to the fourth public factor. Clearly, some information superposes in the four factors. A public factor variable has high load in many original target variables, which is not convenient to explain the public factor variable and to find the economic meaning of each public factor. So it is necessary to make adjustment on this model. This article used the maximum variance method to involve the factor load matrix.

The first public factor variable reflects the information of eight original target variables that reflect national overall economic level, that is, gross social consuming sales, gross industry output, gross second industry output, the national consumption level, GNP, gross tertiary industry output, gross currency supply and gross net import per year; the second reflects the related information of the inflation rate; the third shows the information of the growth rate of the social fixed asset investment; The fourth reflects the situation of the interest rate structure. Each original target variable belongs to some of the four public factor variables.

2) Naming explanation of public factor variable

The first public factor variable reflects national overall economic level; the second public factor variable, the inflation rate; the third public factor variable, the growth rate of the social fixed asset investment; the fourth public factor variable, the interest rate structure, which shows the expected return ratio of the risk securities is mainly related to the national overall economic level, the domestic inflation rate, the growth rate of the social fixed asset investment and the interest rate term structure.

The statistical analysis shows that these four public factor variables are all random variables with average 0 and variance 1, uncorrelated with each other. Therefore, they are very suitably taken as APT factor [6].

3. TEST

Through the analysis above, we have identified four public factor variables and calculated the corresponding value from the 11 original target variables during 1980-2000 according to their scored points. This article selects, according to formula (1), ten stocks such as Huayuan pharmaceutical and takes these 1992-2000 annual return ratio as explained variable, the corresponding 1992-2000 annual data of the four public factor variables as explaining variables, and obtains each stock $a_i, b_{i1}, b_{i2}, b_{i3}, b_{i4}$; According to formula (2), we take a_i as explained variable, $b_{ik}(k=1,2,3,4)$ as explaining variable for multi-variable linear return and obtains the APT model:

$$a_i = 0.189 - 1.051\,\lambda_1 + 0.02067\,\lambda_2 - 0.0233\,\lambda_3 - 0.238\,\lambda_4 \quad (3)$$

Moreover, this article selects 10 stocks such as Yuyuan mall, and uses formula (3) to price them, then takes the difference between the forecast average return ratio of the model and the actual value as the evaluation standard of the model. Since APT is an approximate pricing model and there will be considerable error while applied in individual stock, it is often used in pricing investment portfolio. Therefore this article constructs a simple average-weight portfolio of the ten above stocks and uses $y = \frac{1}{n}\sum_{i=1}^{n} (\bar{r}_{ipredict} - \bar{r}_{ireal})^2$ to measure error. The result is $y = 0.129$.

The test indicates that the APT model (3) obtained has good results in pricing, while there still exists 12.9% pricing error. The explanation is as follows.

China started statistics of macroeconomic indicator late, much data are incomplete and statistical standard is inconsistent. There were only annual data prior to 2000, which results in the lack of sample data. Since the macroeconomic indicator uses annual data, the stock return ratio also has to use it. But China's stock market has existed for only 12 year. When constructing APT, this article only selects the stocks in order to enlarge sample size as much as possible, which results in the small quantity, the unitary category of stocks, the considerable error of the forecast result and the possibility of low significance and goodness of fitting in regression equation. It is believed we can further reduce the pricing error of APT given better sample data.

4. CONCLUSION

Through the introduction of factor analysis, this article synthesizes and simplifies 11 factors, such as GNP, the inflation rate and selects four public factors that reflect the state's overall economic level, the inflation rate, growth rate of the social fixed asset investment and the interest rate. The related statistical analysis shows that four public factor variables selected by factor analysis can reflect most of the original information. Using the four factors, we build an APT, which improves the pricing model by the empirical analysis.

5. REFERENCES

[1] B. Lehmann, D. Modest, “The Empirical Foundations of the Arbitrage Pricing Theory”, Journal of Financial Economics, 1988, Vol.21, 213-254.
[2] E. Fama, Y. Dividend, et al, “Journal of Financial Economics”, 1988, Vol.22: 3-27.
[3] N. Chen, R. Roll, et al, “Economic Forces and the Stock Market, Journal of Business”, 1986, Vol.59: 383-403.
[4] Robert A. Jarrow, Finance theory, Prentice-Hall, 1988
[5] Y.C. John, “The Econometrics of Financial Markers”, Princeton University Press, 1998.
[6] J.E. Ingersoll, “Theory of Financial Decision Making”, Rowman & Littlefield, 1987.

Structuring Method and Key Technology Study on Component- based Comprehensive E-commerce Platform

Xiangzong Wang, Zengliang Liu
Information Engineering School, University of Science & technology Beijing, Beijing, China
Email: wangxz@sinosoftgroup.com, chinaphd@tom.com
Chen Ying
Software School of Hunan University, Changsha, Hunan, China
Email: sprideagle@yahoo.com.cn

ABSTRACT

Basing on the idea of the traditional bazaar and the demanding characteristic of comprehensive e-commerce platform business, a construction method of component-based comprehensive e-commerce platform was put forward from the perspective of the distributed applications. The strategy, idea and principle of the method were also systematically introduced. And the paper provides some key designs in comprehensive e-commerce platform, such as physical framework, the functional logic framework, technical framework and so on. Finally, with an example of ERP gateway, the developing procedure and some key algorithms basing on visual component platform was introduced. The construction method offers a kind of new concept and important reference for practice to explore the comprehensive e-commerce platform.

Keywords: Bazaar Concept, Component, Comprehensive E-Commerce Platform, Framework Modeling, Distributed Applications.

1. INTRODUCTION

With the rapid development of the network technology of the computer, the e-commerce concept is well-known for everyone, but because of the demanding characteristic of higher security and complexity of the trade, e-commerce still stay at information issue, EDI, BtoC level. As the major mode, BtoB way of the trade is still kept at the concept or the trade information processing stage. The implementation method of the real e-commerce urgently demands further probe.

According to dealers' different social roles, there are mainly two kinds of e-commerce modes: BtoC (Business to Consumer), BtoB (Business to Business). Other modes can be regarded as extended from the two main modes, such as BtoG (Business to Government), GtoC (Government to Consumer) etc.

There are various styles about the e-commerce platform structured on the basis of BtoC and BtoB mode. The typical "BtoC" mode of e-commerce platform is generally the type of “one to many", being selling-oriented and incorporation of the products and services trade etc. At present, it has been already quite mature in both technical method and practical application [1].

Because of the limit of historical conditions, there were many kinds of interim forms for "BtoB" mode e-commerce platform in different periods resulting from practical need, such as "one to many" mode, "many to many" mode, purchase-oriented, selling-oriented, products trade, service trade, area type, industry type etc. Finally, according to the natural demand characteristic of e-commerce, the e-commerce platform will develop into a type of form not only with “many to many" mode, but also incorporating product trade and service trade. It contained all functions of other platforms, and it is also applicable for BtoC mode. Therefore, it is sophisticated in technology, difficult in construction and expensive in cost, and becomes one of the research focuses in the field of e-commerce platform at present. So the paper mainly researches on the construction concept and implementation method of such comprehensive e-commerce platform in theory and practice.

2. PRINCIPLE OF STRUCTURING METHOD FOR COMPONENT-BASED COMPREHENSIVE E-COMMERCE PLATFORM

2.1. Question analysis

The development of e-commerce platform is mainly divided into following three stages: At the first stage, it implements whole transaction by off-the-net negotiation and off-the-net payment. The focuses are to set up the information flow platform, and to take information processing, and to carry out information provision or information exchange of trade products or service. The on-line functions include EDI, product exhibition, materials recommending, propaganda, advertisement, etc. At the second stage, its focuses are mainly to realize the online electronic trade and logistics of BtoC mode by dint of one's own or third party's logistics platform, and to set up BtoC type of money flow management platform by means of the bank payment system. At this stage, because of the high security and complexity demands of BtoB mode, the platform can only provide some functions such as the information of supply and demand, online negotiation, online contract signing etc. It fulfils the trade via off –the-net payment. At the third stage, it focuses on online payment and set up good protecting circumstances of e-commerce and constructed the comprehensive type of "virtual electronic trade market" and carry out the online electronic trade in an all-round way.

At present, the development of the e-commerce is the transitional period between the second stage and the third one. It seems that "online payment" were the only key factor influencing realization of the electronic trade of BtoB mode. In fact, due to its inherent complicated characteristic, there exist a lot of profound problems difficult to solve as follows:

1) Security: Because the quantity of the exchanged goods and the amount of money involved in trade are large, it demands more security than BtoC mode.

2) Sociality: The trade is a kind of important social activities, which needs to comply with certain social norms. Compared with BtoC mode, BtoB mode needs to

set up more sound business insurance system in laws and regulations, industry and commerce, revenue, bank, customs, security, credit, etc. Therefore it is not merely a technological puzzle, but also a social issue. And it needs the government's full support and participation. In a sense, the influencing factors of implementation of BtoC mode e-commerce platform is mainly the technological puzzle, but the implementation of BtoB mode e-commerce platform will meet both technological puzzle and social issue.

3) Collaboration: Generally, BtoB mode transfers material and value in the whole supplying chain, which is of great impact to the production and operation to enterprises, so it requires the cooperation and coordination among trading parties, the third party in the logistics service, and all other parties to ensure the security of the circumstance to involve.

4) Frequent changeability of the platform structure: The structure of comprehensive e-commerce platform is dynamic changeable for relating to multi-sides trade and management. On one hand, it needs to insert or delete some back-stage management information of some special trade participants (large-scale enterprises, trade security system, etc.) who become one of the organic components of the platform. On the other hand, because of dynamic change of quantity and geographical distribution of dealers, platform backbone's nodes need to be increased or reduced. Consequently, it causes the platform structure to change frequently.

5) Complexity of resource management: Because the supplies or service involved in the trade vary with the participant's dynamic change, which makes more complexity in resource management platform, it needs coordinated real-time management to the resource information involving administrator and all unknown trade participants. But BtoC mode platform, with characteristic of "one to many", only requires that the platform administrator manage its own resources well, therefore the resource structure is relatively steady.

6) Higher requirement in distributing treatment: All kinds of types of the e-commerce platforms are set up on the basis of heterogeneous distributed processing. And it is satisfied for the platform of "BtoC" mode and "BtoB" mode mainly dealing with information processing to consider its own platform distribution. While the e-commerce platform with the whole virtual electronic trade course needs to do the distribution processing to the ensure information of electronic trade circumstances and large-scale special trader's information, namely distribution processing with more complicated business involved in the trade course.

2.2 Strategy

The brief definition of key element of "traditional bazaar" is as follows: There is an area for market, and there are many product shops, many service shops, the purchasing traders gather at random. There are many banks, insurance system for offering the financial guarantee of the trade. There are traffic police, public security men, etc. whom guarantee trade's security along the street. There’re industry and commerce bureau, price bureau, tax administrations, etc. whom maintain market order, and so on. These key elements form a complete traditional trade bazaar.

Based on the tradition idea of " bazaar ", the main idea and the key elements of the comprehensive e-commerce platform structure are proposed as follows: (1). Electronic network platform - -virtual trade place. (2). Adopting "many to many" mode--platform with many buyers and sellers. (3). Products trade and service trade coexist. (4). With the aid of the standard norm of platform management, it realizes the unified resource management. (5). With the aid of gateway, it realizes information merges among the large-scale special enterprises. (6). To set up third party's authentication organization to guarantee trade's security. (7). With the aid of gateway, it realizes the online bank. (8). With the aid of gateway, it realizes link with information platform of industry and commerce, tax, customs, etc. (9). To merge the business rule, laws and regulations of e-commerce etc. into information processing business and procedure, and offer the guarantee in trade security, etc.

These have succeeded in solving security, sociality, coordination and complexity resource management issue that exist in the comprehensive e-commerce platform. It also solves the problem of “higher requirement in distributing treatment" by multi-stage distribution processing. And it also solves problem of "frequent elasticity of the platform structure" and some other complicated problems of implementation, application and maintenance etc. by means of the visual component-based assembling platform to carry out the comprehensive e-commerce platform.

Therefore the overall maneuver of the comprehensive e-commerce platform is as follows: The technological circumstance of the comprehensive e-commerce platform was set up on the basis of the idea of traditional bazaar by dint of the modern computer and network technology and the method of component-based development. Driven by participant's interest, it sets up e-commerce operational circumstance with safety certification, online payment, logistic system, etc. To set up e-commerce insurance circumstance supported by policies, regulations, standard system, credit system, etc. [2]

2.2. Overall concept

Based on the above-mentioned construction strategies, a kind of platform implementation method of comprehensive component-based e-commerce is proposed. The main idea is: "the prototype system (field products) +the mature products components (COTS) + the visual platform for customizing (being made up development, assemble, integration etc.)". First of all, a “prototype system " is formed in the formula by means of the Domain Engineering. It is the core of the whole system including the system domain architecture model (such as the systematic operating platform, the application server, the basic application supporting platform, the common field application platform, the domain architecture, etc.), the architecture storehouse, the component repository etc. Secondly, while building concrete e-commerce platform, it is divided into foregone part in " prototype system "and the Commercial Off-The-Shelf (COTS) part and the part needed to be customized after analyzing modeling and system demand modeling basing on the prototype system. The COTS generally refers to the mature products that can be bought in the market, and it will be system-grade or sub-system component after packing and it doesn’t include non-application parts that was selected as the systematic supporting part of the prototype when forming the prototype system. The visual customizing platform offers

the component-customized according to needs for application system. The forming ways of all kinds of components include customized developing, assembling, integrating etc. As space is limited, only the key realization technology of the component-based developing method and principle is described briefly in Fig.1..

It can be seen from Fig. 1., there're four stages: business modeling, demand modeling, design, realization in the developing process based on component.

At business modeling stage, the model of the senior level business of e-commerce was set up according to the business demand characteristic of the comprehensive e-commerce platform and the analyzing method of the domain engineering. And the use cases on the senior level business were set up with focusing on interface and activity.

The purpose of demand modeling is to provide the systematic physics framework, logic framework of overall function. It divides and maps the type model of the business into the business component, atom business component according to the criteria for classification of component and based on the business model.

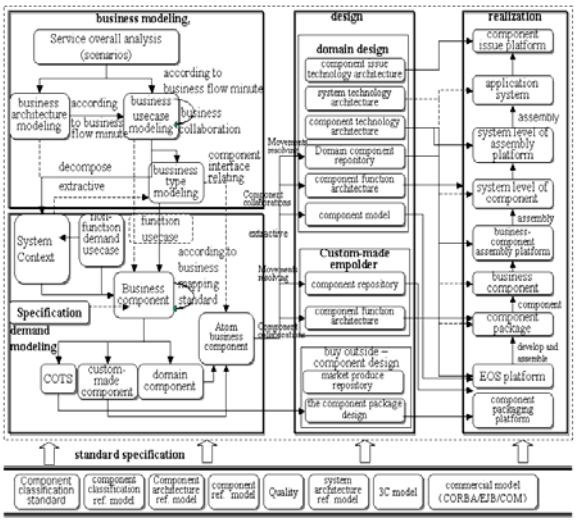


Fig. 1. The method principal and procedure of the component-based e-commerce platform

At the phase of analysis and design, on the basis of the demand model and mature visual component platform, to set up the whole system framework according to component framework technology and method, and design the component model, map atom business component into atom component that may be realized or already existed in the component platform.

At the realization stage, all kinds of components including the application system are developed or assembled or tested with the mature visual component developing and assembling platform.

3. THE REALIZATION FRAMEWORK OF COMPREHENSIVE E-COMMERCE PLATFORM

3.1. Physics framework models

The platform physics framework of the comprehensive e-commerce is shown as Fig.2 including e-commerce circumstance, transaction service platform, socialized logistics platform, e-commerce unified resource management, e-commerce basic supporting platform, gateway service platform etc. [3]

1) The basic e-commerce supporting platform is the foundation of e-commerce platform, including computer network platform, system operating platform, application server platform, visual assembling and developing platform, etc.

2) The resource management platform of e-commerce is used for carrying on unified modeling and management of all kinds of resources involved in e-commerce, and to assure the share of resources as much as possible and the good adaptability to the changing management style and the resource types, thus to meet the frequent changeability of platform structure of e-commerce. The main function of the resource management platform includes the resource modeling, using, maintenance etc., and logistics controlling, products information management, other resource management basis on the resource platform.

3) The trade platform is a key subject of e-commerce construction, including a series of application software functions such as all kinds of electronic trade, market management and controlling, trade service etc.

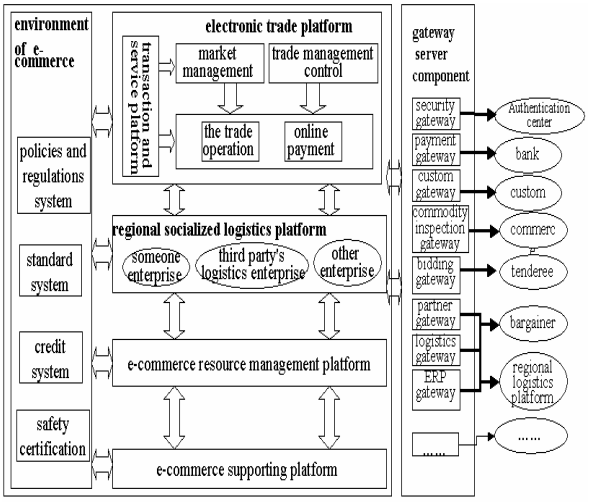


Fig. 2. The physics framework of the comprehensive e-commerce platform

4) The logistics service platform is the foundation of e-commerce platform construction, mainly including the inner logistics platform of the enterprises and regional socialization logistics platform. "Internal logistics" is inside an enterprise. And it is the e-commerce foundation in the electronic trade. "Regional socialization supply logistics" is in a certain area, to integrate local typical manufacturing companies, third party logistics enterprises, enterprises in circulation, etc. into comprehensive e-commerce platform by service gateway mode or client-terminal mode etc. The logistics service platform involves the materialization platform (transport the facility, port, storage, load and unload etc.), manage platform, logistics information platform (management software system of logistics, GPS, GIS in logistics etc.) and so on.

5) E-commerce circumstance construction is the prerequisite for the smooth implementation of e-commerce platform. It will dispel a series of obstacles in the security, credit etc. that the E-commerce brings to

the traders, and promote the wide application of e-commerce by means of government.

6) The service platform of the gateway is mainly to facilitate internal logistics system including all kinds of third parties for service, producers, commerce enterprise, along with banks, customs, storage center, distribution and delivering etc. to seamlessly link to the e-commerce platform. It set up the information channel and carries out automatic connection of the business and information between the e-commerce platform and internal information platform of all kinds of other parties.

3.2 Functional framework model

The functional framework of the comprehensive e-commerce platform is shown as Fig.3.. It mainly includes the trade operation management system, trade management control system, transaction service system, service gateway system, market supervising and administrating system, and online payment system in transaction. The concrete function includes: With the support of the modern logistics system platform, purchase from global supply, sale to global market, electronic warehouse, distributing and delivery etc. and to carry out settlement function, such as online payment, check account etc. and to realize the market management such as the identity authentication, the products authentication, credit authentication, information issue, custom complain etc., and authority management, user's management, credit control, as quality control, providing and delivering control and other market control function. Meanwhile, to offer a series of E-business service tools such as E-MAIL, message board, online auction, advertising service, etc., thus to realize the integrated application of logistics, funds and information flow of the enterprises and institutions on the e-commerce platform. [4]

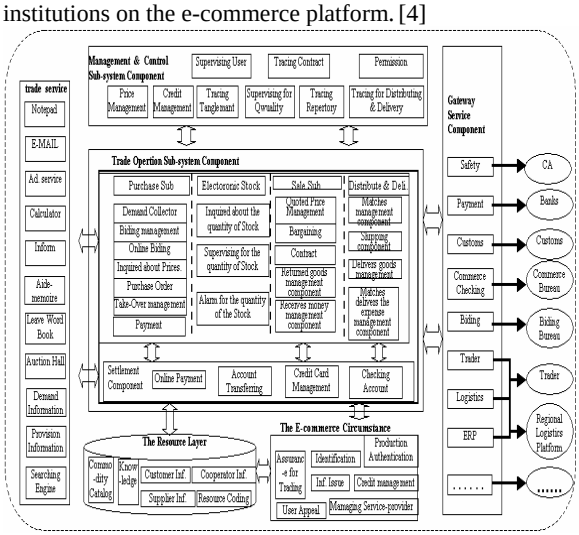


Fig. 3. The Framework of the Function Component of the Comprehensive e-Commerce Platform

Transaction operational management sub-system component fulfils various kinds of commercial activities in the commercial platform and the automatic execution of different auxiliary service activities, and assure the continuity and validity of online commercial activities procedure, which is supervised by its upper strata trade controlling sub-system during performance. The trade target includes product and service. And the trade type includes purchasing and selling.

The trade management control sub-system component fulfils the all-round controlling and supervision during the trade process. The authorized dealers could check the implementation and ordering processing results.

Market management and supervision sub-system component administrate the all kinds of trade resource and services provided by platform. It sets up and supervises to implement the trade rule and the distribution mechanism of all kinds of traders' responsibility, rights, and profit.

Online payment sub-system component is the main body to realize capital flow in the e-commerce platform. The main functions include online payment, account transferring, checking account, credit card management, etc. The business or information flow of the online payment, the account transferring, the management of credit card between the banking system and e-commerce platform are connected by the bank gateway, which submit payment forms to the bank system in way of POST. [5]

3.3 Technological framework model

The technology involved in construction of the e-commerce platform is complicated and various, including communication technology of the network, Web distribution computing technology, EC security technology, payment gateway technology, processing technology of the intelligent information, ERP gateway technology, etc. In order to integrate different technologies in one, the technological framework for e-commerce platform is shown as Fig.4. by referring to the distributed systematic framework model. It is divided into six layers: Client layer (to realize interaction between user and platform, including richer-client and browser), display layer (mainly in WEB way, to execute the procedure of script or dynamic page in the server. It is offered by WEB server provider at this layers of service), service layer (serve layer mainly used to the interaction between domain layer and display layer, offers all kinds of component gateway, such as security gateway, payment gateway, custom gateway, provision and delivery gateway, partner gateway, commodity inspection gateway, bidding gateway, storage gateway, etc.), domain layer (can be also regarded as at business logic layer, include e-commerce circumstance, trade service platform, regional socialized logistics platform, resource management platform of e-commerce, e-commerce supporting platform), data mapping layer(mainly to produce the data component and operation component), data endurable layer (to carry out all kinds of operation and processing in the database system). [6]

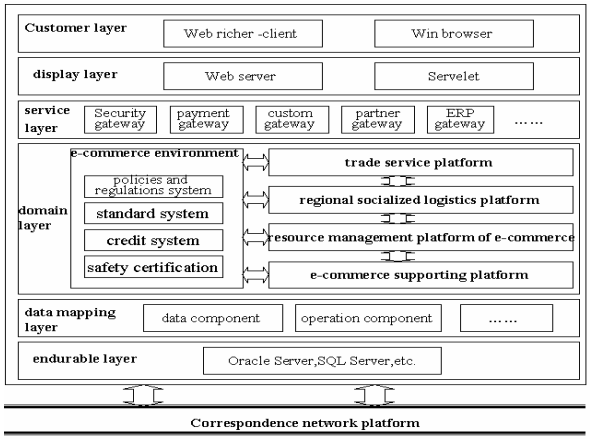


Fig. 4. Component-based e-commerce platform overall technology skeleton chart

4. AN EXAMPLE——IMPLEMENTATION OF ERP GATEWAY

4.1. Technological frame

ERP gateway is used to connect the e-commerce platform and special large-scale enterprise ERP system involved in trade by end-to-end mode. The structure is shown in Fig.5.

1) The business interface component in e-commerce platform and ERP system: It is used to transfer the business procedure and data for every special business, at the same time, to carry out the component assembling based on XML by end-to-end mode.

2) The gateway warehouse for disposes: It is used to offer the integrated service of the gateway, and types of interface information stored in ERP and e-commerce, and the accessing methods for interface. The integrated services of the gateway deal with real-time transferring or service processing through the current account and its configuration information.

3) Interface service layer: This level is used to encapsulate gateway interface service for the business interface in e-commerce or ERP so that the business interface can be accepted by real-time request agent and gateway processing server. In this part it reads the information from the configuration warehouse and transfers into the serving info-package that can be accepted by real-time request agent and gateway processing server, and send the info-package to request agent and gateway processing server.

4) Real-time request agent/gateway service processing. It mainly includes real-time request agent (to be responsible for the request from e-commerce or ERP interface service), gateway server (dealing with some non-real-time business, mainly including the interaction between ERP and e-commerce. It is often backstage running thread).

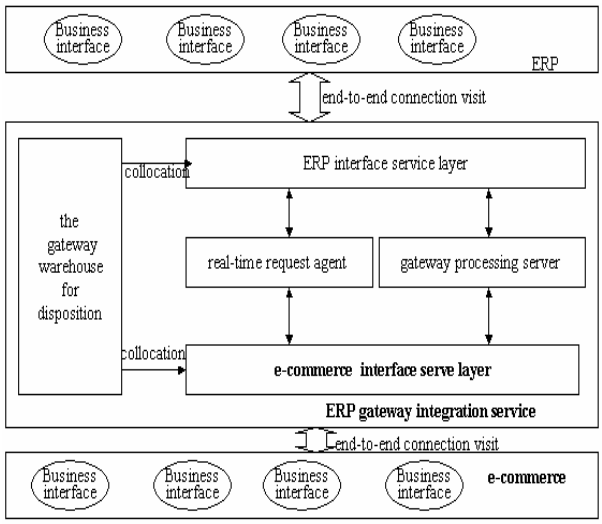


Fig. 5. The frame of the EPR technologic gateway

4.2. Key procedure

On the basis of the mature visual component platform EOS of Primeton shows the ERP gateway component as in Fig.6. Each component unit represents one independent component system including some independent sub-components or sub-component package. [7]

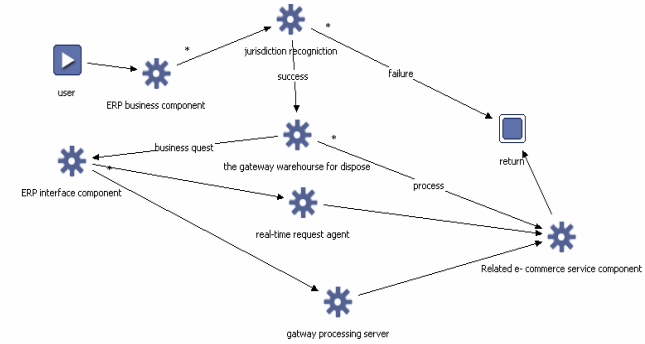


Fig. 6. EPR gateway component elements chart

Taking the user-requested sub-component as example, the component elements of the sub-component are further illustrated in Fig.7..

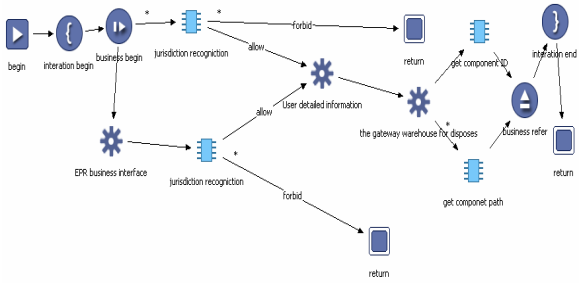


Fig. 7. the visual assembling flow chart for the user—requested component

Part of the realization algorithms:

String CompanName; //Component name

String CompanPaht; //Component visiting address

String CompID //Component id that the users requested

Boolean UserLevel = false; //The initial user's

```
authority is false
    /*According to users' id and component id  if
there's limited authorities of component /
    Void ERPUser (int  userID, String CompID)
    {
        //Through ERP resources to look for this user's
authority
        If (userID exists)
        {
        UserLevel = true;
        }
        else
            {
        UserLevel = false;
        }
    ……
    }
    /*******Over ********/

    /*to get the detailed information of a component,
from warehouse lasting according to component iD
    void getCompanDetail(Stirng CompanID)
      {
      Find this component from the gateway
configuration storehouse
            {
            getCompanXMLData (String CompID)
            }
            else
            {
            If request agent
                {
                getCompanXMLRequestDate(String
CompID)
                }
                Else if turns off and serves the processor
through the network
                {
                getCompanXMLInterfaceDate(String
CompID)
                }
                else
                {
                // failure
                }
            }
      }
……
```

5. CONCLUSIONS

In this paper, aiming at the key problems to construct comprehensive e-commerce platform, the overall concept, principle, method and platform frames are proposed with component-based method and the solution strategy of distribution processing in all kinds of gateways basing on the traditional bazaar idea. And as an example, based on the mature visual component platform, the paper fulfils the ERP gateway by the method. It has been provided a new method and concept in theory and practice for comprehensive e-commerce platform construction. This method has been successfully applied in real design and development of e-commerce platform.

6. REFERENCES

[1] A. continent, high Wei, "The distributed E-business service supplies the original language set up the course of the chain to describe", Computer project, 31(5), March 2005.

[2] L.Q. Huang, "Change of the component of the software. Computer World", The 12th issue. E9. http://www.cn-doc.com/_soft_ect_doc/2005_08_18_23/20050818234122272.htm.

[3] Q. Yong , M.L .Zhao medium-sized in season. "Use the frame to study and realize on the basis of the multi-layer mode of distributing component technology", Computer project, 27(1), June 2001.

[4] Clemens Szyperski Dominik Gruntz Stephan Murer, Component Software: Beyond Object-Oriented Programming, Beijing: Publishing House of Electronics Industry, 2004.

[5] C. John, D. John, UML Components: A Simple Process for Specifying Component-Based Software, Reading: Addison-Wesley, 2001.

[6] C. Guareis, M. van, et al, "A combined approach for component-based software design". In: 3rd Intl.Symposium on Distrbuted Objects and Applications (DOA' 01) [C],2001.

[7] Primeton Technologies, Ltd. EOS5.0 white paper. http://gocom.primeton.com/modules/devcenter/article191.htm, August 2005.

Research On Simulating Optimization Of Distributed Sharing Information Model For Food Supply Chain

Wei Mao[1], Fengtian Zheng [1], Jia Mao [2]
School of Agricultural Economics and Rural Development, Renmin University of China, Beijing, 100872, China[1]
College of Transportation, Jilin University, Changchun, 130022, P.R.China[2]
E-mail: JLmw1@163.com

ABSTRACT

A Synchronization protocol for parallel discrete event simulation is now one of the main research fields' discrete simulation community, its core duty is highly effective management information flow of supply chain. Aiming at food safety of supply chain management work's efficiency was lower; Realizing information resource sharing is an effective way to improve efficiency of supply chain management. The information management system in supply chain was simplified into a discrete parallel system simulation model and a new effective strategy to prevent system deadlock was examined in depth, simulation output result shows that the distributed simulation approach introduced improves computation efficiency meets high efficiency and real time operation systems, according to goal demand real-time, dynamic reasonable dispatch system resource.

Keywords: Sharing Information Model, Simulation, Supply Chain, Food Safety.

1. INTRODUCTION

E-business time forthcoming, information technology widespread application, for the supply chain management which was on the rise brings new turning point, E-business already becomes the essential technical factor in many industrial supply chain conformity. Facing E-business speedy and violent development, how to utilize enterprise E-business technology optimization supply chain management, completes in the enterprise inside check run integration foundation, which extend to the enterprise outside, logistics flow, information flow and fund flow becomes hot topic which the academic circles studies. One kind of original theory arise at the E-business supply chain management, its center duty is highly effective management the information flow of supply chain, realization supply chain worth through the information sharing. This paper introduce supply chain management thought to food enterprise, which uses E-business method that conducted research food supply chain information sharing model take food supply chain as the example.

At the present time information sharing question fundamental research in food supply chain management hot spot mainly concentrates in supply chain information sharing importance, necessary and realization information sharing technical method research. How to uses advanced information technology, which through method improvement and so on information sharing flows management efficiency in the countermeasure research foundation, this paper proposed corresponding model analysis method, constructs important topic which faces based on food supply chain integration management information system.

2. RESEARCH ON DISTRIBUTED SHARING INFORMATION MODEL FOR FOOD SUPPLY CHAIN

A distributed system is a collected works of processors that do not share memory or a clock. A distributed system gives its users entrance to variety of resources that it maintains. Access to shared resource allows computation speedup, as well as improved data availability. The developing direction of high-speed computer network is high bandwidth, low delay, and to offer possibility for parallel processing under distributed circumstances. Information integration technology is one of the hotpots in parallel-processing field. It is necessary to research the algorithm of parallel task distribution of enterprise under distributed circumstances for satisfying requirement to capacity of parallel-processing system.

Food supply chains are formed different organizing carriers and links. Conformity management through food supplies chain, which effectively improve efficiency and service level of food logistics. Through controlling of food safety, ensure favorable running quality system of food safety. The supply chain management reflects the new trend of based on Internet, at same time more request for information sharing is put forward. With the development and popularization of the network technologies, how can the food supply chain management information being shared?

$$I_T = I_C + \sum_{i=1}^{m} \Delta I_i \qquad (1)$$

I_T Express sum total information amount that is utilized by enterprises of food supply chain. I_c Represents food supply chain core enterprises self provide information number. $\sum_{i=1}^{m} \Delta I_i$ Express sharing information amount is providing by other enterprises in food supply chain.

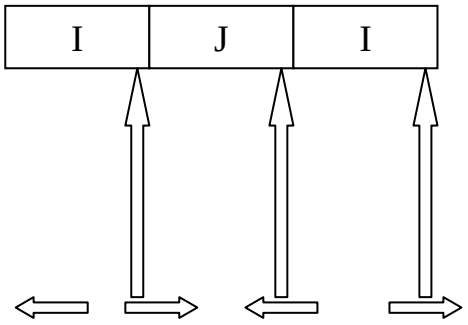
3. OPTIMIZATION STUDY ON SHARING INFORMATION MODEL

The simulate model of information sharing and uses near-optimal enforcement means of task scheduling to simplify logistics information management system into an emulate model of the discrete and parallel system and optimized strategies so that utilization rate of resources in the logistics network and execution efficiency of the concurrent process is improved.

It is significant that the task execution time be predictable to

allow designer to figure out if all urgency will meet their deadline. When several processes compete for a limited number of resources, a situation may arise a process requests a resource and the resource is not available at that time, in which case the process enters a wait state. It may happen that waiting processes will never again change state, because other waiting processes hold the processes that they have requests. This situation is called a deadlock. As follows chart 1showing:

If a process enters into a interface, and while also ready to enter another interface J, it is very possible to fail to enter into it, that means the needful resources can't be obtain, but have to wait in the queue of L, If the process that has occupied the resource is rightly waiting in the queue of L will happen. Because this time exits two processes, neither can advance forward nor reach the step of releasing resource. Show as Fig. 1..



$TASK_1$ $TASK_2$
I preempted by J J resumes I completes

Fig. 1. Queue of process I by process J

Regarding these two kinds according differently lines up the rule correspondence model, ask two kind of customers' standby periods, the response time and party chief's steady distribution

$$\beta_i = \int_0^\infty x dB_i(t)$$

$$\beta_i^{(2)} = \int_0^\infty x^2 dB_i(t) \qquad (2)$$

$$B_i^*(S) = \int_0^\infty e^{-st} dB_i(t) \qquad i = 1,2$$

Ordered resource allocation method by prevention effective dead- lock was put forward by Havender [1, 2]. A real-time computer system is usually make up by the terminal and explicit apparatus, communication network,, many way converter, CPU, magnetic drum and magnetic disc, outside store the apparatus and corresponding processing procedure.

Each process ask for resource according to order that increase progressively strictly, a request resource behind the I process which can only ask for serial number arrange resource those behind, that only request get satisfied at present, could apply for the next resource J . Carry on the following improvement to method of ordered resource allocation: Keep all resource of operating position go on and in the past among memory already the system resource operating position recorded, which ask corresponding weighting of item arithmetic mean value namely new resource frequency utilization form. New resource frequency utilization form not only contains the account of the history but also contain the recent operating position. The following ordered resource allocation j resource calculation formula of the frequency utilization when systems *(n+1)* start once:

$$C_{n+1} = w_1 C_o + w_2 C_n \qquad (3)$$

C_{n+1} express system time start stylish systematic resource frequency utilization *j* number value resource form, C_o express system *n* systematic resource frequency utilization *i* number value resource of the form when the times of start, C_n express system *n* systematic resource frequency utilization *i* application number resource of the form in the times of operation, w_1, w_2 are the weight factor. The formula (3) can be varied into the following form:

$$C'_{n+1} = C_o + w_{2/1} C_n$$

$$\begin{cases} w_1 + w_2 = 1 \\ w_1 > 0, w_2 > 0 \end{cases} \qquad (4)$$

4. REALIZING SIMULATION STRATEGIES

According to computing method of the linear programming, when w_1, w_2 fetching value is close and equal, c_{n+1} value almost solves optimally. Adopt number value to do a large number of simulating experiments, which find $w_{2/1}$ best value is [1.0-1.2] from the experiment; the value can get higher systematic function in the area. As to systematic resource that use often, *i* number value resource of system frequency utilization form relatively close to the constant after the numerous operations. It's according to algorithms of prevention deadlock resource sharing. In realization algorithm, which have adopted an approximate optimizes process method, in order to reduce the application of resources of the process to wait for limitlessly, wait for preferential progression trends process of formation calculate the formula will be defined as follows:

$$S = r_1 \frac{t_w}{t_a} + r_2 \frac{\sum_{i=1}^{p} t_s}{t_q} \qquad (5)$$

$$\sum_{i=1}^{p} t_s = t_1 + t_2 + t_3 + \ldots + t_s \qquad (6)$$

$$\overline{t_\alpha} = \frac{t_1 + t_2 + t_3 + \mathrm{L}\ t_n}{p} \qquad (7)$$

$$\{ r_1 > 0, r_2 < 0$$

$$\left|\frac{r_2}{r_1}\right| \in [0.6-0.8]$$

Among them, t_w express task waiting time that p finish already in waiting for formation, $\overline{t_a}$ express wait for formation all average of waiting time that process finish, $\sum_{i=1}^{p} t_s$ express process p estimate take resource total of time, t_q express process p estimate execution time, r_1, r_2 are the weight factor, which under load terms in the system, it can receive the fetching value of better performance . Provide the detailed description of algorithm below: When arriving in new process, insert the process to wait for and wait in the formation at first, if it appears resource to be idle to begin new process deployment systematically [3,4].

Can find out OEPR algorithm that put foreword have obvious advantages under the same load terms by Fig. 2., average each process has smaller response time, can improve systematic efficiency in OEPR algorithm.

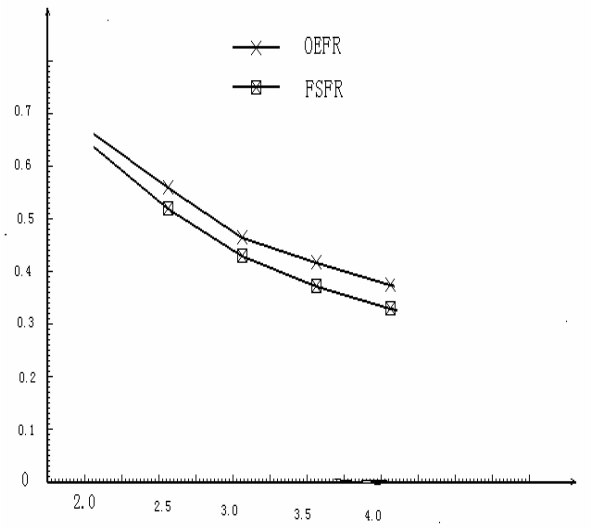


Fig. 2. Contrastive between OEPR optimize and FSPR

Under the background of practical research and simulate_ ion technology of food safety, the simulation model of info_ rmation share of supply chain management will reach an pr_ opose of enhancing system capability, The simulation outp_ uts show improving the efficiency of transportation, guaran_ teeing the standard or requirement of modern supply chain management towards food safely supply [5,6].

5. CONLUSION

In short, it is important to guarantee real time running for food safety. The model of have been put forward in this paper is significant that the scheme dealing with deadlock and realization of errorless multicast especially in respect of guaranteeing real time. The reliability of system improves system efficiency. As optimally designing management information sharing of the supply chain， an emulate model can be abstractly built aiming at the studied system. The information management system is simplified into a discrete and parallel system simulation model, optimization strategies are put forward and efficiency is improved.

It is significance that guaranteeing the standard or requirement of modern supply chain management towards food safely supply. Application of information technology plays an important role in perfecting the supply chain management system for modern logistics businesses. Realizing information resource sharing is an effective way to improve the efficiency of supply chain management. In this paper, the information management system in the supply chain is simplified into a discrete parallel system simulation model and a new effective strategy to prevent system deadlock is examined in depth, having a certain theoretical and practical value.

6. REFERENCES

[1] BanaszakZA, KroghBH, “Deadlock avoidance in flexible manufacturing systems with concurrently competing process flows”, IEEE Trans Robotics Automat, 1990, 6: 724-734.

[2] H. Moharam, “An efficient deadlock-free multicast wormhole algorithm in 2-D mesh multi-computers”. Journal of Systems Architecture, 2000, 46.1073-1091.

[3] R.C. Craig, M.J. Marianne, “Social Responsibility and Supply Chain relationships”. Transportation Research, 2002, (Part E 38). 37-52.

[4] T. Eiichi, N. Michihiko, “Optimal Size and Location Planning of Public Logistics Terminals”. Transportation Research, Part E 5, 1999, 207-222.

[5] S. leahy, R. Murphy, “Determinations of Successful Logistics Relationships: A Third -party Logistics Provider Perspective”. Transportation Journal 5, 1995, 35-39.

[6] B. David, “Douglas Wacbeth. Implementing Collaborations between Organizations an Empirical Study of Supply Chain Patterning”. Journal of Management Study, November 1999, 1003-1017.

Comparative and Optimization of Supply Chain Procurement Strategy in E-Commerce

Guangshu Chang
Research Division for Quality Engineering, Zhengzhou Institute of Aeronautics
Zhengzhou 450015, P.R. China
Management School, Northwestern Polytechnical University
Xi'an 710072, P.R. China
Email: cgs30@163.com

ABSTRACT

In e-commerce era, the complexity and levity of supply chain procurement management become greatly outstanding because the customer demands are changing faster and faster. Advanced information technology, mainly Internet, plays an increasingly important role in procurement management, and brings a global e-procurement (e-purchasing) via Internet. In this paper, a two-stage replenishment model of (*s*, *S*) strategy was established to analyze the operation cost efficiency of the traditional procurement in strategic partnership, e-procurement and the mixed procurement strategy respectively. This model quantified the influence of the Internet on procurement management. Then, the validity of this model was verified through a case study of a state-owned extra-large steel enterprise of China. The study concluded that the implementation of e-procurement technology can effectively reduce supply base and increase the profit margin of both the buyer and the supplier.

Keywords: Procurement, E-Procurement, E-Purchasing, Supply Chain, (*s*, *S*) strategy.

1. INTRODUCTION

For most of the industrial enterprises, the procurement of materials, components and parts is the first step of manufacturing. Therefore, procurement management is very important for reducing the total cost, improving the final product quality, and shortening the operation cycle of supply chain system. Procurement management has been one of the most valuable parts in supply chain management (SCM). It is forecasted that the typical decision-making issues of supply chain in future will focus on the field of procurement management [1]. In e-commerce era, the complexity and levity of procurement management increase greatly because customer demands are changing faster and faster. As a major part of contemporary information technology (IT), the Internet is more and more important in procurement management for improving operation efficiency and competitive advantages of the total supply chain, such as sharing information, reducing cost, shortening cycle, enhancing cooperation, improving service, etc. [2].

Through investigation, Lancioni et al. found that procurement function of the American enterprises is optimized widely because of the utilization of the Internet, for example, GE reduced 40 percent of procurement cost and more than 50 percent procurement personnel [3]. Hence, procurement management using the Internet has become the most important part of e-commerce strategy of the large enterprises [4] and this trend has been developing rapidly in recent years. At the same time, e-procurement (e-purchasing) becomes one of the first-line research fields sponsored by the Institute for Supply Management of USA. Up to now, researchers has studied some related problems of e-procurement. For example, Davila et al. studied the benefits, risks, application fields and implement strategies of e-procurement technologies [5]. Reference [6] discussed some business and technical issues of e-procurement. Reference [7] proposed four kinds of web-based procurement models, including buy-side procurement system, industry B2B exchange, private B2B e-market and third party B2B e-market, and analyzed their value and application respectively. Reference [8] studied the direct and indirect effect of different e-procurement modes on the procurement costs. Reference [9] studied the strategic and operational benefits of electronic integration in B2B procurement process. Reference [10] analyzed the roles of knowledge management for improving the efficiency and benefit of e-procurement system. Reference [11] discussed the strategic, operational and opportunity benefits and seven caveats of effective web-based procurement. Reference [12] explored the impact of adopting e-procurement system on the companies in Singapore. Reference [13] analyzed how and why SMEs in Hong Kong employed e-procurement through four case studies.

However, the research on e-procurement is still in its infancy. The ubiquitous impact of the Internet on supply chains has raised many questions as to the range of its application and the degree of Internet usage in procurement management. For example, how has the Internet been applied to purchasing and what has been its effect on procurement strategies [3]? Furthermore, most literatures are qualitative, explorative or case-based study. Quantitative study is few. In the current quantitative studies on procurement management, most researchers studied it from the view of cost minimum or profit maximum no matter what kind of method they used. Therefore, cost efficiency become the main principle for evaluating traditional procurement and e-procurement strategy. However, they usually focus on the advantages of e-procurement strategy but consider rarely the ones of traditional procurement contract. Although the implementation of e-procurement strategy can effectively reduce procurement cost, optimize procurement process and improve operation efficiency of the overall supply chain system, the long-term strategic partnetship formed with a few key suppliers has its important and unreplacebale advantages, such as improving quality, reducing supply uncertainty, and promoting supplier invovlement and development. Moreover, enterprises usually do not perform large-scale or strategic purchasing through e-procurement because e-procurement technologies are still in their early stages of the traditional technology S-curve and there are risks linked to the adoption of e-procurement. Consequently, most enterprises perform the mixed procurement strategy, which combine both traditional procurement and e-procurement strategy.

The objective of this paper is to analyze and compare the cost efficiency of traditional procurement in strategic partnetship, e-procurement and the mixed procurement strategy through establishing a two-period inventory replenishment model of (*s*, *S*) strategy to quantify the effect of the Internet on procurement management. This paper is organized as follow: first, in section 1, we outline the status quo and documents about e-procurement research. Then in section 2, the two-period inventory replenishment model is established and the optimal solutions to order quantity are provided. After that, in section3, a case study is presented. Finally, in section 4, concluding remarks are offered.

2. MODEL ANALYSIS

In the previous procurement literatures, the order quantity is usually assumed as an exogenous variable. Thus, the main objective of establishing model is to seek suppliers who most effectively satisfy the specific demand [14]. However, in e-commerce environment, the determination of order quantity actually depends on the procurement strategy chose by the buyer.

In a simple supply chain system formed by one buyer and one supplier, a two-period inventory replenishment model of (*s*, *S*) stategy is established based on the work of Peleg et al. [14]. The buyer faces a stochastic demand D_i in period *i*, where *i*=1,2, and $D_i \sim N(\mu_i, \sigma_i^2)$. We use the following notations in our model:

$F_i(X)$: Cumulative density function of demand in period *i*, where *i*=1,2;
$f_i(X)$: Probability density function of demand in period *i*, where *i*=1,2;
X_i: Order quantity in period *i*, where *i*=1,2;
$\underline{X}$: Minimum order quantity;
h_i: Unit inventory holding cost in period *i*, where *i*=1,2;
A_{ij}: Unit Setup cost of the *i*th procurement strategy in period *j*, where *i*=1,2,3, *j*=1,2;
b: Unit stock-outs cost;
p_i: Unit procurement price of the *i*th procurement strategy, where *i*=1,2;
$C_i(X)$: Operation cost of the *i*th procurement strategy, where *i*=1,2,3.

2.1 Traditional Procurement Strategy

In traditional procurement strategy, a long-term supply contract is signed between the buyer and the supplier under the background of strategic partnership. Thus, all procurement in the two periods is performed based on the contract at a pre-negotiated unit price p_1. For simplicity, minimum order quantity will be neglected. Operation cost of the buyer for the two periods is then equal to:

$$C_1(X) = A_{11} + p_1 X_1 + \left[h_1 E_{D_1}(X_1 - D_1) + bE_{D_1}(D_1 - X_1)\right] + E_{D_1}\left[A_{12} + p_1 X_2 + h_2 E_{D_2}(X_1 + X_2 - D_1 - D_2) + bE_{D_2}(D_2 - X_1 - X_2 + D_1)\right] \quad (1)$$

Setting the first derivative of equation (1) with respect to X_2 equal to zero we can get:

$$\frac{\partial C_1(X)}{\partial X_2} = p_1 + (b + h_2)F_2(X_1 + X_2 - D_1) - b = 0 \quad (2)$$

Then:

$$F_2(X_1 + X_2 - D_1) = \frac{b - p_1}{b + h_2} \quad (3)$$

Let $K_1 = X_1 + X_2 - D_1$ and substituting it into equation (3), we obtain:

$$F_2(K_1) = \frac{b - p_1}{b + h_2} \quad (4)$$

Substituting $X_2^* = K_1 + D_1 - X_1$ into equation (1), and setting the first derivative of equation (1) with respect to X_1 equal to zero we can get:

$$\frac{\partial C_1(X)}{\partial X_1} = (p_1 - b) + (b + h_1)F_1(X_1) + F_1(D_1 - X_2)\left[(b + h_2)F_2(K_1) - b\right] = 0 \quad (5)$$

Then:

$$F_1(X_1) = \frac{b - p_1}{b + h_1} + F_1(D_1 - X_2)\left[\frac{b}{b + h_1} - \frac{b + h_2}{b + h_1}F_2(K_1)\right] \quad (6)$$

The expressions for X_1^* and X_2^* can be summarized in Lemma 1:

Lemma 1. Under the traditional procurement strategy in strategic partnership, the optimal order quantities of the supply chain system X_1^* and X_2^* satisfy:

$$X_2^* = K_1 - (X_1^* - D_1) \quad (7)$$

$$F_1(X_1^*) = \frac{b - p_1}{b + h_1} + F_1(D_1 - X_2^*)\left[\frac{b}{b + h_1} - \frac{b + h_2}{b + h_1}F_2(K_1)\right] \quad (8)$$

Where $K_1 = F_2^{-1}\left(\frac{b - p_1}{b + h_2}\right)$.

2.2 E-Procurement Strategy

In e-procurement strategy, all procurement in the two periods is performed through electronic methods, such as e-marketplace, e-procurement system based on EDI, Internet, etc. Without loss of generality, the setup cost of the supplier and buyer in e-procurement strategy are usually neglected because they are very small compared to those in traditional procurement strategy [15], namely A_{21}=A_{22}=0. At the same time, there is usually minimum order quantity in e-procurement strategy [16]. Operation cost of the buyer for the two periods is then equal to:

$$C_2(X) = p_2(X_1 + \underline{X}) + \left[h_1 E_{D_1}(X_1 + \underline{X} - D_1) + bE_{D_1}(D_1 - X_1 - \underline{X})\right] + E_{D_1}\left[p_2(X_2 + \underline{X}) + h_2 E_{D_2}(X_1 + X_2 + \underline{X} - D_1 - D_2) + bE_{D_2}(D_2 - X_1 - X_2 - \underline{X} + D_1)\right] \quad (9)$$

Setting the first derivative of equation (1) with respect to X_2 equal to zero we can get:

$$\frac{\partial C_2(X)}{\partial X_2} = p_2 + (b + h_2)F_2(X_1 + X_2 + \underline{X} - D_1) - b = 0 \quad (10)$$

Then:

$$F_2(X_1 + X_2 + \underline{X} - D_1) = \frac{b - p_2}{b + h_2} \quad (11)$$

Let $K_2 = X_1 + X_2 + \underline{X} - D_1$ and substituting it into equation (11), we obtain:

$$F_2(K_2) = \frac{b - p_2}{b + h_2} \quad (12)$$

Substituting $X_2^* = K_2 + D_1 - X_1 - \underline{X}$ into equation (9), and setting the first derivative of equation (9) with respect to X_1 equal to zero we can get:

$$\frac{\partial C_2(X)}{\partial X_1} = (p_2 - b) + (b + h_1)F_1(X_1 + \underline{X}) + F_1(D_1 + \underline{X} - X_2)[(b + h_2)F_2(K_2) - b] = 0 \quad (13)$$

Then:

$$F_1(X_1 + \underline{X}) = \frac{b - p_2}{b + h_1} + F_1(D_1 + \underline{X} - X_2)\left[\frac{b + h_2}{b + h_1}F_2(K_2) - \frac{b}{b + h_1}\right] \quad (14)$$

The expressions for X_1^* and X_2^* can be summarized in Lemma 2:

Lemma 2. Under the e-procurement strategy, the optimal order quantities of the supply chain system X_1^* and X_2^* satisfy:

$$X_2^* = K_2 - (X_1^* + \underline{X} - D_1) \quad (15)$$

$$F_1(X_1^* + \underline{X}) = \frac{b - p_2}{b + h_1} + F_1(D_1 - X_2^*)\left[\frac{b + h_2}{b + h_1}F_2(K_2) - \frac{b}{b + h_1}\right] \quad (16)$$

Where $K_2 = F_2^{-1}\left(\frac{b - p_2}{b + h_2}\right)$.

2.3 Mixed Procurement Strategy

In the mixed procurement strategy, the buyer purchase material, components and parts from the supplier through different ways in the two periods. In the first period, all procurement is performed based on the long-term contract. In the second period, the buyer will purchase the additional units from the supplier through e-procurement methods. Operation cost of the buyer for the two periods is then equal to:

$$C_3(X) = A_{31} + p_1X_1 + \left[h_1E_{D_1}(X_1 - D_1) + bE_{D_1}(D_1 - X_1)\right] + E_{D_1}\left[p_2(X_2 + \underline{X}) + h_2E_{D_2}(X_1 + X_2 + \underline{X} - D_1 - D_2) + bE_{D_2}(D_2 - X_1 - X_2 - \underline{X} + D_1)\right] \quad (17)$$

Setting the first derivative of equation (1) with respect to X_2 equal to zero we can get:

$$\frac{\partial C_2(X)}{\partial X_2} = p_2 + (b + h_2)F_2(X_1 + X_2 + \underline{X} - D_1) - b = 0 \quad (18)$$

Then:

$$F_2(X_1 + X_2 + \underline{X} - D_1) = \frac{b - p_2}{b + h_2} \quad (19)$$

Let $K_3 = X_1 + X_2 + \underline{X} - D_1$ and substituting it into equation (19), we obtain:

$$F_2(K_3) = \frac{b - p_2}{b + h_2} \quad (20)$$

Substituting $X_2^* = K_2 + D_1 - X_1 - \underline{X}$ into equation (17), and setting the first derivative of equation (17) with respect to X_1 equal to zero we can get:

$$\frac{\partial C_3(X)}{\partial X_1} = (p_1 - b) + (b + h_1)F_1(X_1) + F_1(D_1 + \underline{X} - X_2)[(b + h_2)F_2(K_2) - b] = 0 \quad (21)$$

Then:

$$F_1(X_1 + \underline{X}) = \frac{b - p_2}{b + h_1} + F_1(D_1 + \underline{X} - X_2)\left[\frac{b}{b + h_1} - \frac{b + h_2}{b + h_1}F_2(K_3)\right] \quad (22)$$

The expressions for X_1^* and X_2^* can be summarized in Lemma 3:

Lemma 3. Under the mixed procurement strategy, the optimal order quantities of the supply chain system X_1^* and X_2^* satisfy:

$$X_2^* = K_3 - (X_1^* + \underline{X} - D_1) \quad (23)$$

$$F_1(X_1^* + \underline{X}) = \frac{b - p_1}{b + h_1} + F_1(D_1 - X_2^*)\left[\frac{b + h_2}{b + h_1}F_2(K_3) - \frac{b}{b + h_1}\right] \quad (24)$$

where $K_3 = F_2^{-1}\left(\frac{b - p_2}{b + h_2}\right)$.

3. A CASE STUDY

Here, a state-owned extra-large steel enterprise of China is considered. At present, the company has more than 20 sub-companies including mine engineering, coking, steelmaking, metallurgical construction, machining, science research and development, etc. It can provide various products of more than 30 grades and 2000 specifications. It produces more than 6 million ton iron, 5 million ton steel and 4 million ton steel products per year. Its production needs diversiform input, including 8 grades and 560 specifications materials, e.g. high quality ironstone and coke; 79 grades and 869 specifications components and parts; 7 grades and 46 specifications MRO, e.g. computer and service.

In the past, the company's production management followed this way: laying production plan by production management department (PMD) → making lunar production plan known to all plants → each plant making its schedule and material requirements plan (MPS) and reporting to PMD → examined and approved by PMD → implementing the plans by the plants. We can find that the inventory level of material, components and parts is determined by the plants themselves but not PMD. Since departmental selfishness and information insulation among the plants, the MPSs are usually inaccurate. Moreover, inventory data is static and neglect the material quantity at hand. Therefore, procuerment process is inefficient, such as waste, overstock, and the puchased units' unconformity with production needs, etc.

Since 2001, the company has applied IT through an integrated approach, including E-mail, EDI, ERP, e-tendering, e-informing, e-marketplace, etc., to improve procurement process, help forge close relationships with key suppliers and improve service to users across the business. An overview of e-commerce pilots and future developments under consideration in the company are summarized in Table 1 in Appendix.

According to our model, the calculating results are listed in Table 2. We can find that the optimal order quantities and inventory level in the two periods of e-procurement strategy are all minimal in comparison with the other two strategies. Therefore, the e-procurement strategy should be optimal. However, the company implements the mixed procurement strategy because the current application of IT is incomplete and traditional methods are still necessary in some business and procurement processes. In spite of that, the application of IT and implementation of e-procurement technology has brought evident benefits for the company as shown in Table 3. We can find that there is evident reduction in the level of procurement price, order quantity, holding cost, lead-time, stock-outs cost, etc.

4. CONCLUSION

In e-commerce environment, the determination of order quantity actually depends on the procurement strategy chose by the buyer. Therefore, we establish a two-period replenishment model of (*s*, *S*) strategy to analyze the operation cost of the traditional procurement in strategic partnership, e-procurement and the mixed procurement strategy respectively. Finaly, the validity of our model is verified through a case study of a state-owned extra-large steel enterprise of China. It is found that the implementation of e-procurement technology can effectively reduce supply base and increase the profit margin of both the buyer and the supplier. Thus, it helps improve the relationship between the buyer and the supplier and realize win-win situation.

ACKNOWLEDGEMENT

This research was partially supported by the National Natural Science Foundation of China under Grant 70125004, 70572050 and 70401011.

REFERENCES

[1] Q. Dai, R.J. Kauffman, “To be or not to B2B? An evaluation model for e-procurement channel adoption”, Working Paper, Carlson School of Management, University of Minnesota, USA, 2004.

[2] P. Keskinocak, S. Tayur, “Quantitative analysis for internet-enabled supply chains”, Interfaces, 2001, Vol. 31(2): 70-89.

[3] R.A. Lancioni, M.F. Smith, et al., “Strategic Internet application trends in supply chain management”, Industrial Marketing Management, 2003, Vol. 32(2): 211-217.

[4] A. Barua, P. Konana, et al., “Driving e-business excellence”, Sloan Management Review, 2001, Vol. 43(1): 36-44.

[5] A. Davila, M. Gupta, et al., “Moving procurement systems to the Internet: The adoption and use of e-procurement technology models”, European Management Journal, 2003, Vol. 21(1): 11-23.

[6] T.M. Rajkumer, “E-procurement: Business and technical issues”, Information Systems Management, 2001, Vol. 18(4): 22-36.

[7] C. Subramaniam, M.J. Shaw, “A study of the value and impact of B2B e-commerce: the case of web-based procurement”, International Journal of Electronic Commerce, 2002, Vol. 6(4): 19-40.

[8] L. Boer, J. Harink, et al., “A conceptual model for assessing the impact of electronic procurement”, European Journal of Purchasing & Supply Management, 2002, Vol. 8(1): 25-33.

[9] T. Mukhopadhyay, “Strategic and operational benefits of electronic integration in B2B procurement processes”, Management Science, 2002, Vol. 48(10): 1301-1313.

[10] C. Hsieh, H. Yang, et al., “Roles of knowledge management in online procurement systems”, Industrial Management & Data Systems, 2002, Vol. 102(7): 365-370.

[11] M. Attaran, “The coming age of online procurement”, Industrial Management & Data Systems, 2001, Vol. 101(4): 177-180.

[12] C.B. Kheng, S. Al-Hawamdeh, “The adoption of electronic procurement in Singapore”, Electronic Commerce Research, 2002, Vol. 2(2): 61-73.

[13] J.K.Y. Chan, M.K.O. Lee, “SME e-procurement adoption in Hong Kong: The roles of power, trust and value”, Proceedings of the 36th Annual Hawaii International Conference on System Sciences, Hawaii, USA, 2003, pp.179-188.

[14] B. Peleg, H.L. Lee, et al., “Short-term e-procurement strategies versus long-term contracts”, Production Operation Management, 2002, Vol. 11(4): 458-479.

[15] C.X. Wang, M. Benaroch, “Supply chain coordination in buyer centric B2B electronic markets”, International Journal Production Economics, 2004, Vol. 92(2): 113-124.

[16] M. Popović, “B2B e-Marketplaces”, Available in: http://www.teleactivities.net, 2002.

APPENDIX

Table 1. E-Procurement solution framework

	Current solution	Internet solution (current)	Prediction for future
Production manag.	ERP	SMTP	E-Commerce pilots
Inventory manag.	ERP	E-Procurement	CPFR
Cost reduction	Links with suppliers	E-Procurement	Internet SMTP E-mail links with suppliers
Customer manag.	Link Intranet with Internet	E-Commerce pilots	Online ordering system linked to supplier base
Collaboration with suppliers	Web-based E-mail	SMTP E-mail, online procedure	SMTP E-mail, WWW, FTP
Logistics manag.	ERP	SMTP	SMTP E-mail, WWW, FTP

Table 2. The main indexes of the model

	Traditional procurement	E-Procurement	The mixed procurement
P	1560	1320	-
h_1	132	124	-
h_2	112	103	-
$\underline{X}$	-	2	2
L	-7.5	2.3	-1.1
D_1	1.8	2.2	2.0
D_2	1.6	2.0	1.8
K	9.9	6.7	9.4
X_1^*	6.2	3.7	5.1
X_2^*	5.5	3.2	4.3

Table 3. Comparison before and after the implementation of e-procurement technology

	Before	After
1. The main material (High quality ironstone)		
Unit procurement price (Yuan per ton)	1560	1320
Order quantity per year (Ten thousands ton)	320	260
Unit holding cost (Yuan per ton)	244	227
Unit stock-outs cost (Yuan per ton)	72	65
Unit setup cost (Yuan per ton)	36	12
Lead-time (Days)	-7.5	-1.1
2. The main product (Cold rolling steel sheet)		
Factory price (Yuan)	828	828
Annual output (Ten thousands ton)	96	112
3. Supplier		
Number of suppliers	296	232
Percentage of links with suppliers through the Internet in procurement process (%)	5	39
Percentage of tendering through the Internet (%)	2	97
Profit margin (%)	19	23

Coordination Mechanism within Supply Chain Based on Two-hierarchical Programming Theory*

Cui Hua Zhang
School of Business Administration, Northeastern University
Shenyang, (LiaoNing Province) 110004, P.R.China
Email: chzhang@mail.neu.edu.cn
And
Guang Shu Chang
Zhengzhou Institute of Aeronautics
Zhengzhou (HeNan Province) 450015, China
Email: cgs30@163.com
And
Liang Ge
School of Business Administration , Northeastern University
Shenyang, (LiaoNing Province) 110004, P.R.China
Email: geliang0429@eyou.com

ABSTRACT

This article studies the coordination problem in supply chain. Basic production models for a supplier and a manufacturer are developed. And then a proper mechanism for supply chain coordination is put forward based on two-hierarchical programming theory. In the models, the manufacturer is regarded at the upper layer, while the supplier is at the lower one. The manufacturer can influence the supplier decision through predicting the supplier's supply quantity and punishing the supplier for supply delay. At last, simulation is done. Results indicate that the solution of supply chain optimization based on two-hierarchical programming theory leads to more profit and less cost than that in simple way. Furthermore, the best effect of coordination appears in optimizing the penalty coefficient value.

Keywords: Coordination, Two-hierarchical programming, Penalty coefficient, Cost.

1. INTRODUCTION

With the increased competition between enterprises, demand for supply chain coordination is becoming more and more urgent [1~4]. Two or more enterprises form network union through firm agreement or combined organization for some strategic objectives, and this network union is named as supply chain coordination [5]. The operation cost of supply chain coordination is lower than that of other forms of operation [6, 7]. Monahan studied a two-hierarchical systemic quantity discount problem consisting of a buyer and a seller [8]. He pointed out that the bargainer can increase profits through supplying quantity discount. In some supply chain coordination articles based on inventory control, Porteus and Whang studied the coordination problems between many agents and one manufacturing company under asymmetry information [9]. They educed the best incentive plan to motivate the manager's action, thereby maximizing the corporation's expected profit. Here, the decision makers usually belong to one corporation, which is representative of concentrative supply chain operation. Arntzen used one multistage mixture-integar model for multi-products to optimize a numeral equipment corporation's global supply chain [10]. These approaches treat the supply chain as a vertical integrated corporation. All information is shared inside one company and one single decision maker controls physical distribution. Kolisch studied the integration problem of all kinds of manufacturing systems. He developed one formulation for this problem and reached one heuristic solution [11]. Christoph and Schneeweiss applied hierarchical programming theory to study the coordinative operational mechanism between two neighboring layers in supply chain [12]. They also discussed the meaning of correlated information prediction and the function of controlled variables.

Based on Christoph and Schneeweiss's study, this article is to study the coordination mechanism within two neighboring layers in supply chain and emphasize the application of two-hierarchical programming theory. Then we discuss the penalty coefficient's function. At last, simulation is done to evaluate the effect of a decentralized supply chain's coordination operation.

2. BASIC MODEL FOR SUPPLY CHAIN

In supply chain, two neighboring members constitute a two-hierarchical architecture. And the information between them is not absolutely symmetric. Coordination between the manufacturers's internal assembles and the supplier's exterior supply becomes the primary research problem. It can be regarded as a two-hierarchical programming problem.

The manufacturer orders q_t from the supplier to meet exterior demand V_t. The supplier produces commodities according to the order q_t at their own capacity.

2.1 Manufacturer Model

The manufacturer model is a capacity-planning model. The manufacturer assigns production resources before supply quantity d_t arrives, however, after order quantity q_t is

* This paper is funded by National Natural Science Fund of China under Grant No.70401011.

ascertained. The model is as follows:

$$C^T = \sum_{t=1}^{T}(cY_t + c^+Y_t^+ + c^-Y_t^-) + \sum_{t=1}^{T}(p\hat{d}_t - k\hat{\eta}_t^+ - \lambda\hat{\eta}_t^-) + \sum_{t=1}^{T}h^pI_t + \sum_{t=1}^{T}F\Delta_t \quad (1)$$

s.t

$$Y_{t+1} = Y_t + Y_t^+ - Y_t^- \quad (2)$$

$$Y_t^+ \le Y_t^{\max} \quad (3)$$

$$Y_t^- \le Y_t^{\min} \quad (4)$$

$$Y_1 = Y^{'} \quad (5)$$

$$\alpha \cdot \hat{d}_t \le Y \quad (6)$$

$$I_{t+1} = I_t + \hat{d}_t - Vt \quad , \quad I_1 = 0 \quad (7)$$

$$\Delta_t = V_t - I_t - \hat{d}_t \quad (8)$$

$$\sum_{t=1}^{T}V_t = \sum_{t=1}^{T}q_t \quad (9)$$

$$\hat{d}_{t+1} + \hat{\eta}_{t+1}^+ - \hat{\eta}_{t+1}^- = q_{t+1} + \hat{\eta}_t^+ - \hat{\eta}_t^- \quad (10)$$

$$\hat{d}_1 + \hat{\eta}_1^+ - \hat{\eta}_1^- = q_1 \quad (11)$$

$$\hat{d}_t = AF(q_t) \quad (12)$$

$$Y_t \ge 0, q_t, \hat{d}_t, \hat{\eta}_t^+, \hat{\eta}_t^-, I_t, \Delta_t \ge 0, t \in 1,2,3\cdots T \quad (13)$$

Here, the objective function is concerned with the manufacturers' cost. The manufacturer forecasts the actual supply quantity d_t after the order q_t is ascertained and thereby makes production decision according to the predicted value $\hat{d}_t$. $\sum_{t=1}^{T}(cY_t + c^+Y_t^+ + c^-Y_t^-)$ Denotes the manufacturer's total producing cost. $\sum_{t=1}^{T}p\hat{d}_t$ Represents the total order price. $\sum_{t=1}^{T}(k\hat{\eta}_t^+ + \lambda\hat{\eta}_t^-)$ is a dominant function which is imposed on the supplier by the manufacturer. It can make the supply chain operation more collaborative and the overall cost much lower. $\sum_{t=1}^{T}h^pI_t$ represents inventory cost. $\sum_{t=1}^{T}F\Delta_t$ represents the cost when manufacturers can not satisfy exterior demand. Constraint (2) denotes the relation between two-stage production resource; Constraints (2) and (4) represent the restriction to production resource change. Constraint (6) represents the constraint on production resource; Constraints (7) and (8) represent the balance between the supply and the order; $AF(q_t)$ is an anticipation function in constraint (9), which is to translate the supplier 's capacity modular 's forecast value $\hat{\beta}_t$ into the supplier 's profit model and optimize it to gain $\hat{d}_t$ accordingly. Y_t is production resource in t stage; Y_t^+ is incremental resource at t stage; Y_t^- is reductive resource at t stage, q_t is order quantity in t stage; $\hat{d}_t$ is supply quantity forecasted at t stage; $\hat{\eta}_t^+$ is plus deviation between purchasing quantity and supply quantity; $\hat{\eta}_t^-$ is minus deviation between purchasing quantity and supply quantity; I_t is inventory quantity at t stage. Δ_t is the exterior demand quantity which is not satisfied. The variables involved above are all decision variables. c is unit resource cost; c^+ is unit cost when resource increases; c^- is unit cost when resource reduces; p is unit order price; h^p is unit inventory cost; F is per-unit delay cost; K is per-unit penalty if the supplier delays supply. Because the out-of-stock cost is very high, it is a very important controlled variable. λ is unit penalty if supply is in advance, because it is likely to increase inventory cost. V_t is the number of parts which is required by productive task in t stage, which is derived from exterior demand change; α is production resource which is required by production of per-unit part; $Y_t^{\max}$ and $Y_t^{\min}$ represent the maxima of resource increase and reduction respectively. The variables involved above are all constants.

The manufacturer can organize production schedule based on the forecast value $\hat{d}_t$ after the order is ascertained. The manufacturer's actual production cost equals the cost when the actual order quantity d_t arrives, however, after production schedule is ascertained.

2.2 Supplier Model

The supplier supplies commodities for several manufacturers. Its capacity $\beta_t C$, which is assigned, to every manufacturer is a stochastic variable. The supplier optimizes its own production according to the manufacturer order q_t and the controlled parameter K, λ. Besides, it supplies the quantity d_t.

The supplier's production model is a profit model:

$$C^B = \sum_{t=1}^{T}p\,d_t - \sum_{t=1}^{T}h^sI_t^s - \sum_{t=1}^{T}k\,\eta_t^+ - \sum_{t=1}^{T}\lambda\eta_t^- - \sum_{t=1}^{T}\delta\Delta C_t \quad (14)$$

s.t

$$I_{t+1}^s = I_t^s + Q_t - d_t \quad (15)$$

$$I_1^S = I \tag{16}$$

$$\varepsilon Q_t \le \beta_t C + \Delta C_t \tag{17}$$

$$\Delta C_t \le \Delta C_t^{\max} \tag{18}$$

$$d_{t+1} + \eta_{t+1}^{+} - \eta_{t+1}^{-} = q_{t+1} + \eta_t^{+} - \eta_t^{-} \tag{19}$$

$$d_1 + \eta_1^{+} - \eta_1^{-} = q_1 \tag{20}$$

$$\sum_{t=1}^{T} d_t = \sum_{t=1}^{T} q_t \tag{21}$$

Here, the supplier's objective function is concerning its profit. $\sum_{t=1}^{T} p\, d_t$ is the total order price; $\sum_{t=1}^{T} h^S I_t^S$ stands for the stock; $\sum_{t=1}^{T} k\,\eta_t^{+}$ *and* $\sum_{t=1}^{T} \lambda \eta_t^{-}$ stand for the penalty cost; $\sum_{t=1}^{T} \delta \Delta C_t$ stands for the cost caused by supplier's increaser production ability. d_t is the supply quantity at t stage; I_t^S is the inventory quantity at t stage; η_t^{+} is the delay supply at t stage; η_t^{-} is the supply in advance at t stage; Q_t is the production quantity at t stage; ΔC_t is the supplier's increaser production resource at t stage; The above variables are decision variables; h^s is the unit inventory cost; $I_1^{'}$ is the initial inventory quantity; q_t is the order quantity at t stage which is a variable to the manufacturer but a constant to the supplier; C is the supplier's total capacity; ε is the capacity required to produce per-unit product by the supplier; β_t is the capacity utility coefficient at t stage.

2.3 Supply Chain Model

The supply chain 's total cost function is as follows:

$$C^{TT} = \sum_{t=1}^{T} (cY_t + c^{+}Y_t^{+} + c^{-}Y_t^{-}) + \sum_{t=1}^{T} h^{\mathrm{p}} I_t + \sum_{t=1}^{T} F\Delta_t + \sum_{t=1}^{T} h^S I_t^S + \sum_{t=1}^{T} \delta \Delta C_t \tag{22}$$

Here, $\sum_{t=1}^{T} (p\,\hat{d}_t - k\,\eta_t^{+} - \lambda \eta_t^{-})$ and $\sum_{t=1}^{T} p\, d_t - \sum_{t=1}^{T} k\,\eta_t^{+} - \sum_{t=1}^{T} \lambda \eta_t^{-}$ aren't considered in the total cost because they are the internal costs in supply chain, only affecting the supply chain 's interior assignment of interest. The article attempts to find out the trend of supply chain's total cost.

In practice, supply chain operation is mostly pulled by orders. The manufacturer arranges production plan based on its experience ignoring the supplier the manufacturer forecasts the supply quantity $\hat{d}_t$, which often equals the order quantity q_t. Whereas the supply quantity d_t by the supplier doesn't equal q_t, the increase of deviation $|q_t - d_t|$ can cause more difficulty in decision making, more waste of production resources and more cost of total supply chain [13].

3. SUPPLY CHAIN COORDINATION MECHANISM BASED ON TWO-HIERARCHICAL PROGRAMMING

3.1 Theory of Two-hierarchical Programming

Two-hierarchical programming is a mathematical model for two-hierarchical decision problem, which is a system optimization problem with two-hierarchical recursive structure [14]. The upper layer's objective function and constraints are not only relative with its own decision variables, but also depend on the lower layer problem's optimal solution, while the layer's optimal solution is impacted by the upper layer's decision variables.

The basic mathematical model is as follows:

$$\max_x f_1(x,y) = \mathrm{a}^{\mathrm{T}} x + {}^{\mathrm{T}} y \tag{23}$$

$$\max_y f_2(x,y) = \mathrm{c}^{\mathrm{T}} x + \mathrm{d}^{\mathrm{T}} y \tag{24}$$

$$\text{s.t} \quad Ax + By \le \mathrm{r} \tag{25}$$

$$x \ge 0, y \ge 0 \tag{26}$$

Here, Eq. (23) is the upper layer function, x is the decision variable, y is constant. Eq.(24) is the lower layer function, y is the decision variable, x is constant. After optimizing Eq. (23), the optimum value of x_1 is transmitted to Eq. (24). Then Eq(24) proceeds to optimize its optimal value y_1, which is transmitted to Eq.(23). Thus the optimization cycled to reach the most optimal solution in the end.

3.2 Coordination Mechanism

With the help of two-hierarchical theory, the supplier and the manufacturer can make decisions interactively. The manufacturer forecasts the supply quantity according to its own information after ascertaining the order and arranges its production planning based on $\hat{d}_t$. $X = (Y_t, q_t, k, \lambda, \hat{d}_t)^T$ is the manufacturer decision vector, while $(q_t, k, \lambda)^T$ is transferred to the lower layer supplier. $\hat{d}_t$ is an implicit decision variable which can influence the supplier production directly. The more accurate the forecast, the smaller the value $\left|d_t - \hat{d}_t\right|$, the fewer the waste of the resources and the lower of the supply chain total cost. The manufacturer gets $\hat{d}_t$ through forecasting each stage's assigned capacity coefficient $\hat{\beta}_t$, then transferring q_t to the supplier 's production function to optimize. $Y = (Q_t, d_t, \Delta C_t)^T$ is the supplier's decision vector, while.

d_t is transferred to the manufacturer. The supply chain's neighboring layers continue the order revision under the theory of two-hierarchical programming. And this can lower the supply chain's total cost. And reach the optimal order quantity.

4. SIMULATIONS AND ANALYSIS

4.1 Algorithm Description

Step 1: Substituting the manufacturer's order quantity q_t into the supplier's production function and optimizing it, we can acquire an optimal supply quantity d_t .

Step 2: Substituting q_t and the supplier's capacity coefficient value $\hat{\beta}_t$ into the supplier's production function; we can obtain the supply quantity's forecast value $\hat{d}_t$.

Step 3: Substituting $\hat{d}_t$ into the manufacturer's production function, we can get the manufacturer's production resource arrangement Y_t .

Step 4: Substituting d_t into the planned production function when the actual d_t arrives, we can acquire the manufacturer's actual cost. Further more, we can get the supply chain's total cost.

4.2 Simulation Computation

An assembly enterprise order components from a supplier and proceeds to assemble. The whole production process is divided into five stages. The number of parts ch is required by each stage's productive task can be seen in Table 1.

Table 1. Number of parts required by each stage's production

t	1	2	3	4	5
V_t (Unit)	30	20	35	40	35

Table 2. The manufacturer's production data

c	$Y_t^{\min}$	$Y_t^{\max}$	F	h^p	P
0.5 (Y/h)	30 (h)	30 (h)	10 (Y/unit)	2 (Y/unit)	10 (Y/unit)

λ	α	Y'	c^+	c^-
1 (Y/unit)	2 (h/unit)	40 h	1 (Y/h)	0.2 (Y/h)

Table 3. The supplier's production data

p	h^s	I_1'	C	ε	δ	$\Delta C_t^{\max}$
10 (Y/unit)	1 (Y/unit)	0 (Y/unit)	100 (h)	2 (h/unit)	1.5 (Y/unit)	80 (h)

Table 4. The Supplier's capacity coefficient β_t

t	1	2	3	4	5
β_t	0.65	0.18	0.40	0.50	0.25

Table 5. The value $\hat{\beta}_t$ forecasted by the manufacturer

t	1	2	3	4	5
mean value	0.66	0.21	0.35	0.52	0.27

Table 6. The total cost of the supply chain

	Order combinations					Cost of simple Supply chain Operation	Cost based on Hierarchical theory
	q_1	q_2	q_3	q_4	q_5		
1	30	20	35	40	35	434	392.55
2	25	25	33	42	35	463.12	402.22
3	33.5	26.5	31	43	36	449.32	398.5
4	32	25	30	41	32	502.16	458.35
5	29	21.5	34.5	42	33	416.5	367.56
6	26.5	22.5	36	40.5	34.5	449.65	399.86
7	20.5	23.5	30	43	32	426.84	389.65
8	34	21.5	29.5	44	31	478.36	412.52

Results indicate that the total cost of the supply chain, which is based on the theory of two-hierarchical programming, is always lower than that in simple way. This is because the manufacturer considers the lower layer when it proceeds to optimize is own decisions. Thus the manufacturer's decision becomes more accurate.

Next we consider the controlled variable's effect, regarding one order combination =(30，20，30，35，40，35) and optimizing the supply chain's total cost.

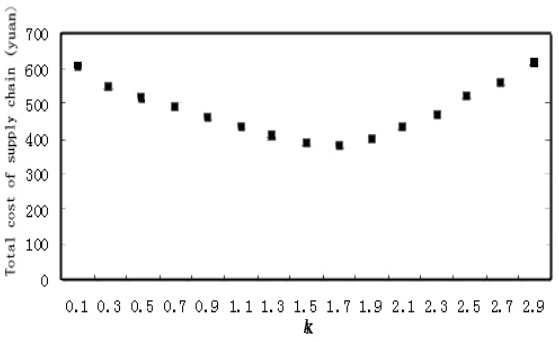


Fig.1 The diversification of supply chain's cost

From Fig. 1. We can find that the supply chain's total costs are different when penalty coefficient adopts different values. While K = 1.7, the supply chain's total cost is the lowest. The article considers the lower layer to ascertain the penalty coefficient based on the theory of two-hierarchical programming. Therefore the optimal value is obtained. As a result, the supply chain's total cost is still lower.

5. CONCLUSION

The article studies the coordination operation mechanism within supply chain between the neighboring layers. Through quantitative analysis，the article indicates that the manufacture can optimize his production plan and the penalty coefficient value based on the theory of two-hierarchical programming. Moreover, it can bring lower cost than the simple supply chain. However, how to ascertain the optimal order combination and the penalty coefficient value, and how to make tighter cooperation and lower cost aren't further discussed. This is to be studied in the future.

6. REFERENCES

[1] D. Pagel, "Management for Optimal Performance through Effective Coordination of the Supply Chain", Production and Inventory Management Journal, No.1, pp.66-70, Jan.1999.

[2] G.J. Van, K. Inderfurth, et al, Zijim, "Material Coordination in Stochastic Multi-echelon Systems", European Journal of Operational Research, Vol. 95, Jan.1996, pp.1-23.

[3] W. Zhao, Y.Z. Wang, "Coordination of Joint Pricing-production Decisions in a Supply Chain", IIE Transactions, Vol. 34, Jan. 2002, pp.701-715.

[4] T.D. Klastorin, K. Moinzadeh, et al, "Coordinating Orders in Supply Chains through Price Discounts", IIE Transactions, Vol.34 (8), Aug. 2002, pp.679-689,

[5] Z.M.Huang, X.Li.Susan, "An Analysis of Manufacturer-retailer Supply Chain Coordination in CooperativeAdvertising", Decision Sciences,Vol. 33(3), pp.469-494, Mar. 2000.

[6] H. MarilynM, P. Lawrence, et al, "Supply Chain Forecasting-Collaborative Forecasting Supports Supply Chain Management",Business Process Management Journal, Vol.6(5), pp.392-450, May 2000.

[7] L. Horvath, "Insight from Industry Collaboration: The Key to Value Creation in Supply Chain Management", International Journal of Supply Chain Management, Vol.6 (5), Oct. 2001, pp. 205-207.

[8] Monahan, "A Quantity Discount Pricing Model to Increase Vendor Profits", Management Science, vol. 30(6), pp.720–726, Dec. 1984.

[9] E.L.Porteus,S.Whang,"OnManufacturing/marketing Incentives", Management Science, Vol. 37(9), Sep. 1991, pp. 1166-1181.

[10] B.C. Arntzen, G..G..Brown, et al, "Global Supply Chain Management at Digital Equipment Corporation", Interfaces, Vol.25, pp.69-93, Jan. 1995.

[11] R. Kolisch, "Integration of Assembly and Fabrication for Make-to-order Production", International Journal of Production Economics, Vol.68, Jan. 2000, pp.287-306.

[12] C. Schneeweiss, K. Zimmer, "Hierarchical Coordination Mechanisms within The Supply Chain", European Journal of Operational Research, Vol. 153, Jan. 2004, pp.687-703.

[13] H. Lee, S. Whang, "Decentralized Multi-echelon Supply Chains: Incentives and Information", Management Science,Vol. 45(5), Oct. 1999, pp.633–640.

[14] C.H. Teng，Z.H. Li, The Theory and Application of Two-hierarchical Programming, Science Press，Beijing, 2002.

Dynamic Electronic Business Application System Based on SOA and Web Services

Wu Deng[1], Xinhua Yang[2], Huimin Zhao[1]
1. Software Institute, Dalian Jiaotong University
No.216 Xingfa Road, Lvshun Development Zone, Dalian, Liaoning, 116052, China
Email: dw7689@163.com , jiqimaozhm@163.com
And
2. Modern Railway Transportation Research Institute, Dalian Jiaotong University
Shahekou Zone, Dalian, Liaoning, 116028, China
Email: xhyang@263.net

ABSTRACT

In allusion to construct business condition based on the traditional techniques indicates some rather prominent problems, this paper brings forward web services which are the kernel and key to the solutions of the new generation dynamic E-business and enterprise application integration, and introduce several key techniques of web services as well as the relationships between them. A service-oriented framework is given and a service-oriented dynamic E-business application system based on SOA and web services is designed in this paper. Therefore, SOA and web services can provide exchange of information and system integration for dynamic E-business, support more complex types of web interaction and collaboration between enterprise trading partners, and consequently advance the developing of dynamic E-business.

Keywords: Web Services, Service-Oriented Architecture, Dynamic E-business Application System, XML, SOAP, WSDL, UDDI.

1. INTRODUCTION

E-business is a business trade activity in which both businesses do not seek a meeting but they carry on through using a simple, quick, low-cost electronic communication mode. And the dynamic E-business is the goal of E-business development; its characteristic is that software defers to commercial process modeling, schemes out reused and flexile components. Constructed commercial condition based on the traditional techniques has two quite prominent problems: The first, because the E-business participants do not use the same technical standard or system framework for service relation, it lacks the uniform commercial publishing and finding mechanism which causes mutually information exchange and cooperation to become extremely difficult. Second, as for regarding system itself, it also exists long development cycle, reconstruction difference, upgrade and maintenance difficulty and so on [1,2]. These all greatly have limited the further developing of E-business. But web services techniques, especially web services based on Service-Oriented Architecture (SOA) have provided the valid way for implementing dynamic E-business application framework. This paper takes web services as a foundation, has constructed a dynamic E-business application system framework based on SOA and web services, and it has expatiated its main techniques.

2. SOA AND WEB SERVICES

2.1 SOA

SOA is a solution for designing and setting up the loosing coupling software system, it can publish business functions in the manner of programmable and accessible services, and the other application can use these services by using the published and findable interface. So, the key conception of SOA is SERVICES, and any application of SOA is regarded as a service to be called and administrated. SOA defines design principles:
1) Functions are divided into less and reused model using modularization;
2) Loosing coupling clients and servers do not need close dependence;
3) Encapsulation encloses function modularization; inside widgets define and encapsulate commendably interfaces.

W3C defined SOA as following: the service provider delivers ultimately wanted results to service users by the aid of accomplishing a set of work. The ultimate results usually change the status of users, provider or both. To some extent, SOA is a model, which is used for designing, developing, deploying and administrating discrete logic units under computer environment. SOA has some advantages, including:
1) Be able to rapidly integrate the third software;
2) Adopt business flow management tools to deploy easily composition system;
3) Be able to update safely independence services through management and control platform;
4) According to services, it can partition development tasks and sustain distributed harmony development.

SOA has three roles, shown as Fig.1.Service broker registers and makes classifications of the published service providers, it also provides search service; Service provider publishes its own service, and also makes response to the request; Service requester seeks requisite service by using service broker, and makes use of the service. The components of SOA must have one or more of the above mentioned roles. These roles carry out such operations as “find”, “publish”, and “bind”. “Find” operation helps service requester to find special service aid by service broker; “Publish” operation helps service provider to register its own function and interface; “Bind” operation helps service requester to use the provided services in deed.

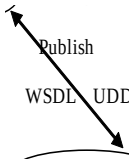


Fig. 1. Web service-based SOA framework

2.2 Web Services and Hierarchy Framework

Web services are a new generation of web application. It combines the advantages of the component-oriented methods and web techniques, and they can describe its own service. It can also publish, locate and transfer modularized application in web. The provided functions of web services may be simple, but it also contains extraordinary complicated business logic. Once web services are deployed, the other applications can find and request them [3]. There are the key techniques of web services, shown as Fig. 2.

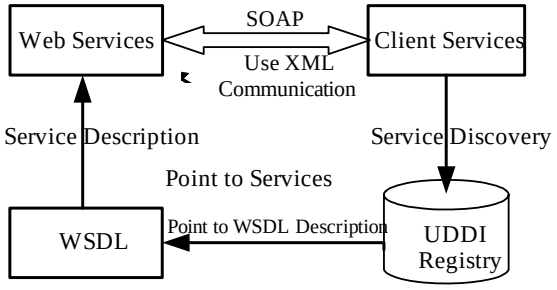


Fig. 2. Key techniques of web services

From the structural aspects, the kernel of web services is service. Web services represent a kind of implementation of SOA, and they are the most popular one. In addition, the three operations of SOA can only process when the components of SOA interact. Therefore some standardized techniques are used in web services, including UDDI, WSDL, HTTP, SOAP, and XML and so on. Web services become the best choice for developing application of SOA [4]. The hierarchy framework of web services includes network, transport, messaging, service description, service publication /discovery, service flow and so on, shown as Table 1.

Table 1. Hierarchy framework of web services

…	…	Management	QoS	Security
Service Flow	WFSL			
Service Publication /Discovery	UDDI			
Service Description	WSDL			
Messaging	XML、SOAP			
Transport	HTTP、FTP、SMTP			
Network	IPV4、IPV6			

Network
There are some internet protocols, which have IPV4, IPV6 and so on.

Transport
Web services transport messages and send data through network in order to implement the interactive purpose of the services request. Therefore, we define the kernel communication mechanism of the transporting data between web services, including HTTP, HTTPS, and SMTP and so on.

Messaging
Describe how to dispose message format. Extensible Markup Language (XML) is a standard meta language for describing data. It is a key point of web services that uses XML to express operation and data, in this way; the different systems could communicate and share data smoothly. Simple Object Access Protocol (SOAP) is a simple communication protocol based on XML in the distributing environment [5]. It is capable of permeating firewall breezily.

Service Description
Service description mainly provides the services interface description information and services deployment information and so on. Web service Description Language (WSDL) is a XML standard that supported by a lot of business associations, such as Microsoft, IBM and so on. By defining a standard XML syntax, WSDL could describe the web services in details, and then describe web services behaviors. The other applications can be informed:
1) How to find a special web services;
2) How to communicate with a web services (using SOAP);
3) How to generate a code to represent a web services.

Service Publication /Discovery
UDDI is a project developed by Microsoft, Ariba, IBM and other IT companies. It aims at how to find web services and how to find questions using web services. What the UDDI represents are Universal Description, Discovery, and Integration [6]. At present, it is becoming the standard for client applications to find web services and publish its own web service.

Service Flow
Service flow is the top of web service application enterprise integration. Because a service flow is usually completed by the mutual cooperation between web services. It associates with web services, completes the system goal according to the service flow logical sequence succession calling. The combination standard may refer to Business Process Execution Language For Web Services (BPEL4WS); workflow standard is brought forward by IBM, BEA, and Microsoft and so on in August 2002.

Security
Process security and coordinated transaction, which related with web services calling, and guarantees the reliable transmission of web services messages. Security is the basic component of the security web services, supports the security model for Kerberos, X509 and so on. At present, the development of this implementation technique is imperfect.

2. 3 The Characteristics of Web Service-based SOA

Communication striding over firewall
Web services exchange data using SOAP-based XML documents and common communication mode, such as

HTTP, FTP, and SMTP and so on. So, the communication based on web service techniques can permeate firewall breezily. Any device, which supports HTTP and XML, could access to web services, and any web services could access to other web services.

Implement easily
A large number of free tools are provided by IT companies, including IBM and Microsoft, to build and deploy web services. Existing JavaBeans and COM components can be easily translated to web services for providing services.

Loosing coupling
Web services have a loosing coupling structure. Any one of web services can change its own running mechanism, but does not affect other services. When web services are updated, the clients who call the services can get the updating with that.

Software along with data can be reused
Web services allow codes to be reused, as well as data that behind the codes.

These characteristics can expand greatly web services functions, truly implement mutual operation, and they may use the loosing coupling model to use and expand each kind of data and the service resources. The specific function is completed by the dynamic binding different services. In researching and designing the dynamic E-business application framework, we find an almost perfect solution for many existing and likely-to-appear problems.

3. DYNAMIC E-BUSINESS APPLICATION SYSTEM BASED ON SOA AND WEB SERVICES

3.1 Traditional Dynamic E-business Framework

The E-business has developed for several stages: From the static homepage, using the interactive network tabulation, carrying on the enterprise to the client (B2C) and the enterprise to the enterprise (B2B), to the developing dynamic E-business. At present, there are still the half enterprises that use static web and simple E-business application. The E-business technologies of previously appeared EDI and EAI all bases on the complex application linking. It links the users, the E-business application and the additional information systems through the program codes[7,8]. There is a traditional E-business framework, shown as Fig. 3.

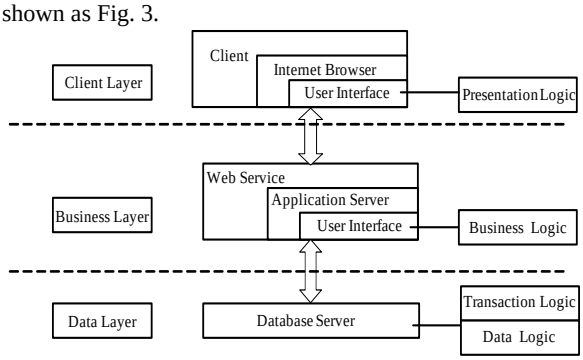


Fig. 3. Traditional E-business framework

Because the traditional E-business framework does not have a nicer integrated ability, can not validly solve the large quantity expense because of the E-business flow changing, and it is unable to validly solve the customizing demand of each category of users. Therefore, it is an optimal scheme that the dynamic E-business of the foundation and implementation should solve the questions. The dynamic E-business is the next-generation E-business, which bases on the emphasizing comprehension and infrastructure of the B2B.It can create the best benefit for the interior and exterior enterprises through adjusting the Internet standard and general infrastructure [9]. There is dynamic E-business application architecture, shown as Fig. 4.

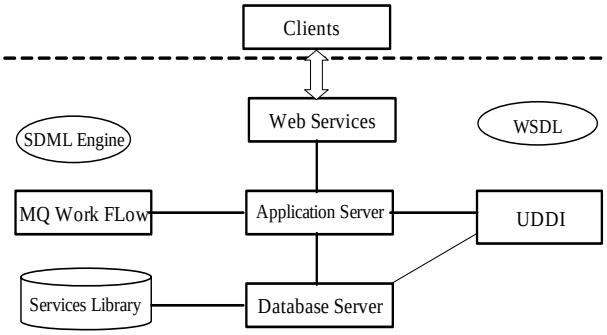


Fig. 4. Dynamic E-business framework

In order to implement the dynamic E-business, the enterprise may establish the massive mutual applications between the services flows and the many partners in the B2B. The services application can directly interact through both sides agreeing interaction or integrated protocol. Due to the complexity of the application integration, the frame is suitable for the kernel application process between the specific commercial partners.

This kind of dynamic E-business framework has some advantages of the relative independent and convenient operation, at the same time it also has some disadvantages, such as too substantial investment, data sharing difference, discommodious system upgrade and information exchange, nonsupport long-distance designing development and so on.

3.2 Dynamic E-business Application System Framework Based on SOA and Web Services

The dynamic E-business based on web services provided application independency, namely the application may automatically, immediately search correlation application of the different commercial entities on Internet. We can select optimal plan through the comparison selection in order to meet greatly the clients' needs. There is a dynamic E-business application system framework based on SOA and web services, shown as Fig. 5.

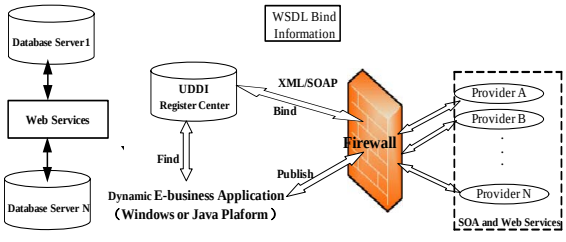


Fig. 5. Dynamic E-business application system framework based on SOA and web services

In the dynamic E-business application system framework based on SOA and web services, each software resource of

web services may be regarded as a pellet software component. It can set up a distributional application program and assemble each kind of E-business program by using the web services pellet software components for the constructed module. A web service can receive the request to complete a series of specific tasks, and use open communication standard to respond to this request in order to guarantee mutually operation. Web services themselves are a possible web services integration system. The dynamic E-business application system framework based on SOA and web services has the following basic principles [10]:

1) The services interface of the software resource must be openly published accessible;
2) The integration of the software resource must be the loosing coupling form;
3) The abundance of the software component resource should improve the flexibility and personalization of the kernel commercial process;
4) The application is constructed by integrating core commercial process and exterior software components /resources;
5) The messages transmission between programs should comply with the open Internet standard;
6) The reusing exterior software resource should reduce the cost and improve productive efficiency for services customers;
7) The software may be sold as services.

The dynamic E-business application system framework based on SOA and web services has such characteristics:

Easy implementation

This point is determined by web services characteristics. As a result of promotion of many companies such as Microsoft, SUN, IBM, now it has already had the massive free tools to fast generate and deploy web services, moreover, it can easily change existing system into other system that can provide services in the way of web services. Under Microsoft Windows.NET Framework, we have implemented the prototype. Windows.NET Framework is windows component used in constructing and running the-next-generation software application program. It supports more than 20 kinds of different programming languages; it can control various insertion operations related to software development; it can conveniently construct, deploy and manage secure and reliable as well as high efficient application program. Using WinForm, WebForm, XML Web Services, ASP.NET, ADO.NET based on the frame of Windows.NET Server 2000 database system, we have implemented the partial related functions.

Easy integration with other system

Once a web service is published, it will be found and used what it serves by the requested service application according to its introduction. Even if the service is changed, the application does not have to change, which makes the integration with other system.

System being independent from developing language and operating system

Due to adopting HTTP and XML, the restriction from special developing language and operating system is inexistent. So, the service such as analysis and material selection services are available for any application, which is developed under any platform and in any language.

4. CONCLUSIONS

The dynamic E-business is the goal of E-business development, and web services are theirs key techniques. This paper has expatiated the deference between the dynamic E-business framework based on SOA and web services and the traditional E-business using the frame lies in that the former uses web services to take its frame the standard. XML, SOAP, WSDL and UDDI technologies are the web service kernel, they provide open, supporting plug and play, the application of independent of any language and platform for web service, thus it promotes the development of the dynamic E-business. The dynamic E-business may integrate enterprise's value chain in highly efficient way, thus it endows with more flexibilities, productivity and creativity for the enterprises. Only when the enterprises set up the dynamic E-business system, are they able to adapt to the constantly changing market, can they grasp market opportunities. As a result of doing so, it brings enormous benefits for the commercial integration.

1. REFERENCES

[1] China E-commerce Yearbook. (2002 and 2003). Beijing: China E-commerce Year Book Editing House (in Chinese).

[2] J.Y. Zhu, B.Z. Yang, "Research paper of e-business development all over the world", United Nations Trade and Development Board [M]. 2003.

[3] R. Anzbock, S. Dustdar, et al, "Software Configuration, Distribution, and Deployment of Web Services". SEKE July 2002 Italy.

[4] L. Arsanjani, B. Hailpern, "Web Services Promises and Compromises," March 2003 Queue, Volume 1 Issue 1.

[5] UDDI version 2. 04 API specifications UDDI published specification, UDDI. Org [S]. http://uddi.org/pubs/ProgrammersAPI2V2.042Published 220020719. PDF, 2002.

[6] E. Naveen, C. David, et al, "Enterprise Application Integration in the electronic business world," Computer Standards &Interfaces 25(2003): 69-82.

[7] D. Fensel, C. Bsussler, "The Web service modeling framework WSMF,"Electric Commerce Research and Application 2002, 1(2): 113-137.

[8] T.B. Chen, W.Wu, "Design and implementation for dynamic e-business integration based on J2EE-enabled Web services," Dalian Ligong Daxue Xuebao/Journal of Dalian University of Technology, 2005, 9.

[9] H.F. Bian, J.H. Xu, et al, "Agent and web service based on dynamic business process integration,"Tsinghua Univ.(Sci&Tech), No.3: 429-432, 2003.

[10] Y. Rao, B.Q. Feng, "New generations of network architecture-web services review," Computer Science, 2004, 31(5): 1- 4 (in Chinese).

Quantum-behaved Particle Swarm Optimization for Security Constrained Economic Dispatch

Daojun Wang, Wenbo Xu
Center of Intelligent and High Performance Computing, School of Information Technology,
Southern Yangtze University, Wuxi, 214122
Email: wdj_513@hotmail.com

ABSTRACT

This paper proposes a quantum-behaved particle swarm optimization (QPSO) method for solving the economic dispatch (ED) problem in power systems. Many inequality constraint and equality constraint characteristics of the generate units and the system, such as the generation of each unit should between its minimum ($P_{Gi}^{\min}$) and maximum ($P_{Gi}^{\max}$), the total system generation includes load demand of the system and the transmission losses and so on, they all indicate certain nonlinear characteristics of the system. The proposed approach has been tested on a representative system: IEEE 30 bus, the experiments indicate that the QPSO algorithm is better than the PSO algorithm and Hopfield neural network.

Keywords: Economic Dispatch, Security Constrained, Quantum-behaved Particle Swarm Optimization.

1. INTRODUCTION

Economic dispatch (ED) problem is one of the fundamental issues in power system operation, which pertains to optimum generation in an interconnected power system to minimize the cost of generation subject to relevant system constraints [1,2]. This paper investigates the application of Quantum-behaved Particle Swarm Algorithms (QPSO) for determining to the optimal loading of generates units in a power system (IEEE 30 BUS SYSTEM).

A wide variety of optimization techniques have been applied in solving the ED problems such as Linear Programming (LP), Quadratic Programming (QP), Simulated Annealing (SA) and Genetic Algorithms (GA) [3, 4]. LP methods are fast but associated with the piecewise linear cost approximation; QP methods have a problem of convergence and algorithmic complexity; SA and GA methods are new probabilistic heuristic algorithms which have been successfully used to solve ED problem, GA method is usually faster than SA method because of its parallel search ability, but recent research has identified that it has degradation with highly epistatic objective functions [4].

Particle Swarm Optimization (PSO) is a modern population based evolutionary search technique, which first introduced by Kennedy and Eberhart, it has successfully optimized a wide range of continuous functions and have generated higher quality solutions in shorter time, at the same time, it has more stable convergence characteristic than other stochastic methods [5]. Quantum-behaved Particle Swarm Optimization (QPSO) is first proposed by Sun, it is more effective for actual social system, because a social system is far more complex than that formulated by anyone of equation and even the thinking mode of an individual of the social system is so intricate that a linear evolvement equation is not sufficient to depict [6, 7]. In Sun's paper, the experiment results indicate that the QPSO algorithm work s better than standard PSO on several benchmark functions and it's a promising algorithm. The experiment results of ED problem are shown in Section IV, and it prove that QPSO is better than PSO in this case.

2. PROBLEM DESCRIPTION

The ED problem is an optimization problem that determines the power output of each effective bus to satisfy the system load demand that will result in a least cost system state.

2.1 ED Problem Formulation

The objective of the ED problem is to minimize the total system cost, and so the objective function is:

$$C = \min(\sum_{i=1}^{N_g} F_i(P_{Gi}))$$

where $F_i(P_{Gi})$ is the cost function of the i[th] generating unit, P_{Gi} is the real output of the i[th] generating unit, and N_g is the total number of generate units in this power system.

The cost function of each generate unit is related to the actual power injected to the system, and is typically modeled by smooth quadratic function as:

$$F_i(P_{Gi}) = a_i + b_i P_{Gi} + c_i P_{Gi}^2$$

Where a_i, b_i and c_i are the cost coefficients of the i[th] generating unit [2, 8-13].

2.2 System Transmission Losses

The most popular approach for finding an approximate value of the losses is by way of Kron's loss formula as follow, which represents the losses as a function of the output level of the system generate unit:

$$P_L = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} P_{Gi} B_{ij} P_{Gj} + \sum_{i=1}^{N_g} P_{Gi} B_{i0} + B_{00}$$

Where B_{ij}, B_{i0}, B_{00} are known the loss coefficient method or the B- coefficient method. Using the matrix notation, the loss formula can be expressed as [3]:

$$P_L = P^T[B]P + B_0 P + B_{00}$$

2.3 Inequality Constraint

The generation of each unit should between its minimum ($P_{Gi}^{\min}$) and maximum ($P_{Gi}^{\max}$) production limits, which directly related to the design of the machine. That is, the following inequality constraint for each generate unit should satisfied a pair of inequality constraints, as follows:

$$P_{Gi}^{\min} < P_{Gi} < P_{Gi}^{\max}, \text{ i =1,..., } N_g$$

Where P_{Gi} is the real output of the i[th] generating unit [2, 3, 9-13].

2.4 Equality Constraint

The total system generation includes load demand of the system and the transmission losses, that is:

$$\sum_{i=1}^{N_g} P_{Gi} = P_D + P_L$$

Where $\sum_{i=1}^{N_g} P_{Gi}$ is the total system generation, P_D is the total system load demand, P_L is the transmission losses [2, 3, 12, 13].

3. Particle Swarm Optimization and Quantum-behaved Particle Swarm Optimization

3.1 Overview

Particle Swarm Optimization (PSO) algorithm, originally introduced by Kennedy and Eberhart in 1995, is a population-based evolutionary computation technique. It is motivated by the behavior of organisms such as fishing schooling and bird flock. In a PSO system, each particle corresponding to individual of the organism is a candidate solution to the problem at hand. Particles of the population fly around in a multi-dimensional search space, to find out an optimal or sub-optimal solution by competition as well as by cooperation among them [5].

Quantum-behaved Particle Swarm Optimization (QPSO) algorithm, proposed by Jun Sun, the search space and solution space of the problem are two spaces of different, and the search space of an individual particle at each iteration is the whole feasible solution space of the problem, the experiment in results Sun's paper indicate that the QPSO works better than standard PSO on several benchmark functions [6, 7].

3.2 Particle Swarm Optimization (PSO)

The PSO algorithm maintains a population of particles, where each particle represents a potential solution to an optimization problem. Let S be the size of the swarm, each particle i can be represented as an object in the D-dimensional space as:

$$X_i(t) = (X_{i1}(t), X_{i2}(t), ...X_{iD}(t))$$

and $V_i(t) = (V_{i1}(t), V_{i2}(t), ...V_{iD}(t))$.

The particles move according to the following equation:

$$V_i(t+1) = \omega V_i(t) + c_1 * rand1() * (pbest_i - X_i(t)) + c_2 * rand2() * (gbest - X_i(t)),$$

i=1,2…S,

$$X_i(t+1) = X_i(t) + V_i(t+1),$$

Where ω is the weight factor, c_1 and c_2 are the acceleration constant, $rand1()$ and $rand2()$ are the uniform random value in the range [0,1], $V_i(t)$ is the velocity of particle i at iteration t, $X_i(t)$ is the current position of particle i at iteration t.

The weight factor ω provides a balance between global and local explorations. The constants c_1 and c_2 represent the weighting of the stochastic acceleration terms that pull each particle toward the pbest and gbest position. Low values allow particles to roam far from the target regions before being tugged back. On the other hand, high values result in abrupt movement toward, or past, target regions [5].

3.3 Quantum-behaved Particle Swarm Optimization (QPSO)

The dynamic behavior of the particle in Quantum-behaved Particle Swarm Optimization (QPSO) is widely divergent from that of the particle in traditional PSO system in that the exact values of V and X can't be determined simultaneously. In QPSO, the particle moves according to the following equation:

$$mbest = \frac{1}{S}\sum_{i=1}^{S} pbest_i = (\frac{1}{S}\sum_{i=1}^{S} pbest_{i1}, \frac{1}{S}\sum_{i=1}^{S} pbest_{i2}, ... \frac{1}{S}\sum_{i=1}^{S} pbest_{iD})$$

$$p_i = \varphi * pbest_i + (1-\varphi) * gbest,$$

φ =rand1()

$$X_i = p_i \pm \alpha * |mbest - X_i| * In(1/u)$$

u=rand2()

$$\alpha = (\alpha_1 - \alpha_2) * \frac{(MAXITER - t)}{MAXITER} + \alpha_2$$

Where mbest is the mean best position among the particles, p_i is a stochastic point between $pbest_i$ and $gbest$ stand for the local attractor of the particle, φ and u are random values in the range [0,1], α is a parameter of QPSO that is called Contraction-Expansion Coefficient, α_1 and α_2 .are the initial and final values of the parameter α , t is the current iteration number and MAXITER is the maximum number of allowable iterations [6,7].

The only parameter in the algorithm is Contraction-Expansion Coefficient, α , which is called Creativity Desire of the particle and works on individual particles convergence speed and the performance of the algorithm [7].

3.4 Implementation of the QPSO algorithm

The proposed QPSO approach has been tested on IEEE 30 bus system using the following steps:

1) Choose the population size and number of generation
2) Select the real power of each generate unit except slack bus generate units in the system, as state variables (X_i)
3) Generate randomly S particles { $X_i(0)$, i= 1, 2, …

S}, where $X_{ik}(0)$ is generated by random selecting a values with uniform probability over the k^{th} optimized parameter search space $[X_{\min}, X_{\max}]$

4) Evaluate the fitness of each particle according to the objective function including the total generation cost $\sum_{i=1}^{N_g} F_i(P_{Gi})$ and the security penalty functions. The constraints includes the load equality constraint and inequality constraint for each generate unit, so the F_T is given by:

$$F_T = \sum_{i=1}^{N_g} F_i(P_{Gi}) + K_1 * abs(\sum_{i=1}^{N_g} P_{Gi} - P_D + P_L) + \sum_{i=1}^{N_g} (K_2 * abs(P_{Gi} - P_{\min})) + \sum_{i=1}^{N_g} (K_3 * abs(P_{Gi} - P_{\max}))$$

Where K_1 is a constant penalty factor for the load equality constraint, K_2 and K_3 are the inequality constraint penalty factors given by:

$$K_{2i} = \begin{cases} 0, if: P_i \geq P_{i,\min} \\ 1, if: P_i < P_{i,\min} \end{cases},$$

and $$K_{3i} = \begin{cases} 0, if: P_i \leq P_{i,\max} \\ 1, if: P_i > P_{i,\max} \end{cases}$$

5) For each particle, set the initial position as its pbest position, and set the particle of best fitness as the gbest position.
6) Set time counter t=1
7) Calculate the mbest by:

$$mbest = (\frac{1}{S}\sum_{i=1}^{S} pbest_{i1}, \frac{1}{S}\sum_{i=1}^{S} pbest_{i2}, ... \frac{1}{S}\sum_{i=1}^{S} pbest_{iD})$$

8) Recalculate the position of each particle by:

$$X_i = p_i \pm \alpha * |mbest - X_i| * In(1/u)$$

$$p_i = \varphi * pbest_i + (1-\varphi) * gbest$$

9) For each particle X_i, it's the fitness is better than the $pbest_i$, update $pbest_i$; update gbest if the best fitness of this generation is better.
10) Set time counter t=t+1, if t> MAXITER, then the program stop; else go to step7.

4. Results and Analysis

The proposed method (QPSO) has been applied to IEEE 30 BUS system compared with the PSO method and the Hopfield Neural network.

Table 1 shows that the fitness of PSO method and QPSO method in 50 times:

Table 1. 50 times date

TIME	1	2	3	4	5
PSO	805.63	805.45	804.78	803.80	806.93
QPSO	807.28	802.45	802.35	803.77	803.25
TIME	6	7	8	9	10
PSO	808.30	806.01	809.30	805.67	803.26
QPSO	804.50	802.79	802.62	802.42	803.69
TIME	11	12	13	14	15
PSO	804.41	802.47	812.22	802.68	802.64
QPSO	804.83	805.71	808.99	803.00	802.13
TIME	16	17	18	19	20
PSO	803.20	813.05	805.32	802.62	804.61
QPSO	803.43	805.28	803.47	803.00	801.95
TIME	21	22	23	24	25
PSO	803.19	801.99	803.76	813.57	803.24
QPSO	805.69	804.16	803.55	802.34	804.06
TIME	26	27	28	29	30
PSO	811.26	807.21	804.92	803.65	804.76
QPSO	805.40	803.02	802.09	802.90	810.67
TIME	31	32	33	34	35
PSO	809.13	806.30	803.85	822.46	816.55
QPSO	802.89	808.88	804.23	812.09	803.39
TIME	36	37	38	39	40
PSO	803.54	803.79	814.26	807.91	804.18
QPSO	811.37	802.76	803.77	802.24	805.93
TIME	41	42	43	44	45
PSO	807.80	816.47	804.39	807.77	803.84
QPSO	803.08	802.46	809.56	803.99	802.63
TIME	46	47	48	49	50
PSO	805.88	802.15	801.96	817.66	813.03
QPSO	802.27	802.10	806.23	812.24	804.95

Fig.1 show that the average evolution of PSO method and QPSO method.

It is evident that the QPSO method is better than PSO method both in speed and final fitness. The mean fitness of QPSO is 804.5568 compared with 806.8566 using PSO. Referring to standard deviation, QPSO is 2.7420 and PSO is 5.2159.

Table 2 shows the best fitness of PSO and QPSO compared with the Hopfield neural network:

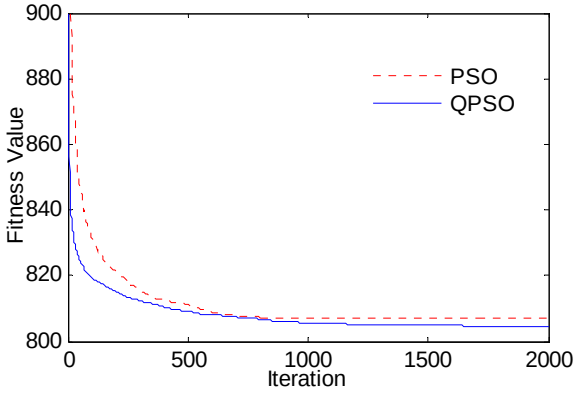


Fig. 1. AVERAGE RESULTS OF 50 TIMES

Table 2. THREE TYPES ALGORITHM

	Hopfield	PSO	QPSO
Best Fitness	801.9642	801.9602	801.9482

5. Conclusion

This paper presents an enhanced PSO method--QPSO method for the security constraint economic dispatch problem. The proposed method utilizes an enhanced global searching method for complex problem, and we successful used it in the ED problem. The test system is IEEE 30 Bus system. The results show that QPSO is better than PSO method.

6. REFERENCES

[1] J.W. Allen, F.W. Bruce, Power Generation, “Operation and Control, Second Edition,” Tsinghua University Press, Beijing, 2003

[2] S.H. Rukmi, “Active Security Constrained Optimal Power Flow Using Modified Hopfield Neurak Network,” Doctor Paper, Dalhouseie University, 2001

[3] K. Price, "Differential Evolution: A Fast and Simple Numerical Optimizer," 1996 Biennial Conference of the North American, IEEE, 1996, pp. 524-527.

[4] E.P.G. Rau'l, R.C.M. Jose, “Differential evolution based economic environmental power dispatch”, Proceedings of the 37th Annual North American, IEEE, 2005, pp. 191-197.

[5] J. Kennedy, R.C. Eberhart, “Particle swarm optimization,” in Proceedings of the IEEE International Conference on Neural Networks, IV,1995, pp. 1942-1948.

[6] J. Sun, W.B. Xu, et al, “A Global Search Strategy of Quantum-Behaved Particle Swarm Optimization,” Conference on Cybernetics and Intelligent Systems, Proceedings of the 2004 IEEE, Singapore, 2004, pp.111-116.

[7] W.B. Xu, J. Sun, “Adaptive Parameter Selection of Quantum-Behaved Particle Swarm Optimization on Global Level,” ICIC 2005, Springer-Verlag Berlin Heidelberg 2005

[8] K.P. Rohi, K.S. Swarup, “Partilce Swarm Optimization for Security Constrainted Economic Dispatch,” ICISIP 2004, IEEE, 2004, pp.7-12.

[9] P. Jong-Bae, Member, IEEE, Ki-Song Lee et al, “A Particle Swarm Optimization for Economic Dispatch With Nonsmooth Cost Functions,” IEEE IEEE Trans. on Power Systems, VOL. 20, NO. 1, IEEE, 2005, pp.34-42.

[10] Y. Hirotaka, K. Kenichi, et al, “A Particle Swarm Optimization for reactive power and voltage control considering voltage security assessment,” IEEE Trans. on Power Systems, Vol.15,No.4, IEEE, 2001, pp. 1232-1239.

[11] G. Zwe-Lee, “Particle Swarm Optimization to Solving the Economic Dispatch Considering the Generator Constraints,” IEEE Trans. on Power Systems, VOL. 18, NO. 3, IEEE, 2003, pp.1187-1195.

[12] K.P. Rohit, K.S. Swarup, “Partilce Swarm Optimization for Security Constrainted Economic Dispatch,” ICISIP 2004, IEEE, 2004, pp.7-12.

[13] S. Orero, M.R. Irving, “A Genetic Algorithm Modelling Framework and Solution Technique for Short Team Optimal Hydrothermal Cheduking,” IEEE Transactions on Power Systems, Vol. 13, No. 2, IEEE, 1998, pp.501-518.

Research on the Coal Logistics Transport Corridor Based on Supply Chain Management*

Jing Yu
College of Communications and Transportation;
Shanghai Maritime University, Shanghai 200135 , China
Email: jingyu@cct.shmtu.edu.cn

ABSTRACT

In order to reduce the coal logistics cost, distribute the coal logistics evenly in transport network and enlarge transport capability, the problems on the coal logistics transport corridor and coal logistics distribution center location are researched based on the theory of supply chain management. The model of coal logistics distribution centers location is given comprehensively considering the cost factors and non-cost factors. Then based on order-transport-inventory chain, the mathematical model is established.

Keywords: Coal Transport Corridor, Distribution Center Location, Mathematical Model, Supply Chain Management, Fuzzy Judgments.

1. INTRODUCTION

As a basic source of energy in China, coal is dominant in production and consumption of primary energy, which will not change for a long time. Meanwhile, the geographic distribution of coal production and consumption is very unbalanced in China, and the basic direction of transportation of the coal flow is from northern or western China to southern or eastern part of the country. This induces high demand for coal transportation. The quantity of coal is about 45 percent in the gross domestic freight transportation, and the average transit distance is about 550 kilometers. Now the coal production and consumption are highly dependent on transportation [1, 4].

The problems of transportation bottleneck and transportation incapability are emerging in carrying coal for the moment. This makes it necessary to optimize the transportation system by using advanced methodology, such as Supply Chain Management (SCM) theory. However, little attention is given in applying SCM to solve the problems of coal logistics [3, 5].

The problems on transport corridor of coal logistics are very important in guaranteeing the normal operation of the coal enterprise production. Consequently, under the idea of SCM, such problems can be studied from a perspective of coal logistics along transportation corridor.

2. THE CONCEPT OF TRANSPOR CORRIDOR ON COAL LOGISTICS AND THE STRUCTURE

2.1 The concept of transport corridor

Transport corridor is based on the concept of transportation network structure, and it relates to transportation network but different from it. Transport corridor forms a banding aisle, which is joined by congested traffic flows. In function, transport corridor is the framework of the transportation network, which connects the main hubs, stations and distribution points, performing the function of exchanging traffic flow among the transportation network [2].

We can comprehend the concept from the following aspects:

1) Coal transport corridors connect coal production area with highly concentrated consuming zones, and complemented by parallel transportation line that is different in transportation modes. They offer logistics services of coal transportation together and have strong attraction to the coal consumers.

2) Transport corridors include not only all kinds of transportation lines, but also harbor docks, hubs, distribution centers and corresponding backup service facilities.

2.2 The structure of transport corridor

Each transport corridor consists of node subsystem and thread subsystem, as shown in Fig. 1. .

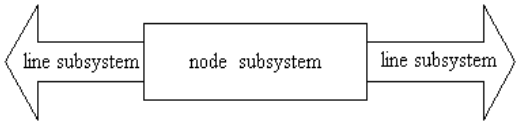


Fig. 1. Basic component of transport corridor

Transport corridor of coal logistics is also made up of node subsystem and thread subsystem. After establishing the coal logistics distribution centers, coal is flowing along the path: coal enterprise to coal logistics distribution centers, and finally to consumers. Because the path from coal logistics distribution centers to the consumers is relatively smooth, the problems of transport corridor from coal enterprises to logistics distribution centers are mainly discussed in this paper. And logistics distribution centers can be considered terminal consumers. So the location of coal distribution centers and the structure of coal transport corridor are important problems under study.

3. THE OPTIMIZING MODEL FOR THE COAL DISTRIBUTION CENTER LOCATION

Coal logistics distribution center becomes the biggest consumer that is the main part of coal logistics node subsystem in transport corridor. Therefore, the location of coal distribution center is one of the most important problems need to be solved.

There are many factors determining the location of coal logistics distribution center, which can be divided into cost factors and non-cost factors. The capturing of these factors is given as follows [6, 8].

3.1 The cost factors

The cost factors that can be analyzed quantitatively mainly

* Authors: Yu Jing, Ph.D, female, Associate Professor, interested in system programming and management on transportation and logistics.

include the transportation cost, distribution cost and the expense of logistics facility. As for the cost factors, the mathematical model can be established to optimize the location in order to obtain the lowest logistics cost. The objective function can be show as follows:

$$\min f(x)= \sum_{hijk}(A_{hij}+B_{hjk})\times X_{hijk}+\sum_{j}F_j\times Z_j+\sum_{hj}S_{hj}(\sum_{ik}X_{hijk})+\sum_{hk}D_{hk}(T_{hk}) \quad (1)$$

where,

h - the patterns of coal products, (1,2,…, H).

i - the coal enterprise, (1,2,…, I).

j - the coal distribution center, (1,2,…, J).

k - the coal consumer in distribution area,(1,2,…, K).

A_{hij} - the units transportation cost when transporting coal product h from i to j .

B_{hjk} - the units transportation cost when transporting coal product h from j to k .

X_{hijk} - the amount of transporting coal product h to consumer k from i to j .

F_j - the average fixed management cost of coal product in the distribution center.

Z_j - control parameters, if $\Sigma_{hik}X_{hijk}$ >0, it is equal to one or else zero.

$S_{hj}(\Sigma_{ik}X_{hijk})$ - partly variable cost which is produced by j persevering coal product h .

$D_{hk}(T_{hk})$ - the losing cost because of delay time when distributing product h to k .

$f(x)$ - the total cost of logistics cost.

3.2 The non-cost factors

The non-cost factors cannot be analyzed in quantity, but can produce influence that is to be reckoned with on the location choice of the coal logistics distribution center. These non-cost factors consist of regional economy condition, transportation condition, environment factors, the requirement of environment protection, city planning factors, area policy factors and so on.

We can adopt fuzzy math theory to analyze the effect of the non-cost factors based on the above Eq. (1). The main idea is to consider the non-cost factors as the factor set and the theoretically optimized resolution quantities as evaluation set (supposing there are1, 2, 3…, n optimized locations). The subjection degree vectors corresponded with the non-cost factors can be obtained based on the optimized location. Then comprehensive fuzzy evaluation and sorting are made for all the optimized resolutions with the fuzzy closeness degree, and finally get the optimized feasible solution.

The following is the main steps.

1) Factors set

Factors set is the collection of regional economy condition, transportation condition, environment factors, investment condition, city planning, area policy and other non-cost factors, which have important influence on the location of coal logistics distribution centers. If u_i stands for corresponding non-cost factors, then the factors set is:

$$U=\{u_1,u_2,\ldots,u_m\} \quad (2)$$

2) Weights set

The distribution of weights is largely a matter of judgment. When the location of coal logistics distribution center is different, the corresponding weights are also different. Therefore, we can apply the method of AHP and Delphi to distribute weights based on different locations. If a_i stands for the weights of u_i , then the weights set is:

$$A=\frac{a_1}{u_1}+\frac{a_2}{u_2}+\cdots+\frac{a_m}{u_m}=(a_1,a_2,\ldots,a_m) \quad (3)$$

3) Evaluation set

Building the comprehensive fuzzy evaluation is to modify the deficiency of Eq. (1), which ignores the non-cost factors. Take the theoretically optimized locations as the elements of evaluation set, and then choose the best plan while considering the non-cost factors v_i as constraints.

$$V=\{v_1,v_2,\ldots,v_n\} \quad (4)$$

4) Evaluation of single element

The process of the evaluation of single element is the process of fuzzy mapping. $f: U\rightarrow\varphi(V)$, and it also means to get evaluation matrix R using the closeness degree $r_{ij}\in[0,1]$ of u_i to v_j when evaluating.

$$R=(r_{ij})_{m\times n} \quad (5)$$

5) Fuzzy integration evaluation

Integrated evaluation is fuzzy linear conversion from U to V , $B=A\cdot R$, $B=(b_1,b_2,\ldots,b_n)$ is the result of evaluation . The bigger the value, the higher the feasibility of coal logistics distribution center location.

4. THE OPTIMIZED MODEL OF COAL LOGISTICS TRANSPORT CORRIDOR STRUCTURE

Under the idea and theory of SCM, coal transportation is not an independent link, but a part of supply chain that has close relation with order and storage. They influence one another. The amount of storage affects the amount of order, and the amount of order decides the quantity of transportation. In order to reduce the cost of coal logistics transportation, balance the distributing amount of coal logistics in the road network and enlarge the transportation capability, the mathematical model is built based on order-transport-inventory chain, comprehensively considering coal price, coal type, transportation ability, transportation fee, transportation reliability, the capability and fee of storage, and some other factors[7].

The following are the notations of variables of the model:

I - the amount of the selected coal enterprise.

J - the amount of coal logistics distribution centre.

M - transportation mode in broad sense.

K - the amount of coal's category.

T - the amount of time unit.

A_{it}^{k} - the coal amount of enterprise i supplying coal type

k in the unit of t.

B_{jt}^{k} - the amount of coal logistics distribution centre j requiring coal type k in the unit of t.

C_{ijtm}^{k} - the unit transportation cost from i to j produced by ferrying coal category k through the transportation mode m.

d_{jt}^{k} - the unit storage cost of j storing coal category k during the period of t.

P_j - the storage capability of j.

α_{it}^{k} - the price of coal category k produced in the period of t.

ω_{j0} - the initial storage of j.

q^{k} - the transportation capability required of unit coal category k.

γ_{ijtm}^{k} - the capability of transportation from i to j in the period of t adopting the transportation mode m.

Δt - the length of each time period.

X_{ijmq}^{k} - the amount of the coal type k ordered and sent from i to j in the beginning of q.

V_{jt}^{k} - the amount of coal type k stored by j in the end of t.

η_{ijmqt} - control parameters, if int $(\beta_{ijtm}/\Delta t) = t - q$, $\eta = 1$, else, $\eta = 0$.

It shows whether the booked coal at the begging of q is arrived from i to j in the end of t.

If it arrived in the period of time, the fees of order and transportation turn into the cost of logistics, else they cannot be calculated.

The objective function can be expressed as follows:

$$\min\left\{\sum_{k=1}^{K}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{m=1}^{M}\sum_{q=1}^{T}\eta_{ijmqt}(\alpha_{iq}^{k}+C_{ijtm}^{k})X_{ijmqt}^{k}+\sum_{t-1}^{T}\sum_{k=1}^{K}\sum_{j=1}^{J}d_{jt}^{k}V_{jt}^{k}\right\} \quad (6)$$

Subject to:

$$\begin{cases} V_{jt-1}^{k}+\sum_{m=1}^{M}\sum_{i=1}^{I}\sum_{q=1}^{t}\eta_{ijmqt}X_{ijmqt}^{k}-V_{jt}^{k}\geq B_{jt}^{k} & (\forall\ k, j, t) \\ \sum_{m=1}^{M}\sum_{j=1}^{J}\sum_{q=1}^{t}\eta_{ijmqt}X_{ijmqt}^{k}\leq A_{it}^{k} & (\forall\ i,\ k, t) \\ \sum_{k=1}^{K}V_{jt}^{k}\leq P_j & (\forall\ j,\ t) \\ \sum_{k=1}^{K}q^{k}\eta_{ijmt}X_{ijmt}^{k}\leq\gamma_{ijtm}^{k} & (\forall\ i,\ j, t, m) \\ V_{j0}^{k}=\omega_{j0}^{k} & (\forall\ j, k) \\ V_{jt}^{k}\geq\omega_{jt}^{k} & (\forall\ j, t, k) \\ X_{ijmqt}^{k}, V_{jt}^{k}\geq 0 & (\forall\ i, j, k,\ m, q) \end{cases}$$

Eq. (6) is a multi-goods network flow model. There are lots of algorithms to solve the model, and the method of dual matrix is generally used.

5. CONCLUSION

1) At present, the research on coal logistics is rare, especially on transport corridor of coal logistics based on SCM. It is necessary to restructure the coal logistics mode and establish the model of coal logistics distribution center location.

2) Considering the cost factors and non-cost factors comprehensively, the model of coal logistics distribution centers is given based on fuzzy theory.

3) Comprehensively considering coal price, coal type, transportation capability, transportation fee, transportation reliability, the storage capability and storage fee, and some other factors, the mathematical model is built based on order-transport-inventory chain, for the objective of reducing the cost of coal logistics shipment , balancing the distributing amount of coal logistics in the road network and expanding the transportation capability.

6. REFERENCES

[1] J. Yu, Q.X. Cai, “The Optimization Model of The Selecting Location on The Marshalling Station of Mine Area Railway”, Journal of China University of Mining and Technology, 2003,32(4): 452-454.

[2] X.F. Tang, “The Research of Coal Logistics Transport Corridor Problems”，XuZhou: The College of Energy and Safety Engineering of CUMT, 2005.

[3] Y.G. Li, G.P. Xia, “Batch Order Model in Transnational Supply Chain”, CIMS, 2005, 11(1): 79-84.

[4] J. Yu, G.Z. Liu, “The Reliability Study of Transportation Process in Mining Loading Station”, Journal of Liaoning Technical University, 2005, 24(supple): 250-251.

[5] S.H. Ma, Y. Lin, et al, “SupplyChain Management”. Beijing: China Machine Press, 2000.

[6] X.L. Meng, B.M. Cui, “The Model of Selecting Logistics Center Location With Fuzzy Judgment and Arithmetic Design”, Gansu Science, 2004, 20(1): 85-91.

[7] S.H. Ma, Q.Q. Meng, “Review of Supply Chain Logistics Capability Researches”, CIMS, 2005, 11(3): 301-307.

[8] P.F. Yan, “Research of the Planning and Design of Logistics Center Based on Railway System”, Logistics Scitechnol, 2006, 29(3): 76-78.

* Authors: Yu Jing, Ph.D, female, Associate Professor, interested in system programming and management on transportation and logistics.

Evaluative Research into E-government Sites Based on BP Neural Network

Ying Liang, Chun Hua Ju, Dong Sheng Liu
Computer and Information Engineering College, Zhejiang Gongshang University
149 jiaogong Rood Hangzhou, Zhejiang P.R.China, 310035
Email: ly.806@hotmail.com juchunhua@hotmail.com lds1118@hotmail.com

ABSTRACT

After analyzing various evaluation index systems of E-government sites, issues were studied such as too many indices, hardly-quantified and easily- correlated indices. Principal Component Analysis Method was adopted firstly to screen out the main one from a great many evaluation factors. The dimension was declined effectively on condition that evaluation information was reserved. The E-governmental website's quantitative evaluation is gained by the self-study function of the BP neural network. With the application and verification in E-government sites of Hangzhou 2005, it has proved that the model has a good universal property, robustness and reliability. It supplies a new valid way of evaluation in E-government sites.

Keywords: BP neural network, E-government Evaluation, Principal Component Analysis.

1. INTRODUCTION

With the development of Internet technology, E-government website becomes the main method by which the government can enhance the supervision abilities and provide more public services. In recent years, China's electronic government website has made great progress, but still exists many problems, such as the low exponent of handling affairs on-line, overemphasis on the government's central function instead of the users'. It becomes a major concern for many scholars and governments about how to evaluate the E-government website objectively and accurately, how to help it develop quickly and healthily and how to solve the existent problems. It is the comprehensive evaluation into the E-government website that can solve the above-mentioned problems.

Currently, many domestic and international literatures focus on the research of E-government website evaluation, such as the Accenture Consulting adopted the concept of "macro-mature degree" while evaluating the E-government development level [1]. In 2004, CCW Research carried on evaluation considering the following two aspects: services and the quantity of website [2]. In 2006, the CCID Consulting tried to evaluate websites from the service of contents [3].But from above-mentioned research, most traditional evaluation methods did not have unified viable standard and quantitative expression. Some evaluation indices are blurry[1] and Layer Analysis Method (AHP) is used simply. Complex non-linear relations between each index and weight can't be solved effectively [3]. Actually, the E-government website evaluation is a very complicated process, not only a great many evaluation factors are concerned, but also the relativity and non-linear relations of each evaluation factor can not be solved by the traditional evaluation method.

The BP neural network, with the characteristic of non-linear approach, the stronger ability of adaptation, studying, proceeding together and reasoning logically, can resolve the problem effectively. For this, a kind of evaluation model was put forward based on the BP neural network. Principal Component Analysis was adopted firstly on condition that evaluation information was reserved to filter out the main factor from a great deal of evaluation factors in order to decline the dimension effectively. E-government website's quantitative evaluation is gained by the self- study function of the BP neural network. With the application and verification in E-government sites of Hang Zhou 2005, it has proved that the model had a good universal property, robustness and reliability, which supplies a new valid way of evaluation in E-government sites.

2. BP NEURAL NETWORK MODEL BASED ON THE PRINCIPAL COMPONENT ANALYSIS

2.1 BP Neural Network Theory

Mr. Rumelhart, Hinton and Williams presented the back-propagation neural networks model [4], which is a highly complicated non- linear dynamics system. It is a multi-layer feed forward neural network and it is often called the BP neural network for weights are changed by back propagation algorithm. It solves the study problems of hidden layer in multi-layer networks. The BP neural network includes three layers generally, an input layer, one or more hidden layers, an output layer. Fig. 1. Shows the most basic three- layer BP neural network.

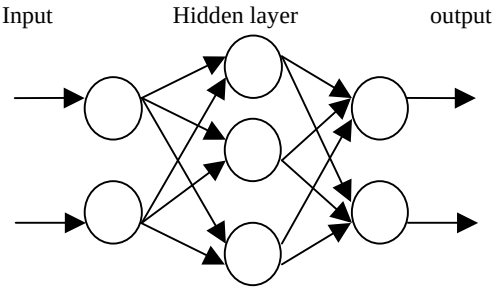


Fig. 1. BP Neural Network.

BP Neutral Network Based on the Principal Component Analysis

* This work was supported by the grand foundation of National Social Science Fund under Grant No 05BTJ019, foundation of National Doctor Fund under Grant NO 20050353003 and the project of Modern Business Trade Center of Zhejiang Gongshang University.

The Principal Component Analysis approach is a kind of multivariate analysis in order to reduce dimension and change the synthesized indices into a few principal factors not related to each other. It studies about how to explain the multi-variable covariance through several Principal Component (namely the linear combination of the original data). It eliminates correlativity between evaluation indices availably to avoid information overlapping between indices.

Step1: The Standardization of the original data bank (the $x_{ij})_{m \times k.}$ (m is the number of the sample; k is the number of the input factor).

Step2: Correlation matrix equation is built to get the characteristic value and characteristic vectors. The correlation coefficient can be gotten by the standardization value. K rank correlation matrix can be acquired by k values by which eigenvalue λ_i can be acquired. K eigenvectors include k component values $a_{ij.}$

Step3: The principal components are selected. The contribute rates of the first Principal Component to total variance are calculated. The contribute rates of variance is $\lambda_i \Big/ \sum_{j=1}^{k} \lambda_j$ .K Principal Components are made sequence by the contribution rates. The Principal Component with the largest contribution rates is called the first primary component, next is called the first second component and the rest can be deduced accordingly. The number of Principal Components are decided by it's contribute rates. The one whose contribute rates of variance is bigger than 85% can represent the majority information of k original values.

Step4: The Principal Component equation is built to compute the Principal Component value, each Principal Component equation is

$$c_i = \sum_{j=1}^{k} a_j x_j \qquad (1)$$

And a_j(j=1,2, …,k) is the component value corresponded to the eigenvalue λ_j and x_j($j=1,2, \ldots, k$) is the standardization value of each value . Each main component is computed to form new training, test and validation set.

The influencing factors were very complicated to the E-government evaluation. In order to avoid the loss of the important indices, the qualitative analysis method was often adopted to bring all influencing factors into the systems. There were 54 indices in the system of E-government evaluation of Hangzhou in 2004. If experts give marks one by one, it will be a waste of manpower and if all these indices are input into the neural network, it will increase the complications and reduce the performances of the network. So Principal Component Analysis was adopted to carry on the processing.

Software of matlab was used to call the princomp function of Principal Component.

[Pc, score, latent, tsqare] = princomp () (2)

The above formula was input to get the corresponding coefficients. Pc: Principal Component coefficient, score: Principal Component scores, latent: eigenvalues, tsqare: demographic.

2.2 Correlation Analysis in the Result of the Model

The quality of the training should be evaluated after carrying on the Principal Component Analysis and the training of the BP neural network. In the general circumstance, the disciplinal results (output) will not be the same as the target, but the disciplinal result should be within the permitted scope of the error given. If it is satisfied, training quantity is thought to be qualified, otherwise unqualified. At this time, linear regression method is adopted to make an intuitionistic evaluation of the neural network. The correlation coefficient scope of the valuation is settled to 0.7 and the actual output achieves ideal simulation result if it is bigger than 0.7.

3. APPLIED EXAMPLE

3.1 The Assurance of the Principal Component

According to the request of E-government evaluation in Hang Zhou City 2004, 54 indices can be fixed, as shown in Table 1. Parts of evaluation data in 2004 were adopted to carry on the evaluation.

11 Principal Components could represent the majority of information provided by the original data after PCA in the software of Matlab. The eigenvalue and the contribution rates of each Principal Component were shown in Table 2.

Because some indices data have certain relativity and the information reflected exist certain degree of overlapping, we make use of 11 principal component indices not related as the new set (the linear combination of the original 54 indices), which reserve the majority of original information. It reduces the input number of BP neural network to 11, and raises the BP neural network convergence speed and efficiencies. The corresponding relation of the Principal component indices and original indices is shown in Table 3.

Table 1. Indices

	Indices
1	Integrity of Organization Information
2	Effectiveness of Information
3	Science of Categorical Catalogue
4	Available of Important Information
5	Diversity of language
6	Integrity of Operational Unit
7	Operational Information Maturity
8	Detailed Degree of the Way of Contact
9	Single Logging of Citizen Mail
10	Column of Consultation and Appeal
11	Telephone and Mailbox for Appealing
12	Making Suggestions
13	Information in Common Use
14	Diversity of Search ways
15	Website's Navigation Function
16	Form Multiplicity of the Information
17	Information Service of Public Utilities
18	Characteristic Service
19	Networking Ratio of Handling Affairs
20	Responses Rate to Citizen's Appeals
21	Query of Affairs Information
22	Degree of One-stop Service
23	Multi-outlets Communications
24	Feedback Time of Information
25	Norm and Affinity of Service
26	Customized Information Service
27	Amity Information
28	Privacy and Safeties Hint
29	Service of Special Community
30	Convenience of the Service
31	On-line Supervision
32	Consultation before the Policy- making

33	Discussion of Special Subject	44	Data Instauration
34	Opinion Poll of the Public	45	Monitor of Network Safety
35	Reading Quantity	46	Norm of Domain Name
36	Ranks in Search Engine	47	Register Speed of the Website
37	Rank in the Same Class Website	48	Visual Effect of Web Page
38	Index of Page Rank	49	Function of Website Link
39	Universal Use of Website	50	Validity of Website Link
40	Help Information of the Website	51	Website Counter
41	Control of Customer Interview	52	Information of Maintenance Supervisor
42	Filtration of Sensitive Information	53	ICP Number
43	Data Safety	54	Public Evaluation

Table 2. The Contribution Rate of the Primary Component

Serial Number	Eigenvalue	Contribution Rate	Total Contribution Rate
The First Component	53.056	0.2182	0.2182
The Second Component	44.212	0.1818	0.4
The Third Component	33.87	0.1393	0.5393
The Fourth Component	27.466	0.1130	0.6523
The Fifth Component	20.651	0.0849	0.7392
The Sixth Component	18	0.0740	0.8112
The Seventh Component	16.996	0.0699	0.8811
The Eighth Component	13.475	0.0554	0.9365
The Ninth Component	7.096	0.0292	0.9657
The Tenth Component	4.729	0.0195	0.9852
The Eleventh Component	3.58	0.0147	0.9999

Table 3. Indices System Structure

	Principal Component	Characteristic Content
1	Basic Service of Information(X_1)	Integrity of Organization Information、Effectiveness of Information、Science of Categorical Catalogue、Available of Important Information、Diversity of language
2	Basic Processing of Business(X_2)	Integrity of Operational Unit、Operational Information Maturity
3	Basic Interaction (X_3)	Detailed Degree of the Way of Contact、Single Logging of Citizen Mail、Column of Consultation and Appeal、the telephone and mailbox for appealing、Making Suggestions、Information in Common Use
4	developmental service of information (X_4)	Diversity of Search ways、Website's Navigation Function、Form Multiplicity of the Information、Information Service of Public Utilities、Characteristic Service
5	Developmental Processing of Business (X_5)	Networking Ratio of Handling Affairs、Responses Rate to Citizen's Appeals、Query of Affairs Information、Degree of One-stop Service
6	Developmental Interaction (X_6)	Multi-outlets Communications、Feedback Time of Information、Norm and Affinity of Service
7	Perfect Service of Information(X_7)	Customized Information Service、Amity Information、Privacy and Safeties Hint
8	Perfect Processing of Business(X_8)	Service of Special Community、Convenience of the Service、On-line Supervision
9	Perfect Interaction (X_9)	Consultation before the Policy- making、Discussion of Special Subject、Opinion Poll of the Public
		Reading Quantity、Ranks in Search Engine、Rank in the Same Class Website、Index of Page Rank、

10	The third Square Technique Index (X_{10})	Universal Use of Website 、Help Information of the Website、Control of Customer Interview 、 Filtration of Sensitive Information、Data Safety、Data Instauration、Monitor of Network Safety、Norm of Domain Name、Register Speed of the Website、Visual Effect of Web Page、Function of Website Link 、 Validity of Website Link 、 Website Counter 、 Information of Maintenance Supervisor、ICP Number
11	Public questionnaire(X_{11})	Public Evaluation

3.2 Training of the BP Neural Network

BP neural network is not good with slow convergent speed and low efficiency, but improved BP neural networks based on self-adaptive *lr* are adopted accordingly to increase the stability and raise the speed and accuracy.

Training steps:

Step1: Corresponding data were gathered through the judgments made by experts [5].The experts in electronics realm were invited to give the appraising marks of 11 indices and the final marks of the 60 sub- website in Hangzhou City. And expert's knowledge and the experience were distributed in the coupling weighting value at this time.

Step2: E-government evaluation model based on BP neural network is carried out by the software of matlab65. 11 indices selected in step 1 were as the input and the final marks of the website as the output .The number of the hidden layer nodes could be decided firstly according to the empirical formula, here adopting "trying" method [6]. Small hidden layer nodes were chosen firstly and the training was stopped if the epochs were too much or the network had not achieved the convergence in the regulated times, then the hidden layer nodes would be increased and trained again. Four circumstances as 6, 8, 10, and 11 of the hidden node numbers were experimented and the hidden layer node number is 10 at last. The established model structure of BP network in this text was: 10 ×10 ×1, the system error was 10^{-4} and the initial value of the weight matrix wasthe random matrix of 11 ×10 and 10×1.

Step3: the training set was selected randomly from 20 government website in Hangzhou city 2005. Data of another 20 websites were selected as the test sample and the data in 10 websites were selected as the sample to predict. The Training set is shown in Table 4.

Step4: Training of the network. The training sample was input into the system and Matlab neural network tool box with the traingda study method was used to train the sample. The procedure was over after study for 6699 times, the error declines to 0.0001. Table 4 and 5 indicate the weighting value and the threshold between layers.

Table 4. Training Set of BP Neural Network

Serial Number of the Sample	Indices Mode	Scores
	$\mathbf{x_1}$ $\mathbf{x_2}$ $\mathbf{x_3}$ $\mathbf{x_4}$ $\mathbf{x_5}$ $\mathbf{x_6}$ $\mathbf{x_7}$ $\mathbf{x_8}$ $\mathbf{x_9}$ $\mathbf{x_{10}}$ $\mathbf{x_{11}}$	$\mathbf{y_1}$
1	[8.7 6.3 2.5 6.3 9.8 5.2 8.8 4.4 6.2 9.1 4.4]	6.389
2	[3.5 6.5 4.3 9.7 2.5 4.2 8.5 4.6 6.7 7.8 2.5]	4.6053
3	[8.9 9.3 2.5 4.6 7.8 6.9 2.3 1.6 3.3 3.6 8.9]	6.9056
4	[8.2 3.6 4.6 9.6 5.6 8.3 5.6 4.5 9.6 4.2 5.2]	5.7379
5	[6.6 3.6 5.6 4.6 8.6 9.6 9.8 7.8 5.6 6.6 9.6]	7.3656
6	[6.9 9.4 8.5 7.3 6.3 2.5 8.6 9.6 8.5 9.6 6.3]	7.3906
7	[3.6 5.6 8.9 8.6 4.5 5.6 9.6 7.5 3.6 2.3 1.8]	4.1338
8	[6.9 8.6 9.4 5.5 6.5 7.5 8.8 9.5 4.4 5.5 2.2]	5.6602
9	[8.8 6.6 4.5 9.6 8.5 6.2 3.2 4.5 5.6 3.2 3.6]	5.5306
10	[9.6 8.8 7.6 7.5 7.2 8.6 9.5 8.2 6.7 7.8 8.6]	8.3962
11	[6.9 8.5 7.5 8.3 8.2 8.1 9.6 9.4 9.5 8.2 7.2]	7.7482
12	[6.9 8.5 7.5 8.3 9.5 8.2 8.1 4.6 8.3 7.6 7.2]	7.6130
13	[8.5 8.2 8.3 7.6 7.2 7.1 4.6 8.3 7.6 7.2 6.3]	7.3150
14	[8.5 9.5 8.6 9.7 7.6 9.2 8.2 8.4 7.2 6.3 6.3]	7.6367
15	[8.5 9.5 8.8 8.6 9.7 8.2 7.3 5.6 5.8 6.3 6.3]	7.6149
16	[9.6 8.5 8.6 7.3 4.8 9.6 7.5 8.6 7.5 9.8 6.6]	7.9614
17	[5.6 6.9 9.8 9.6 9.2 4.5 7.6 8.8 8.6 8.5 7.5]	7.6403
18	[9.6 8.5 7.6 9.2 8.1 7.3 9.5 8.6 8.5 7.5 9.8]	8.8172
19	[9.6 8.5 8.6 7.3 4.8 9.6 7.5 8.6 7.5 9.8 6.6]	7.9614
20	[5.6 6.9 9.8 9.6 9.2 4.5 7.6 8.5 9.6 9.5 8.6]	8.1176

Table 5. The Weight Value and Threshold between Hidden Layer and Input layer After Training

	1	2	3	4	5	6
1	0.2659	-0.1496	0.0558	-0.2166	0.3178	-0.1915
2	0.2449	0.3836	0.1800	-0.1223	-0.0005	-0.4656
3	0.6700	0.2316	-0.1195	0.2770	-0.0813	-0.1300
4	-0.0610	0.5032	0.0597	0.4390	0.2224	0.3672
5	-0.0245	0.3713	0.2766	-0.2610	0.4283	0.2521
6	-0.1919	0.0057	-0.0039	0.0657	0.4574	0.6679
7	0.2139	0.0604	0.0831	0.0784	0.2992	0.3127
8	0.1346	0.2369	-0.0865	0.1380	-0.1867	-0.5608
9	-0.0141	-0.2832	0.2633	0.4998	-0.3086	-0.0303
10	-0.6392	0.5916	0.2202	-0.0912	0.2225	0.0238
	7	8	9	10	11	Threshold
1	-0.3270	0.3161	0.6357	-0.1139	-0.2918	-7.3450
2	0.1297	-0.4308	-0.5846	0.0802	0.1876	0.7318
3	-0.6252	0.2398	0.0224	0.1330	0.3468	-7.3126
4	0.0150	0.4314	-0.1849	-0.3770	-0.5725	-7.9261
5	0.1380	-0.0230	0.4709	0.4064	-0.1149	-10.3864
6	0.6080	-0.0692	-0.0373	0.3982	-0.0681	-0.6664
7	0.2834	0.1774	0.0228	0.3283	0.1821	-15.3241
8	0.3257	0.0586	-0.7058	0.1963	-0.3333	1.0722
9	0.0403	0.3557	-0.0818	0.3319	0.6193	-1.6866
10	0.3133	-0.1000	-0.0168	-0.2578	0.4380	-4.9638

Table 6. The Weight Value and Threshold between the Hidden Layer and Output Layer after Training

1	2	3	4	5	6
-0.3296	0.2340	1.3617	-0.7126	1.4932	1.7015
7	8	9	10	Threshold	
1.4606	-0.8864	1.3219	0.9329	1.2606	

3.3 Test and Validation of the Model

The model should be tested after the training. 11 parameters of the unstudied 20 websites were input into the network and after the imitation, homologous output could be gotten. In order to test the regression level between the actual output and the expected output, we carried on the linear regression method. The beeline is $A = 1.1204X - 1.0287$ and related coefficient is 0.9043, shown in the Fig. 2. The regression indicates that the actual output is effective.

At the same time, in order to test the model, 11 parameters of the last 10 websites were input, after imitation the corresponding output can be achieved. Comparison between actual data and the stimulant results was shown in Table 7.The regression analysis was made after BP training. Regression line is $A = 0.8604X - 1.0409$ and related coefficient is 0.9256, shown in Fig. 3. The fact indicates that: As long as the actual input mode is close to study sample's input mode in a certain train, it can produce the output near to the study sample's result, so-called the self associations function. It indicates that the actual output of the model is identical to the expected one and the model has a good extension.

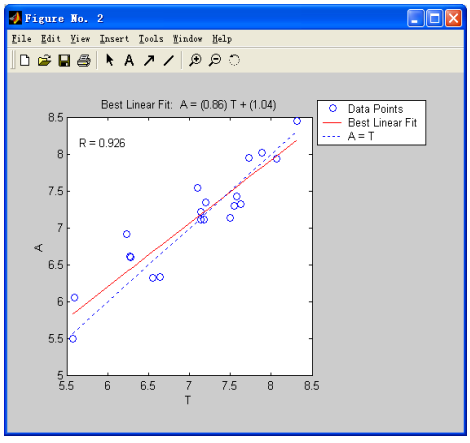


Fig. 2. Linear Regression of Actual Value and Validated Value

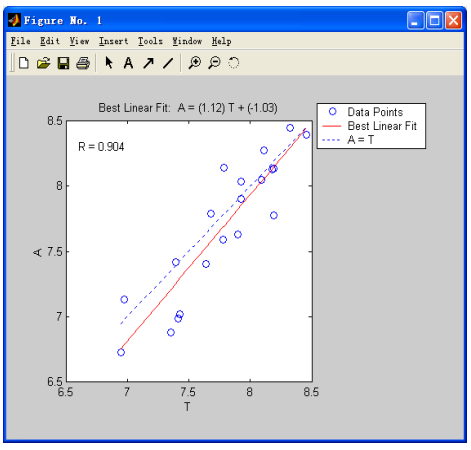


Fig. 3. Linear Regression of Actual Value and Validated Value

Table 7. The Test Sample (Comparison Table of the Original Data and the Forecast Data)

	41	42	43	44	45	46	47	48	49	50
T	7.9241	7.6385	8.1904	8.4547	7.7822	6.9453	7.4124	7.3508	7.6805	8.3231
P	7.8998	7.4066	7.7725	8.3924	7.5882	6.7236	6.9859	6.8808	7.7912	8.4480
	51	52	53	54	55	56	57	58	59	60
T	8.1103	8.1936	7.9283	7.8975	8.1747	7.4299	7.7827	7.3907	8.0904	6.9752
P	8.2756	8.1340	8.0341	7.6305	8.1278	7.0152	8.1394	7.4162	8.0471	7.1280

4. CONCLUSION

The evaluation model of the BP neural network based on Principal Component Analysis was used in the E-government website evaluation, which raises the performance of the network. It weakens the random and appraiser's subjectivity when fixing the index weight in evaluation, which guarantees the objectivity and scientific of the evaluation result. Once the network has been trained perfectly, it can be used to solve the same kind of problem, so the model based on BP neural network has an extensive foreground and application value.

5. REFERENCES

[1] Y.F.Yang, Q.H.Bai, "E-government Evaluation Iindex System", Computer Applications and Software, Aug.2004, Vol.21 (8): 61-65.

[2] CCW Research, "The Government Website Valuation Index Sign System Elucidation". http://www.ccw.com.cn/dc/show/2004pt_show/htm2004/20041102_16ssw.asp.

[3] "The Chinese Government Website Evaluation Results", E-government , January 2006,Vol.7(51): 23～30.

[4] C.H.Dong, Matlab Neural Network and the Application, Beijing: National Defense Industry Publishing Company，2005.

[5] G.Q.Huang, Y.D.Shi, "The Marketing Risk Valuation Model Based on the Expert Systems and the BP Neural Network", Journal of Information, January 2005, Vol.7 (1): 77-79.

[6] S.E.Yang, L.Huang, "The Finance Alert Model of Listed Company Based on BP Neural Network", Journal of Systems Science and Systems Engineering, , January 2005, Vol.7(2):12-17.

Collaborative Filtering Recommendation Algorithm Based on Sample Reduction

Linqi Gao
School of Management, Tianjin University & Tianjin Normal University
Tianjin 300387, China
Email: gaolqi@mail.tjnu.edu.cn
Congdong Li
School of Management, Tianjin University
Tianjin 300072, China
Email: licd2000@tju.edu.cn

ABSTRACT

Personalized Recommender System has become an important research field in Electronic Commerce, and the clustering of customers is the basis to produce the recommendation. The customers clustering in recommender system has some unique characters, such as the extreme sparsity of user rating data and huge sample space. Traditional Collaborative Filtering (CF) algorithm works poorly in this situation. To improve the quantity of recommending, sample reduction method is proposed in CF schema to lessen the sample space in both row and column aspects before carrying on the classification. It has the following features: (1) a Restraint Function is introduced into the basic CF model to solve the problem of sparsity of user rating data; (2) selective sampling and look-ahead framework are combined, based on the nearest neighbors algorithm, to reduce the number of samples while maintaining the quality of classification. At last, experiments are designed on the basis of MoveLens data set; recall and precision are applied as evaluating guidelines. Compared with general CF, the proposed algorithm has higher quality of recommendation.

Keywords: Nearest Neighbor Algorithm, Selective Sampling, Collaborative Filtering, Recommendation System, and Electronic Commerce.

1. INTRODUCTION

With the development of Electronic Commerce, Personalized Recommender System has become an important research item [1]. From estimating the requirement of customer, proves the suitable product and services for individual. Personalized Recommender System aims to solving the customer's overloading of information, and improving the performance of e-Commerce system. At present, many e-Commerce systems have employed different types of Personalized Recommender System, such as eBay, Amazon, DangDang book store, and so on.

The recommendation system not only help customer, but also enhance the customer's satisfaction to the business. In general, the recommender system has the following effects in Electronic Commerce environment.

1) Help user to search useful information

Perhaps, the customer having the explicit shopping goal can gain support from the search-engine to find their needs. But most customers that only having imprecise needs are very difficult with patience to search item by item in millions of web page, and decide whether it is interested or not. Through the appropriate recommendation, recommender system may turn a browser into a buyer.

2) Promote sales

Recommendation can be produced according to the product that customer has bought, such as the similar type of discount-CD. It may effectively promote the overlapping sale, and provide customer a better choice of commodity or service.

3) Personalized service

In fact, a successful recommender system could establish a personalized shop for each customer. It is adjusted and maintained according to customer's characteristics.

4) Enhance the loyalty of customer

If the recommender system can provide higher service though its recommendation, it will attract customer to use it more. So, the personalized service can build a reliable relation between the merchant and the customer.

Currently, more and more AI technology has been applied to improve the capability of recommender system. To gain exact and real-time recommendation, some recommending methods have been constructed based an different theory, such as Collaborative Filtering algorithm, Bayesian Network, Association Rule Mining, Clustering, Hurting Graph, Knowledge-based Recommender, etc [2,3,4,5].

In these recommending methods, Collaborative Filtering algorithm is a successful method, widely applied in many e-Commerce systems, such as recommending movies or news for user. CF algorithm evaluates the current customer's near neighbors according to the rating data. Through neighbors' rating data, the current customer's evaluation for a new product can be forecasted, then, the recommendation for current customer can be obtained.

But, in a realistic EC system, the magnitudes of customers and products are all huge. How to deal with the sparse and huge sample space is a key issue to make exact recommendation. Traditional CF algorithm works poorly in this situation.

2. BASIC COLLABORATIVE FILTERING MODEL

The CF recommendation acts according to other users' viewpoint to produce recommendation tabulates to the goal user. Its basic thought is based on a supposition: If user grade to some product quite similarly, then they grade to other product also similar [6]. CF recommendation model use

* The Research is supported by (1) the National Science Fund of China under Grant No. 70402009; (2) University Science and Technology Development Fund of Tianjin Education Committee under Grant No. 20030613; (3) Doctor Research Fund of Tianjin Normal University under Grant No. 52RL005.

statistics technology to search goal user's recent neighbors, and then forecast goal user's grading according to neighbors' grades thus has the recommendation.

The rating data of customer forms a $m \times n$ matrix $A(m,n)$. In the matrix, row *m* represents customer *m*, column *n* represents item *n*. The element $R_{i,j}$ in i^{th} row and j^{th} column represents the rating data of customer *i* about item *j*, like Table 1 show.

Table 1. Customer Rating Data Matrix

	Item $_1$	Item $_2$	……	Item $_n$
Customer $_1$	$R_{1,1}$	$R_{1,2}$		$R_{1,n}$
Customer $_2$	$R_{2,1}$	$R_{2,2}$		$R_{2,n}$
……				

To find user's neighbors, the similarity among users must be measured. Some measure algorithms often are applied, such as cosine similarity, correlation similarity and adjusted cosine similarity, to measure users' similarity degree.

The precise computation of goal user's neighbors is the successful key of CF model. It is necessary to attempt new algorithm foundation to increase the recommendation precision through better classification.

So, the similarity measurement among customers and customers' clustering are the foundation of Personalized Recommender model. Many clustering method have been applied in many application, but in e-Commerce system, the customers' clustering has its unique features.

Firstly, the magnitudes of customers and products are huge. The sample space for clustering is a typical sparse space. Traditional clustering methods work poor in this application field; make near neighbor set of a given customer is not exact. So, the quality of recommendation decreases dramatically.

Secondly, because the samples space is huge, labeling the samples is a costly process. At the same time, the processing speed decreases. It is not favors for providing service to on-line user.

In order to solve these two problems, it is needs to compress samples while maintaining the quality of classification. In the training process, providing the learning algorithm with some control over the inputs is an effective solution. Sometimes, it is called active learning.

3. DESIGN OF SR-CF ALGORITHM

The overall thought proposed in the article is to lessen the Customer Rating Data Matrix (Table 1) in both row and column aspects before carrying on the NN classification. The compression of column aspect is to resolve sacristy of user rating data. It is called the features compression. On the other hand, the compression of column aspect is to reduce number of samples based on selective sampling technology.

3.1 Restraint Function

Let *X* be an examples space, a set of objects described by a finite collection of features. Given a set of labeled examples $\{(x_1,\varpi_1),\cdots,(x_n,\varpi_n)\} \subseteq X \times \{0,1\}$ and a similarity function *d* over *X*, the nearest neighbor rule decides that $x \in X$ belongs to the category ω_j of its nearest neighbor x_j, where

$$j = \arg\min_{i=1,\ldots,n} d(x_i, x) \quad (1)$$

Euclidean metric is the similarity function frequently used. But in recommender system, the similarity measurement among customers is carried on the user's appraisal matrix to commodity. Because the number of customers and commodity are extremely huge, a customer only rates a few commodities. The matrix is extremely sparse. When measure their similarity, three kinds of situations will be encountered.

1) Regarding item *j*, both of the customer *i* and *k* have rated, then, the similar degree may be compared;

2) Regarding item *j*, the customer *i* has rated, but customer *k* has not, this time if calculating their similarity on item *j*, the similar degree will be reduced. Obviously, these rating data are different. The customer *k* has not rated item *j*, maybe he has not used this product, but not regard rating data is 0 on item *j*.

3) Regarding item *j*, both of customer *i* and *k* have not rated. In this situation, if calculating their similarity on item *j*, the similar degree will be increased. Because their rating data are not the identical value – 0, in majority situation.

If cannot restrain the latter two kind of situations, similarity evaluating will be out of true, thus the quality of recommendation will be dropped.

It is consistent with human's classifying process. The humanity does not need have the complete characteristic component to carry on the classification. Therefore, it is necessary to introduce a "restraint function". When comparing two customers, those components which having rating data be selected to calculate the similar degree.

The restraint function is a decision equation.

$$R(R_{i,j}) = \begin{cases} 1 & \text{if user } i \text{ rated item } j \\ 0 & \text{if user } i \text{ not rated item } j \end{cases} \quad (2)$$

When calculating the distance between of customer *I* and *k*, restraint function *R* is applied, keeps the computation in those item which customer rated. See equation (3).

$$d(i,k) = \sqrt{\sum_{j=1}^{n} R(R_{i,j}) R(R_{k,j}) (R_{i,j} - R_{k,j})^2} \quad (3)$$

3.2 Selective Sampling

The sampling technique is one kind of information processing widely applied. The selective sampling approach is more common in practice [7]. The learner selects an unlabeled example from the given set and asks the teacher to label it.

Consider the following selective sampling paradigm. For examples space *X*, let $f: X \to \{0,1\}$ be a teacher that labels instances as 0 or 1. A learning algorithm takes a set of labeled examples, $\{(x_1, f(x_1)),\cdots,(x_n, f(x_n))\}$, and returns a hypothesis $h: X \to \{0,1\}$. *X* is an *n*-dimension Euclid space constituted by users, $X = R^n$.

Let *X* be an unlabeled training set, a set of objects drawn randomly from *X* according to distribution *p*. Let $D = \{(x_1, f(x_1)) \mid x_i \in X, i = 1,\cdots,n\}$ be the training data — a set of labeled examples from *X*. A selective sampling algorithm *S*, receives *X* and *D* as input and returns an unlabeled element of *X*.

The process of learning with selective sampling can be described as an iterative procedure. The selective sampling procedure is called to obtain an unlabeled example and the teacher is called to label that example, during every loop. The labeled example is added to the set of currently available labeled examples and the updated set is given to the learning procedure, then the classifier is updated.

Teacher process applies Equation (1) and (3), label that unlabeled example according to the label of its nearest labeled

neighbor.

In sampling process, in order to ensure sample selected valid, two features should be considered.

1) Uncertainty of classification. It is means the probability of miss-classification for a sample. Generally, if a sample has two neighbors that have diverse label and same distance, its uncertainty of classification is higher.

2) Affection of sampling. The classifying result about selected sample has affection about other un-labeled sample. If reducing the affection through cautious sampling, so the precision of classification will be enhances.

So, the main consideration is not only the uncertainty of the particular unlabeled sample, but rather the effect that the labeling of this point may have on its neighborhood. The look-ahead selective sampling algorithm described in the next section takes these considerations into account.

3.3 Look-ahead Selective Sampling Algorithm

The basic thought of look-ahead selective sampling algorithm is estimating the expected utility of un-labeled sample before selecting samples [8]. Then, sample that has the highest expected utility is selected to classify.

For a customer *x*, which has the real classification *f* and hypothesis classification *h*, Let $I_{h,f}$ be a binary function, where $I_{h,f}(x)=1$ if and only if $h(x)=f(x)$, otherwise, $I_{h,f}(x)=0$. Let $\alpha_h(f)$ is the accuracy of hypothesis *h* relative to *f*.

$$\alpha_h(f)=E_x[I_{h,f}(x)]=\int_X I_{h,f}(x)p(x)dx \qquad (4)$$

$P(x)$ is the probability density function specifying the instance distribution over *X*. Thus $\alpha_h(f)$ is the probability that $h(x)$ is equal to $f(x)$ on sample *x*.

$$\alpha_h(f)=1-err(h)=P(h(x)=f) \qquad (5)$$

The *h* is the classification of customer *x* produced by a learning algorithm *L*. Let $U_L(D)$ denote the expected accuracy of hypothesis $h=L(D)$. This expected accuracy is taken relative to the posterior distribution on given labeled sequence *D*, which is defined by Bayesian rule from the prior distribution.

$$U_L(D)=E_f[\alpha_h(f)]=E_f[E_x[I_{h,f}]]=E_x[E_f[I_{h,f}]] \qquad (6)$$

The $E_f[I_{h,f}]$ is the probability $P(f(x)=0\mid D)$ that algorithm *L* consistent with *D* is equal to $h=L(D)$ at specific sample *x*. So,

$$U_L(D)=\int_X P(f(x)=h(x)\mid D)p(x)dx \qquad (7)$$

Suppose the class probabilities $P(f(x)=0\mid D)$ and $P(f(x)=1\mid D)$ are given. With a finite sequence *D* selected from *X*, Equation (7) can be computed by

$$U_L(D)\approx\frac{1}{|X|}\sum_{x\in X}P(f(x)=h(x)\mid D) \qquad (8)$$

Equation (8) is chosen as utility measurement translating the problem of evaluating the expected classifier accuracy to the problem of estimating the class probabilities.

From Equation (8), the sample category probability bust be calculated firstly to gain the utility $U_L(D)$, because the classifying algorithm is unknown. Let θ_1, θ_2 be the classifications of sample x_1, x_2. Their values are 0 or 1. Because x_1 and x_2 have same distribution, so their expected values are equality.

$$E[\theta_1]=E[\theta_2]=\overline{\theta} \qquad (9)$$

The covariance between θ_1 and θ_2 is determined only by the distance between x_1 and x_2.

$$Cov[\theta_1,\theta_2]=E[(\theta_1-\overline{\theta})(\theta_2-\overline{\theta})]=\gamma(d(x_1,x_2)) \qquad (10)$$

γ is an approximation to describe covariance function. Let γ be an exponent function [9], *d* is the distance between two points and σ is the max distance in sample space *X*.

$$\gamma(d)=0.25e^{-d/\sigma} \qquad (11)$$

A linear function is used to estimate the conditional class probability of un-labeled sample x_0 [10].

$$P(\theta_0\mid\theta_1,...,\theta_n)=\alpha_0+\sum_{i=1}^{n}\alpha_i\theta_i \qquad (12)$$

θB_{1B}… θB_{nB} is the known classification associated with samples x_1… x_n. α_i, i=1, …, n are coefficients. The optimal linear estimator minimizes the Mean Square Error (MSE). So,

$$P(\theta_0\mid\theta_1,...,\theta_n)=E(\theta_0)+\partial\cdot(\theta-E(\theta)) \qquad (13)$$

∂ is an *n*-dimensional coefficients vector specified by the covariance values, and

$$\begin{aligned}&\partial=R^{-1}\cdot r\\&R_{ij}=E[(\theta_i-E(\theta))(\theta_j-E(\theta))]\\&r_i=E[(\theta_0-E(\theta))(\theta_i-E(\theta))]\end{aligned} \qquad (14)$$

R is a $n\times n$ matrix, and ∂, γ is *n*-dimensional vectors. The values of *R* and γ **is** specified by the random field model.

$$\begin{aligned}&R_{ij}=Cov[\theta_i,\theta_j]=\gamma(d(x_i,x_j))\\&r_i=Cov[\theta_0,\theta_i]=\gamma(d(x_0,x_i))\end{aligned} \qquad (15)$$

Because selective sampling algorithm estimates class probability of customer based on nearest neighbor, it is feasible to suppose: (1) the priori class probabilities are equal; (2) only two nearest neighbors are considered, other points are negligible.

$$\begin{aligned}&P(\theta_0=1\mid\theta_1=0,\theta_2=0)=\frac{1}{2}+\frac{-\gamma(d_{01})-\gamma(d_{02})}{\frac{1}{2}+2\gamma(d_{12})}\\&P(\theta_0=1\mid\theta_1=0,\theta_2=1)=\frac{1}{2}+\frac{-\gamma(d_{01})+\gamma(d_{02})}{\frac{1}{2}-2\gamma(d_{12})}\\&P(\theta_0=1\mid\theta_1=1,\theta_2=0)=\frac{1}{2}+\frac{\gamma(d_{01})-\gamma(d_{02})}{\frac{1}{2}-2\gamma(d_{12})}\\&P(\theta_0=1\mid\theta_1=1,\theta_2=1)=\frac{1}{2}+\frac{\gamma(d_{01})+\gamma(d_{02})}{\frac{1}{2}+2\gamma(d_{12})}\end{aligned} \qquad (16)$$

So, through calculating the expected utility of sample and selecting the sample which having higher utility, the number of samples is reduced.

4. EXPERIMENTS AND ITS ANALYSIS

The goal of customers clustering is improving recommendation. So, the validating basis of algorithm is the quality of recommendation, is not the cluster itself. This is the characteristic of experiment in the article.

Using data set provided in the MovieLens (http://movielens.umn.edu/), experiment is designed. From rating data database, randomly select 5,000 terms as the experimental data set, contains 352 users and 150 movies.

Uses PRECISION and RECALL as evaluation criteria of experiment those frequently applied in the information retrieval domain.

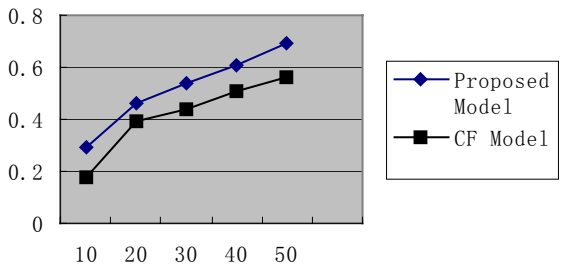
RECALL is percent of movies that user likes which correctly recommends.

$$recall = \frac{|test \cap top - N|}{|test|} \quad (17)$$

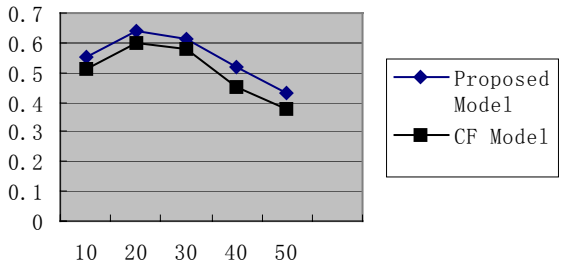
PRECISON is percent of recommended movies liked by user.

$$precision = \frac{|test \cap top - N|}{N} \quad (18)$$

Comparing the model proposed in this article with the general CF model, experimental results show in Figure 1 and Figure 2.



X-coordinate expresses the number of recommendation
Y-coordinate expresses RECALL
Fig. 1. RECALL comparison



X-coordinate expresses the number of recommendation
Y-coordinate expresses PRECISON
Fig. 2. PRECISON comparison

The experimental results show the proposed method has higher RECALL and PRECISION than traditional CF method, thus can enhance the performance of recommendation.

5. CONCLUSIONS

The clustering of customer is the basis of recommender algorithm, it is necessary through analyzing feature of application to design clustering algorithm. A Collaborative Filtering Recommendation Algorithm is proposed based on look-ahead selective sampling technology. It has the following features: (1) a Restraint Function is introduced in the basic CF model to solve the problem of sparsity of user rating data; (2) at basis of the nearest neighbors algorithm, selective sampling and look-ahead method are combined to recommender system to reduce the number of samples while maintaining the quality of classification.

In the future, it is necessary to design more experiment and roundly examine the proposed algorithm, including its quality and time expense.

REFERENCES

[1] R. Bergmann, P. Cunningham, "Acquiring Customer's Requirement in Electronic Commerce," Artificial Intelligence Review, 18:163-169,2002.

[2] Y.H. Cho, J.K. Kim, "Application of Web usage mining and product taxonomy to collaborative recommendations in e-commerce," Expert Systems With Application, 26: 2333-246, 2004.

[3] A. Mild, T. Reutterer, "An improved collaborative filtering approach for predicting cross-category purchase based on binary market basket data," Journal of Retailing and Consumer Services, 2003,10: 123-133.

[4] C. Kwok-Wai, "Learning user similarity and rating style for collaborative recommendation," Information Retrieval, 2004,7: 395-410.

[5] L.Q. Gao, L.Z. Li, "Production recommender method based on customer's behavior," Computer Engineering and Application, 2005,3:188-190.

[6] B. Sarwar, G. Karypis, et al, "Item-Based collaborative filtering recommendation algorithms," Proc. of the 10th International World Wide Web Conference, 2001,285-295.

[7] Y. Freund, H. S. Seung, et al, "Selective sampling using the query by committee algorithm," Machine Learning, 1997,28(2/3), 133–168

[8] S.J. Russell, P. Norvig, Artificial Intelligence, A Modern Approach. Prentice Hall, 1995.

[9] Y. Eldar, M. Lindenbaum, et al, "The farthest point strategy for progressive image sampling," IEEE Transactions on Image Processing, 6:9,1997, 1305–1315.

[10] A. Papoulis, Probability, Random Variables, and Stochastic Processes, 3rd ed. McGraw-Hill, 1991.

A New Two-Stage Heuristic For Vehicle Routing Problem with Time Windows

Weimin Liu
School of Mathematical Science, South China University of Technology, Guangzhou, 510641, China
Email: wm_Liu@hotmail.com

And

Zhifeng Hao
School of Mathematical Science, South China University of Technology, Guangzhou, 510641, China
Nation Mobile Communications Research Laboratory, Southeast University, Nanjing,210096,China
Email: mazfhao@scut.edu.cn

And

Ming Chen
Nation Mobile Communications Research Laboratory, Southeast University, Nanjing,210096,China

ABSTRACT

The vehicle routing problem with time windows is a famous combinatorial optimization problem which has been widely applied in logistics. The problem is to serve all customers at minimum cost without violating their time window constraints and the vehicle capacity constraint. The *m*-VRPTW is a variant of the VRPTW in which only a limited number of vehicles are available, whose feasible solution may include some unserved customers due to the limited vehicles. The primary objective of *m*-VRPTW is to maximize the number of customers served, while the second is to minimize the total travel distance. In this paper, we first time introduce adaptive k-means to this problem, with an Ejection Pool to hold temporarily unserved customers and an Adaptive Memory to keep good solutions. Then, we minimize the total distance using an improved AMP algorithm. The experimental results showed consistently good performance of our algorithm when compared with other methods.

Keywords: m-VRPTW, k-means, ejection pool, adaptive memory.

1. INTRODUCTION

Vehicle routing problem with time windows is a famous NP-hard combinatorial optimization problem which has been applied widely into practice. VRPTW can be described as follow: given a directed graph G=(V,E),where V denotes *N* customers who are represented by the numbers *1,….,N* and the depot represented by the number 0,while the set of edges E denotes routes of vehicles. Each customer has a specific time window $[e_i,l_i]$, $i=0,1,2,...N$. All routes must leave the depot after time e_0 and return before time l_0. A vehicle has to wait if it arrives at the customer *i* before e_i , while the customer can't be served if the vehicle arrive after l_i. A set of *K* vehicles with identical capacity *Q* and each customer with demand $q_i \geq 0$ are given. With all customers served by only once, the total service demands of the customers in a route are not allowed to exceed *Q*. The travel cost between vertex *i* and *j* is denoted by c_{ij}. Travel costs satisfy the triangular inequality

$$c_{ij} + c_{jk} \geq c_{ik} \qquad \text{Eq. (1)}$$

Since this problem was firstly proposed by Dantzing and Ramser[1] in 1959,it has received dramatic attention in the recent decades and been proved to be a NP-hard problem by Savelsbergh[2]. In the recent decades, a tremendous amount of work in the field has been published. Some exact approaches have been reported for VRPTW of small size, such as Branch and Bound(K-tress approaches)[3] , Lagrangian relaxation methods[4][5][6], and so on. Exact approaches have to cost much time to find optimum solutions. Current VRPTW optimization models are useful for a variety of practical applications. However, some practical issues have not yet been addressed. Many heuristics and meta-heuristics methods have been proposed to gain better results, for instance, Ant Algorithms[7], Constraint Programming Algorithm[8], Two-stage method[9], MILP[10]and so on. All these methods aim to explore the minimum cost with unlimited number of vehicles.

Recently, a new variant of the VRPTW, the *m*-VRPTW was proposed by Lau et al.[11] to address real world constraints such as a limited fleet. Given *m* and a VRPTW instance, the first object of *m*-VRPTW is to maximize the number of served customers with *m* vehicles as many as possible, and the second object of this problem is to minimize total distances. Since *m*-VRPTW has only limited number of vehicles, we can distinguish it with VRPTW by whether it allows unserved customers or not. Lau[11] proposed a *tabu* search approach characterized by a holding list and a mechanism to force dense packing within a route. Andrew Lim[12] proposed an effective two-stage method to solve this problem.

In this paper, we present a new two-stage heuristic for the *m*-VRPTW. We first time propose adaptive k-means and FSFS(first start-time first service) to create initial solutions, because it's reasonable to consider that solution will be easily found if customers 'sit' in the right cluster. In the step of maximizing served customers we use two data structures called Ejection Pool(EP) to hold temporarily unserved

customers and Adaptive Memory to keep the best solution found so far. Unserved customers in the EP should be re-inserted back into routes iteratively, and some existing served customers may be kicked out and added into the EP. We update the adaptive memory during the processing of space search, which will save much cost for the second step of minimizing travel distances. The experimental results demonstrate that our approach is an effective way in the maximization of customers served. In the second step, we resort to IAMP which is based on AMP[13] to minimizing the total distance. So far, the known best results of C1 class of Solomon's benchmarks[14] on the minimization of total costs was published by AMP algorithm. Our algorithm shows good performance when compared with other methods.

2. ALGORITHM FOR *M*-VRPTW

Our algorithm consists of three main components, the procedure to generate initial good solutions, the procedure to maximize served customers with *m* vehicles and the procedure to minimize total distances. The first step is based on simple k-means and LNS which was firstly proposed by Paul Shaw[8].LNS has been proved to be an effective algorithm for VRPTW, especially in the minimization of travel costs, see [9].However, in this paper, the idea of LNS is significantly applied in maximizing served customers rather than minimizing travel costs. In order to improve the efficiency, we propose a new operator called *Transfer* to improve our initial solutions. The algorithm firstly focuses on maximizing served customers using the EP and adaptive memory data structures. Then we aim to reduce travel distances at the minimum cost by IAMP algorithm.

Initial solutions

In this section, we focus on the generation of our initial solutions of *m*-VRPTW. At first, we have known that the distribution of customers is greatly impacted by their time window constraints, which is the main difficulty of this problem. Due to the low complexity of k-means, it can cluster fast. We use simple k-means to cluster customers at first according to their locations. Since simple k-means depended greatly upon the selection of initial centers, we randomly generate the initial centers to attain various clusters, which can be useful to keep the diversity of solutions. Even though some clusters may still contain some obvious infeasible customers. Thus, it will work more effective if we can do some improvement during the clustering process. Here, we apply a constraint to clusters as follow:

$$. \ distance(c_i, center(\sigma)) \le \alpha \bar{d}, \qquad \text{Eq. (2)}$$

where α is a parameter , c_i is a customer in route σ and $\bar{d}$ is the average distance of route σ.

Initially, except the depot, we use k-means upon customers to generate *m*-clusters. Then we introduce our FSFS idea in this way. We firstly sort customers in *m*-clusters respectively according to their ready-time with not descending order and then try to arrange start-time for each customer. Here, the start-time is the probable time to serve a customer. Suppose c_i is the pre-successor of c_j , we define the start-time as follow:

$$Start\text{-}Time(c_j) \begin{cases} \mathbf{max}\{Start\text{-}time(c_i)+Service\text{-}time(c_i)+ c_{ij}, e_j\} \\ \quad , Start\text{-}time(c_i) + Service\text{-}time(c_i) + c_{ij} \le l_j \\ \infty \ , otherwise \end{cases} \qquad \text{Eq. (3)}$$

The algorithm generates some partial or complete solutions in the light of this method.

However, some customers whose time windows are wide should have been located behind, may be arranged before other customers, which may lead to infeasibility. As we know, customers with short time windows are easily violated, therefore it is reasonable for us to suppose that locations of such customers are more stable than those with wide. And their infeasibility is always due to wrong arrangement. That is, there are too many customers before them. In order to deal with such matters, in this paper, we propose a new operator which we called *Transfer*. First, we presume *i* is the first infeasible customer and *j* is the last customer in the route σ respectively. Then we define the *Transfer* operator as follow:

Transfer: For customer *i*, place *i* after *j* , remove arcs (i^- ,*i*) and (*i*, i^+) ,and then add arc (i^- , i^+), where i^- is the pre-successor of *i* and i^+ is the successor of *i*.

We applied *Transfer* operator after the process of arrangement of start-time. In this way, we can separate customers with short time windows and those with wide. What's more, customers with short time windows have located their 'stable' positions. As a result, we can generate some good initial solutions. Moreover, we discovered the new operator performed better if a low bound of vehicles to server all customers was given. It's because there may exist only customers with tight time windows in the routes due to the insufficient vehicles. We have discovered that the algorithm would be more effective by the use of the new operator.

Maximizing the served customers

Given *m* vehicles and an initial solution generated by procedure above, we resort to EP and adaptive memory data structures to maximize served customers. All other routes are removed from the solution and those customers served in these routes are added into the EP. And a best solution so far is kept in the adaptive memory.

Since it was impossible to serve all customers with insufficient vehicles, we'd better analyze some latent conditions. Here, we found the bound of the number of customers served, which is significant to lead us to explore the best solution. We give the approximation of the number of customers served as follow:

$$Number\_Served \le \max\{m\frac{l_0}{\theta T}, \frac{mQ}{d}\}, \quad 0<\theta\le 1 \qquad \text{Eq. (4)}$$

where $T \in \{\sum_{\min} ServiceTime(c_i), \sum_{\max} ServiceTime(c_i)\}$

In each iteration, the procedure selects an unserved customer in the EP with the maximum relativeness defined as Equation (5) and updates the adaptive memory as well. The relativeness is not a new idea, it has been used in LNS which is a good method to deal with the VRPTW firstly proposed by Paul Shaw[8]. The procedure insert a customer in the EP into an existing route according to the relativeness, the value of which is a heuristic measurement of the relationship between route σ and customer *i*. For customers in Ejection Pool, C'= $\{c_1,c_2,\ldots c_s\}$,we define the relativeness between customer c_i and route σ_j as follows:

$$relativeness_{i,\sigma_j} = \frac{1}{c_{ij} + v_{ij}} \qquad \text{Eq. (5)}$$

Where $c_{ij} = \overline{d} / \max d_i$, $\overline{d}$ is the average distance between customer c_i and customers in route σ_j, v_{ij} =1 if customer c_i and customer c_j have similar time windows.

Although our relativeness defined is similar with the one proposed by Pawl Shaw in formulation, each variable is different. Generally, a solution of *m*-VRPTW contains some unserved customers because only a limited of vehicles is available. In order to ensure the efficiency, we sort the relativeness for each customer in Ejection Pool with ascending order. By selecting a route with the maximum relativeness, the algorithm intends to insert a customer into the route. Finally, some customers in the EP with very difficult to insert into routes will resort to the use of some classical and newly neighborhood operators. The pseudo-code of our algorithm is given by the following procedure:

Procedure Maximization
  Use improved k-means to generate clusters;
  Try to arrange start-time for each customer by using FSFS and *Transfer* operator;
  Add infeasible customers into the EP;
  Update the adaptive memory;
  while ((EP is not empty) and (no improvement within several run-times)) **do**
    Select a customer *i* from the EP;
    Determine the right route and position to insert;
    Remove *i* from the EP if it is feasible;
    Add infeasible customers into the EP;
    Apply classical operators *2-Exchange, Or-Exchange, Relocation ,Crossover and Exchange* to routes;
    Update the adaptive memory;
  end while
end

After the candidate customer, denoted as *u*, is selected, the algorithm has to determine the best position of the target route to insert the customer *i*. Here, we use Solomon's I1 heuristic insertion in [14] because it had been proved to be the best insertion. The insertion cost, $c(i,u,j)$ of inserting u between i and j is defined as Equation (6). d_{iu}, d_{uj} and d_{ij} are distances between i and u, u and j, and i and j, respectively. Since inserting u into the target route may lead to violating other customers' time window constraints, it is reasonable to consider the service delay as well. For this purpose, $c_{12}(i,u,j)$ shows the idea as follow:

$$c(i, u, j) = \alpha_1 c_{11}(i,u,j) + \alpha_2 c_{12}(i,u,j), \quad \text{Eq. (6)}$$
where $\alpha_1 + \alpha_2 = 1$, $\alpha_1 \geqslant 0$, $\alpha_2 \geqslant 0$,

$$c_{11}(i,u,j) = d_{iu} + d_{uj} - d_{ij}, \quad \text{Eq. (7)}$$

$$c_{12}(i,u,j) = max(t_u - l_u, 0) + max(t_u + s_u + d_{uj} - l_j, 0). \quad \text{Eq. (8)}$$

In Equation (8), t_u is the earliest time to serve u after u is inserted between i and j by temporarily relaxing the late time window constraint of u, s_u is the given service time of u and l_j is the latest arrive time of j such that the time window constraints of j and those customers after j will not be violated. The final insertion position is chosen to be the *p*-th element in the candidate list, and $p = \lfloor random(0,1)^e \times s \rfloor$, where e>1 and s is the size of the candidate list.

After several insertions, most of customers in the EP have been inserted into target routes accordingly. However, there may be still few customers difficult to insert into the routes. Then, we use some classical and newly operators to improve the procedure, for instance, in [15,16,17]: *Exchange, Relocation, Crossover, 2-Exchange and Or-Exchange, ICROSS.* We show details about these operators. We denote i^+ as the successor of i ,while i^- the pre-successor.

Exchange: Exchange the position of customers *i* and *j*.

2-Exchange: For two customer *i* and *j* on the same route, *i* is visited before *j*, remove arcs (i,i^+), (j,j^+), add arcs(i,j) and reverse the orientation of the arcs between i^+ and *j*.

Or-Exchange: Remove a sequence of 1, 2 or 3 customers from a route and reinsert the sequence elsewhere on the same or a different route.

Relocation: For customers *i* and *j*, place *i* after *j*.

Crossover: Exchange the successors of customer *i* and *j*.

ICROSS: It is an extension of the *Cross* operator, both original and the reversed order for the customers are considered.

These traditional operators have been proved to be effective to VRPTW in many papers. By using these neighborhood operators, good solutions can be obtained without much difficulty. Then, we must update the adaptive memory to keep the best current solutions for the next procedure. The algorithm greatly improves the solution quality in term of the number of customers served when compared with other methods, which is demonstrated by the experimental results in the following section.

Minimizing distance cost

In the first stage, the procedure of maximization of served customers has obtained some good solutions, several of whom have been close to the optimum. Thus, we can continue to explore the solution space without much cost. In this stage, we improve the AMP algorithm which was firstly proposed by Rochat and Taillard[5]. Its main strategy is to keep good solutions throughout the searching process. We have shown the adaptive memory data structure to demonstrate this idea. Since the algorithm can obtain good solutions in the first stage, we don't have to deal with routes with sole customer separately. Because such routes barely generate after the first stage. Then, we insert those routes which contain less than two or three customers into other routes. Moreover, we must also take into consideration that there may be some identical solutions obtained in the adaptive memory. Thus we must do some further search in the solution space. Here, we use operators mentioned in the Sec2.2 once again. The pseudo code of IAMP is as follow:

Procedure IAMP( )
  while(exists routes contain less than 3 customers)
    o:= RandomSelect(operator);
    Re-insert routes with less than 3 customers into other routes;
  end while
  Divide all the solution, T: = {all routes};
  Label each route with the value of the solution to which it belongs;
  Sort T by increasing values of labels;
  while (exists descending solutions) **do**
    T':= T , S:= $\varnothing$;
    Choose t∈T', probabilistically, based on its current relative evaluation;

Table 1. Number of customers served for C1

	Vehicles	*10*	*9*	*8*	*7*	*6*	*5*	*4*
C101	LST	100	92	84	75	66	57	47
	LZ	100	92	84	75	66	57	47
	Our Alg.	100	92	84	75	66	57	47
C102	LST	100	99	92	84	72	60	49
	LZ	100	99	93	84	72	60	49
	Our Alg.	100	99	93	84	72	60	49
C103	LST	100	99	93	84	72	60	48
	LZ	100	99	94	84	73	60	49
	Our Alg.	100	99	94	84	**75**	60	48
C104	LST	100	99	94	84	72	60	48
	LZ	100	99	95	84	72	61	48
	Our Alg.	100	99	95	84	72	61	**49**
C105	LST	100	92	84	75	66	57	47
	LZ	100	93	84	75	66	57	47
	Our Alg.	100	93	**85**	75	66	**58**	47
C106	LST	100	93	86	77	68	58	47
	LZ	100	94	86	76	67	58	47
	Our Alg.	100	93	86	77	68	58	47
C107	LST	100	93	85	76	66	57	47
	LZ	100	94	86	77	67	57	47
	Our Alg.	100	94	86	77	67	57	**48**
C108	LST	100	96	87	78	69	59	48
	LZ	100	98	89	80	70	59	48
	Our Alg.	100	98	87	80	70	59	48
C109	LST	100	99	92	82	72	60	48
	LZ	100	99	93	83	72	61	49
	Our Alg.	100	99	93	83	72	61	48
Total	LST = 4854 , LT= 4875 , Our Alg.=4878							

S:=S ∪ { t };
Remove from T' all the tours including one or few customers belonging to t ;
if(S is a partial solution) **then** Complete S ;
Improve S;
Label the improved solutions and Sort them, and Update T;

end while
Update the adaptive memory;
end

In this algorithm, we use different operators randomly but not sequentially since widely use of operators may be probable to reduce the quality of solutions, and the complexity will increase rapidly with sequentially use of operators. To strengthen the efficiency, the procedure mixes some operators if two operators work together can improve the solution better than the use of sole operator. Our experimental results have shown its efficiency. More detail about AMP algorithm can be found in [13].

In our distance minimization procedure, although other heuristic such as *tabu search* procedure is able to explore the solutions, it does not have the ability to go through the infeasible solution space.

3. EXPERIMENTAL RESULTS

This section describes experimental results on our algorithm. The entire code of the algorithm is more than 3000 lines. All results are given on a P4 2.4G, 512MB RAM using C++ language on VC6.0. All numbers used were double precision floating points. In order to compare our algorithm with others', we have experiments on our algorithm with the Solomon's 56 benchmarks available at *http://www.cba.neu.edu/~solomon/problems.html*. The size of the adaptive memory is set to be 20, and the run-times of adaptive k-means is no more than 100 times. The control parameters were set as α=1.3, α_1= 0.5, *s*=13, *e*=50.

For the *m*-VRPTW, our algorithm is compared with earlier methods including LST[11] and LZ[12] in Table 1. The C class of the problem sets of the Solomon benchmarks has especially shown the efficiency of our newly proposed operator. The best running time to find the best solution is only within 15sec, such as C101, C105, C107, C109, R201, R205, and so on. The results demonstrate the effectiveness of our algorithm. The new best solutions are highlighted in table. Meanwhile, other Solomon instances were tested as well and Table 2 shows the results.

Table 2. Number of customers served for R1 and RC1

m	19	18	17	16	15									
R101	100	99	**99**	95	93									
R102	-		100	99	98									
m	14	13	12	11	10	9	8	7	6	5	4	3	2	1
R101	89	86	**82**	**77**	**71**	65	59	53	46	39	32	25	17	9
R102	**96**	**94**	**93**	**89**	**85**	**82**	**75**	**68**	**61**	53	45	36	25	13
R103	-	100	**99**	**98**	**95**	**91**	**85**	**77**	**69**	**59**	**49**	**39**	**27**	14
R104	-	-	-	-	-	100	**93**	**84**	**74**	**64**	52	41	28	14
R105	100	98	96	91	**87**	**81**	**75**	**67**	59	50	41	32	22	12
R106	-	-	100	98	**95**	**90**	**84**	**75**	**66**	58	49	37	26	14
R107	-	-	-	-	100	**96**	**90**	81	71	62	51	**40**	27	14
R108	-	-	-	-	-	100	**94**	**86**	**76**	65	53	41	**28**	**15**
R109	-	-	-	100	97	**90**	**87**	**75**	66	57	47	**37**	25	13
R110	-	-	-	-	100	**95**	**88**	**78**	69	60	49	37	26	14
R111	-	-	-	-	100	96	89	80	71	61	51	39	27	14
R112	-	-	-	-	-	100	92	83	73	62	51	39	27	14
RC101	100	97	93	**90**	84	**78**	71	63	55	48	40	31	21	11
RC102	-	-	100	**97**	**93**	81	**79**	71	62	54	45	34	23	12
RC103	-	-	-	100	98	**93**	**86**	**78**	**68**	**58**	47	36	24	11
RC104	-	-	-	-	100	96	90	79	71	**62**	50	38	26	13
RC105	-	100	98	96	**92**	**85**	**78**	70	62	53	44	34	23	12
RC106	-	-	-	100	95	**89**	**81**	71	62	53	**44**	33	23	12
RC107	-	-	-	100	**99**	**92**	83	75	67	**57**	46	35	24	12
RC108	-	-	-	-	100	**96**	**86**	**81**	69	58	**48**	36	25	13

To further test the robustness of our algorithm, Table 3 shows our computational results for *m*-VRPTW instances of Solomon's benchmarks. A new best result for Solomon's instances, RC107, was found by our algorithm. According to Table 3, our algorithm can generate equivalent or better solutions in 39 instances. For other instances, 8 are close to the best. 2 are at most 4% worse. Therefore, our algorithm is a good method compared with previous methods.

4. CONCLUSION

This paper presented a new two stage heuristic algorithm for the vehicle routing problem with time windows and a limited number of vehicles. Our clustering process in the first stage is proved to be able to get good initial solution for the next stage. By introducing FSFS idea and new operator *Transfer,* the algorithm can be strengthened significantly. The use of the EP and the adaptive memory data structures has played great roles in the algorithm. In the procedure of minimization of distance cost, our improved AMP can get to the best solutions faster. Experimental results have proved the effectiveness and robustness of our algorithm. The algorithm can return either the best published solutions or solutions very close in quality on all Solomon benchmarks.

Table 3 Comparison of best results for VRPTW

instance	*Previous best result*		*Our best results*		*Gap*
R101	19	1645.79	19	650.81	0.31%
R102	17	1486.12	17	1498.32	0.8%
R103	13	1292.68	13	1295.76	0.3%
R104	9	1007.24	9	1007.65	0%
R105	14	1377.11	14	1377.11	0%
R106	12	1251.98	12	1251.98	0%
R107	10	1104.66	10	1150.86	4.1%
R108	9	960.88	9	980.95	2.1%
R109	11	1194.73	11	1215.68	1.7%
R110	10	118.59	10	1119.00	0.03%
R111	10	1096.72	10	1096.72	0%
R112	9	982.14	9	982.14	0%
C101	10	828.94	10	828.94	0%
C102	10	828.94	10	828.94	0%
C103	10	828.06	10	828.06	0%
C104	10	828.78	10	828.78	0%
C105	10	828.94	10	828.94	0%
C106	10	828.94	10	828.94	0%
C107	10	828.94	10	828.94	0%
C108	10	828.94	10	828.94	0%
C109	10	828.94	10	828.94	0%
RC101	14	1696.94	14	1696.95	0.01%
RC102	12	1554.75	12	1558.57	0.2%
RC103	11	1261.67	11	1261.67	0%
RC104	10	1135.48	10	1135.68	0..01%
RC105	13	1629.44	13	1629.44	0%
RC106	11	1424.73	11	1424.73	0%
RC107	11	1230.48	11	**1225.32**	**-0. 4%**
RC108	10	1139.82	10	1139.96	0..01%
R201	4	1252.37	4	1252.37	0%
R202	3	1191.70	3	1191.70	0%
R203	3	939.54	3	941.41	0.1%
R204	2	825.52	2	825.52	0%
R205	3	994.42	3	997.25	0.3%
R206	3	906.14	3	906.14	0%
R207	2	893.33	2	901.48	0.9%
R208	2	726.75	2	726.75	0%
R209	3	909.16	3	909.16	0%
R210	3	939.34	3	954.12	1.6%
R211	2	892.71	2	906.19	1.5%
C201	3	591.56	3	591.56	0%
C202	3	591.56	3	591.56	0%
C203	3	591.17	3	591.17	0%
C204	3	590.60	3	590.60	0%
C205	3	588.88	3	588.88	0%
C206	3	588.49	3	588.29	0%
C207	3	588.29	3	588.29	0%
C208	3	588.32	3	588.32	0%
RC201	4	1406.91	4	1406.94	0.002%
RC202	3	1367.09	3	1368.44	0.09%
RC203	3	1049.62	3	1058.37	0.8%
RC204	3	798.41	3	798.46	0.01%
RC205	4	1297.19	4	297.65	0%
RC206	3	1146.32	3	146.32	0%
RC207	3	1061.14	3	1061.14	0%
RC208	3	828.14	3	828.44	0.03%

5. ACKNOWLEDGEMENTS

This work has been supported by the National Natural Science Foundation of China (10471045, 60433020), the program for New Century Excellent Talents in University (NCET-05-0734), Natural Science Foundation of Guangdong Province (031360, 04020079), Excellent Young Teachers Program of Ministry of Education of China, Fok Ying Tong Education Foundation (91005), Social Science Research Foundation of MOE (2005-241), Key Technology Research and Development Program of Guangdong Province (2005B10101010, 2005B701011 18), open research fund of National Mobile Communications Research Laboratory (A200605), Key Technology Research and Development Program of Tianhe District (051G041) and Natural Science Foundation of South China University of Technology (B13-E5050190).

6. REFERENCES

[1] G. B. Dantzig and R.H. Ramser. “The Truck Dispatching Problem”. Management Science 6, 1959, 80–91.

[2] Savelsbergh, M., “Local search in routing problems with time windows”. Annals of Operations Research 4, 1985, 285 – 305.

[3] M. L. Fisher, “Optimal Solution of Vehicle Routing Problems Using Minimum K-trees”, Operations Research 42, 1994, 626-642.

[4] Desrosiers.J, Sauve′.M., Soumis.F.,“Lagrangian relaxation methods for solving the minimum fleet size multiple traveling salesman problem with time windows”. Management Science 34, 1988, 1005–1022.

[5] Desrochers, M., Desrosiers, J., Solomon, M.A., “A new optimization algorithm for vehicle routing problem with time windows”. Operations Research 40, 1992, 342 – 354.

[6] Brian Kallehauge,Jesper Larsen and Oli B.G.Madsen, “Lagrangian duality applied to the vehicle routing problem with time windows”. Computers & Operations Research 33, 2006, 1464-1487.

[7] L. M. Gambardella, E. Taillard, G. Agazzi, "MACS-VRPTW: A Multiple Ant Colony System for Vehicle Routing Problems with Time Windows", In D. Corne, M. Dorigo and F. Glover, editors, New Ideas in Optimization. McGraw-Hill, 1999.

[8] P.Shaw,”Using Constraint Programming and Local Search Methods to Slove Vehicle Routing Problems.” Principles and Practice of Constraint Programming, 1998, 417-431.

[9] R.Bent and P.Van Hentenryck, “A Two-Stage Hybrid Local search for the Vehicle Routing Problem with Time Windows”, Technical Report

CS-01-06,Departmen of Computer Sicence,BrownUniversity,2001.

[10] Rodolfo Dondo, Jaime Cerda, "A cluster-based optimization approach for the multi-depot heterogeneous fleet vehicle routing problem with time windows". European Journal of Operational Research,2006.

[11] H. C. Lau, M. Sim, and K. Teo. "Vehicle routing problem with time windows and a limited number of vehicles". European Journal of Operational Research, 2003, 148:559 – 569.

[12] Andrew Lim and Xingwen Zhang. "A Two-Stage Heuristic for the Vehicle Routing Problem with Time Windows and a Limited Number of Vehicles" International Conference on System Science,2005

[13] Y. Rochat and È. D. Taillard. "Probabilistic Diversification and Intensification in Local Search for Vehicle Routing". Journal of Heuristics, 1995, 1:147-167.

[14] Solomon M.M. "Algorithms for the vehicle routing and scheduling problem with time window constraints." Operation Research 35,1987, 254-265.

[15] B.De Backer, V.Furnon, P.Shaw, P.Kilby, and P.prosser. "Solving Vehicle Routing Problems Using constraint Programming and Metaheuristics". Journal of Heuristcs, 2000, 6:501-523.

[16] G.Kindervater, M.Savelsbergh. "Vehicle Routing: Handiling Edge Exchanges."In E.Aarts and J.K. Lenstra, editors, Local Search in combinatorial Optimization, chapter 10, John Wiley & Sons Ltd,1997, pp. 337-360.

[17] B-O.Braysy, "A Reactive Variable Neighborhood Search Algorithm for the Vehicle Routing Problem with Time Windows," Working Paper, University of Vaasa,Finland, 2001.

Understanding Corporate B2B Web Sites' Effectiveness

Zhengrong Pan, Mingguang Zhang
School of Electrical and Information Engineering,Lanzhou University of Technology
Lanzhou, Gansu, 730050, China
Email:zhangmingg8@163.com

ABSTRACT

The globalization of economic activities has led multinational business-to-business (B2B) firms to use their corporate web sites extensively to communicate with their stakeholders. This paper examines three factors of information that influence corporate web sites' effectiveness and the components and limitation of web service in a B2B context. Results indicate that informativeness, usability and quality of information are important factors for users. And the Web Services architecture provides emerging standards and technologies for most areas, but it still has significant limitations.

Keywords: Corporate web site; B2B; Informativeness; Usability; Quality of information; Web Services.

1. INTRODUCTION

In recent years, many firms have been undergoing profound transformations in the pursuit of reducing costs and providing better services to customers. Whole value chains between firms and their customers and suppliers are being re-engineered. Inter-organizational systems have been an essential part of managing these changing relationships. In addition, the increasingly internationalization of products and companies has created the need for cross-border inter-organizational relationships that rely on B2B e-commerce systems. B2B e-commerce has been identified as an emerging trend, with companies planning strategies to capture benefits. It is estimated that worldwide B2B e-commerce transactions will reach $2775 billion in 2004—a growth of over $2500 billion since 2000 (eMarketer, 2001). There are also estimates that by 2004, B2B sales on the Internet will be six times larger than B2C sales (Haig, 2001). E-commerce may become the primary low-cost method for B2B transactions. Buyers and sellers benefit from the productivity and profitability improvements associated with e-commerce (Deeter-Schmelz et al., 2001). With a relatively small investment, companies can achieve remarkable savings: e.g., leaders in e-commerce achieved 41% improvement in cycle time and a 10% reduction in staff and can generate up to a 13:1 ROI by using the full potential of B2B e-commerce (Kearney, 2002).

A firm's corporate web site often acts as a gateway for visitors to find critical information that shapes their perceptions about the company as a whole as well as the company's offerings. If visitors to a corporate web site do not find helpful and relevant information, they are likely to form a poor perception about the company and may ultimately visit and transact with competitors who provide the necessary information. In a B2B context, the information aspects in a corporate web site become even more important due to the financial risk and the complexity of the exchange process with typical B2B transactions (Peppers & Rogers, 2001). The impact of the Internet in international marketing is expected to be much greater for B2B than for B2C (Samiee, 1998).Therefore, this paper focuses on the information content of corporate web sites and the characteristics of information important for corporate web site effectiveness firstly.

Secondly, the World Wide Web has enabled businesses to share documents across a generalized, global network. It facilitate EC, for example, sellers have been able to publish company and product information via their Websites, and to some degree, search engines have allowed buyers to find and analyze this information. Yet such searches are not reliable because of the diverse systems and data presentations. Moreover, sellers on the WWW generally do not use industry-wide standard transaction templates for accessing product information and executing purchase transactions. This has limited the ability of automated services to find sellers and conduct automated transactions. So in order to improve Web Services, the components and limitation of web service in a B2B context must be studied.

2. BUSINESS-TO-BUSINESS ELECTRONIC COMMERCE

Although B2B e-commerce is a popular subject, systematic empirical evidence regarding firm and innovation characteristics that are crucial in its adoption is lacking. Rather, research has focused on a range of other topics including: how the establishment of virtual intermediaries is influencing B2B e-commerce; what firms can do to enhance the successful implementation of the Internet; how the Internet is being used to manage supply chain components; what marketing activities are influenced by the Internet; how to select B2B e-commerce models; what abilities determine participation in electronic markets; how B2B e-commerce influences coordination costs; what issues should be considered before a firm embarks on B2B ecommerce; what firms can do to increase the effectiveness of the World Wide Web as a marketing tool; what structural dimensions result in effective digital markets (e.g., online auctions, bid systems); and how inter firm communications have been influenced by the Internet (e.g., Boyle & Alwitt, 1999; Dou & Chou, 2002; Lynn, Lipp, Akgu¨n, & Cortez,2002; Wilson & Abel, 2002). Our understanding of B2B ecommerce phenomena is still at a very early stage with additional research needed.B2B e-commerce is defined as a supply chain innovation that generates cross-firm process integration. It is the use of the World Wide Web to secure the trading of goods, information, and services before, during, and after the sale (Barnes-Vieyra & Claycomb, 2001). It includes electronic data interchange (EDI) and Web- and Internet-based applications. B2B e-commerce has changed the manner in which cross-firm transactions are conducted, resulting in lower costs and improved supply chain management. Online marketplaces widen the spectrum of buyers. By bringing together large numbers of buyers and sellers and by automating transactions, electronic markets give sellers access to new customers, expand the choices available to buyers, and reduce transaction costs. Specifically, B2B e-commerce fosters: choice-flexibility in

marketplaces, a wider selection of supply chain partners; speed-faster decision-making, time to market, and transactions; streamlining-increased efficiency and productivity because of less paperwork, elimination of bureaucratic systems, better control of inventory; and response-instant response to unforeseen problems and opportunities (Haig, 2001).

3. FACTORS OF INFORMATION IN A CORPORATE B2B WEB SITE

A B2B corporate web site is a gateway between a company and its prospects, customers, and other stakeholders. This differs from B2C web sites in that the prospects and customers are other businesses rather than end consumers. The goal of this paper is to broach the factors of information that influence web site effectiveness.

A B2B firm's corporate web site is typically used to provide general information about the company, trade or application-specific information and news, and product information (Robbins & Stylianou, 2003). Most multinational firms provide links from their corporate web sites to their country-specific web sites where the company operates. Often, the actual commercial transactions (involving order-placement, payment, tracking of shipment, etc.) are carried out in those country-specific web sites because of the idiosyncratic nature of the business environment (laws and regulations, taxes, etc.) in each country and varying company policies (with regard to pricing, packaging, delivery, etc.) in different countries. Thus, the effectiveness of a B2B firm's corporate web site may be more dependent on the ability of the web site to provide relevant information to generate enough interest to convert a visitor into a customer or partner. In this manuscript, three factors of information (informativeness, usability of information, and quality of information) in a corporate B2B web site are considered. These factors of information were selected based on prior research, content analysis of corporate web sites.

3.1. Informativeness

Marketing practitioners and academic researchers contend that one of the primary purposes of a company's web site (corporate or country-specific) is to provide information to prospects, customers, and other stakeholders (Chen, Clifford, & Wells, 2002; Eighmey, 1997; Lohse, Bellman, & Johnson, 2000; Palmer, 2002; Robbins & Stylianou, 2003). A content analysis of corporate web sites of Fortune 500 companies indicated that a large amount of content (information) was provided on these sites (Robbins & Stylianou, 2003). Anot_ her content analysis revealed that firms' country- specific web sites are being used as part of an integrated communic_ ation strategy to serve higher objectives such as creating de_ sire and action (Sheehan & Doherty, 2001). Consequently, the ability of a web site to make a visitor feel that the web site has communicated something of value is viewed as one of the most important predictors of web site effectiveness. Informativeness is the ability of a web site to make inform_ ation available. In this sense, informativeness may be view_ ed as static information or content available on a web site (Hoffman & Novak, 1996). Usability of information, on the other hand, reduces consumers' search costs by helping to access the information available on a web site more efficie_ ntly (Bakos, 1997). Furthermore, it must be noted that a site may score high on informativeness regardless of the manner in which it is presented. Thus, concerns of information ove_ rload or formatting are unrelated to the ability of the site to provide information of value. Lastly, informativeness is co_ nceptualized as a perceptual construct. Therefore, informat_ iveness is not the same as the actual amount of information available on a web site, even though these are expected to be correlated. In an exploratory study by Keeney (1999), a need to maximize product information was expressed as one of t_ he main objectives related to e-commerce. Palmer (2002) f_ und content (as perceived by student juries) of corporate w_ eb sites to be a significant predictor of the web sites' success. Chen and Wells (1999) found perceived informativeness of a web site to be the second most important factor in explain_ ning variance in visitors' attitudes toward the web site.

Therefore, The higher the level of perceived informative_ eness in a corporate web site, the higher is the web site's effectiveness.

3. 2. Usability

Usability relates to users' perceived ability to utilize the information provided by a web site. These perceptions are often strongly influenced by the web site interface design, specifically its navigational and interactive features (Palmer, 2002). Good navigational design and use of interactive features in a web site help users to find relevant information easily and quickly, thereby making information more usable. Thus, usability is viewed at an abstract level as pertaining to the pragmatics of how a user perceives and interacts with a web site (Barnes & Vidgen, 2002). Many researchers have proposed and found empirical support for the notion that more interactivity usually results in a more satisfying web site experience, which in turn increases the effectiveness of web sites (Novak, Hoffman, & Yung, 2000).Berthon, Pitt, and Watson (1996) suggested that the usability of a web site is critical in converting site visitors from blookersQ to bbuyers.Q In other words, usability at a web site influences the web site's effectiveness. Ghose and Dou (1998) studied interactive functions in web sites and found that the greater the degree of interactivity (leading to higher usability of information) in a web site, the higher is the web site's attractiveness. An exploratory study revealed ease of use as one of the concerns of web site users. Specifically, maximizing ease of user interface, making access easy and simple, and finding desired products were mentioned as major objectives (Keeney, 1999). Clearly, an efficiently executed web site design that enhances ease of use is an important factor in determining web site effectiveness (Eighmey, 1997). On the other hand, factors such as web page complexity that lower ease of use also lower attitude toward the site (Stevenson, Bruner, & Kumar, 2000). Similarly, in a study based on B2B sites, navigability was found to improve commitment toward the site (Bauer, Grether, & Leach, 2002). Finally, Palmer (2002) also found interactivity and navigation to have significant association with a corporate web site's success. Therefore, the higher the level of perceived usability in a corporate web site, the higher is the web site's effectiveness.

3. 3. Quality of information

This refers to the quality of the content in a web site and the suitability of the content from a user's perspective (Barnes & Vidgen, 2002). Just as the perceived amount of information (informativeness) is important, the quality of information in a corporate web site is also important in shaping visitors' initial perceptions about the company.

Because of advances in information technology, providing timely and accurate information is not supposed to be difficult for organizations. Equally important is that customers' expectations today are extremely high, and they are unwilling to accept outdated, inaccurate, unbelievable information in a web site. Empirical evidence suggests that a web site that provides outdated or inaccurate information leads to visitor dissatisfaction and abandonment of the web site (Barnes & Vidgen, 2002; Palmer, 2002; Wang, 1998). In other words, poor quality of information is detrimental to a web site user's experience leading to reduction of web site effectiveness. Similarly, relevant and accurate information at a web site has been found to improve visitors' experience at a web site, leading to higher web site effectiveness (Robbins & Stylianou, 2003). Therefore, the higher the level of perceived information quality in a corporate web site, the higher is the web site's effectiveness.

4. WEB SERVICES IN A CORPORATE B2B WEB SITE

Web Services focus mainly on data type standards, schema expression languages, and common communication methods. This platform takes advantage of the ubiquity of the WWW by using, XML, and UDDI. Web Services use open standards and have been submitted to the World Wide Web Consortium (W3C) [1]. While UDDI attempts to meet the needs of categorization, schemas, and transaction templates, it does not provide sufficient support. In order to meet all the requirements for a successful architecture, UDDI needs to be combined with other frameworks, such as RosettaNet [2], ebXML [3], Universal Business Language (UBL), or the Semantic Web.

4.1. Limitations of Web Services

While the Web Services architecture represents a step forward, limitations still exist for automated services.

1) Business categorization is unreliable and variable

The UDDI registry is able to locate organizations that belong to specific business types; however, since organizations can register with a variety of categorization schemes, the UDDI registry does not support economical and reliable searching for all businesses of a given type. Searchers must query several different business categorization schemes to find businesses of a specific type.

Jewell and Chappell have written the following about the anticipated limited market reach of the UDDI registry:

It's probably not realistic to expect software to dynamically discover and use new businesses on the fly in the near future. Realistically, human analysts need to browse a UDDI portal that allows customized searches and queries to discover the businesses they are interested in working with. It's more likely that software will contain the logic necessary to locate and integrate with Web Services for companies that have been predetermined. It's also likely that businesses will set up private UDDI registries that they can share with their approved partners to facilitate B2B integration.

2) Product and service representations are nonexistent or inconsistent

Because of the UDDI yellow pages, the registry can help searchers find businesses that offer a given class of products or services, but it does not support automated searches for specific products or comparisons of products and prices across vendors. For example, with UDDI it is possible to find companies that manufacture TVs, but not find all vendors who sell high-definition, stereo, 27 in. color TVs. Moreover, it is not possible for an automated search client to collect the model numbers, prices, and features of all those offered by the registered manufacturers. WSDL definitions do not exist for categories of products and services that should have equivalent or similar descriptive fields. While some industries may standardize their WSDL signatures or use existing ones from the existing pool, formal involvement with the registry does not provide any incentive for participants to standardize or adopt standards defined by other organizations. This lack of standardization significantly impedes its usefulness [4].

3)Transaction templates are nonexistent or inconsistent

While UDDI provides the tModel structure [UDDI Version 3], which can be used by many businesses, the structure allows any number of external schemes for categorization. Since any registered entity can define tModels, many different specifications for the same business or product will exist.

4)Discovery services are limited because of a lack of standards

There are important implications for deficiencies for both the seller and the searcher. Companies lack defined product and service representation schemas and transaction template definitions to guide the development of automated commerce support software[5]. This hinders the development of automation for both sellers and buyers, because different sellers expose different automation interfaces. Sellers' lack of standards leads to problems on the client side: it is difficult to search and discover competing vendor services. Without common field names and transaction templates, search clients cannot be developed that effectively exploit these fields [6]. Moreover, indexing services cannot use standard interfaces to collect product and service information across vendors. Because of this heterogeneity, clients must be programmed to interact with specific seller interfaces, making the network fragile and extremely difficult to maintain.

In summary, the Web Services architecture provides emerging standards and technologies for most areas, but it still has significant limitations. So we must study the complete components of Web Services.

4.2. Web Service components

The Web Services architecture involves the Web Services Description Language (WSDL), Simple Object Access Protocol (SOAP), and Universal Description, Discovery and Integration registry (UDDI) [7].

1) Web Services description language

WSDL specification provides a set of rules for defining XML schema. The WSDL specification is a machine-readable fingerprint that describes an automated service and its attributes and is loosely analogous to an interface or header file used to describe the interface and behavior of a module in a program. It was developed by Microsoft, Ariba, and IBM and has been submitted to the W3C. Client software can query services for their WSDL definition. If the client software is prepared to make use of the services, it can interact with them through specific calls to the services. WSDL defines XML definitions for basic data types, including specific common data types that correspond to specific data fields.

2) Simple Object Access Protocol

SOAP is responsible for transferring XML encoded information from one computer to another. Because it uses HTTP, Web servers allow it to pass through firewalls with

relative ease, though companies are currently exploring ways to maintain adequate security when using it [8]. SOAP also supports standard data types that can be used for requests made to services and provides asynchronous messaging and event notification to help the host and client programs communicate.

WSDL and SOAP are widely supported in many different languages. Implementation libraries exist in languages such as Java, .NET, Perl, Python, Visual Basic, etc. [9]. Together WSDL and SOAP provide the framework for the definition and execution of remote calls on services such as enterprise objects that can be used to both publish data and execute transactions. While both WSDL and SOAP support the use of standard data types, no standard exists for which standard data types will correspond to specific data fields used by different service APIs. Also, field names like ''product code'' or ''product description'' have not been standardized.
3) Universal description, discovery, and integration registry system

The UDDI registry provides a central location for registering and finding services within the Web Services architecture. Currently, public services created by IBM, Microsoft, SAP, and HP replicate registrations and provide redundant lookup services. Because of registration replication, participants need to register with only one registry to be included in all UDDI servers. UDDI has been criticized because it relies too heavily on a centralized registry [10]. Moreover, it may take time to develop functioning public directories that could be used to conduct business [11]. While this technology appears to have potential, the lack of standards is a limitation.

5. CONCLUSION

Based on the review of academic and trade press literature, three dimensions of information (informativeness, usability, and quality of information) were identified that might influence a corporate B2B web site's effectiveness. Using prior research and pretests, reliable and valid perceptual measures for these constructs were broached.

While the Web Services architecture is a technological step forward, the lack of required standards limits its usefulness and widespread adoption. The technology om/pages/search/group60381.adp.

toolbox of today is sufficient to support EC, but the standards that must be developed are conceptual standards required for efficient technological implementation. In particular, a shared set of APIs should support representation definition from top-down industry consortiums as well as from ottom-up participants. It should enable a standard API for specific transactions and industries over time, while allowing for individual and changing needs.

6. ACKNOWLEDGEMENTS

This work was supported by the National 863 Scientific Project Development Fundation P.R. China under grant no.2002AA415270.

7. REFERENCES

[1] W3C, Web Services Architecture Working Group, 2002. http:// www.w3.org/2002/ws/arch/.
[2] RosettaNet, RosettaNet Background Information, 2001. http:// www.rosettanet.org/background/.
[3] ebXML, ebXML Technical Architecture Specification, 2001. http://www.ebxml.org/specs/ebTA.pdf.
[4] T. O'Reilly, Amazon Web Services API, 2002. http://www.oreillynet. com/cs/weblog/view/wlg/1707.
[5] G. Chakrabortya, P. Srivastavaa, et al. "Understanding corporate B2B web sites' effectiveness from North American and European perspective", Industrial Marketing Management, 2005, 34 (2005): 420-429.
[6] C.A. Conan, L.D. Douglas, et al. "Marketplace and technology standards for B2B e-commerce: progress, challenges, and the state of the art", Information & Management, 2005, 42 (2005): 865-875.
[7] T. Bellwood, L. Cle´ment, et al. UDDI Version 3 Specifications Document, UDDI Working Group, 2002. http://www.uddi.org/pubs/uddi-v3.00-published-20020 719.htm.
[8] C.C. Albrecht, "How clean is the future of SOAP?", Communications of the ACM, 2004, 47(2): 66-68.
[9] T. O'Reilly, Web Services Center, 2002. http://webservices.oreilly. com/.
[10] M. Baker, UDDI v3 Announced; did anybody notice? O'Reilly Network Weblogs, 2002. http://www.oreillynet.com/pub/wlg/ 1680.
[11] S. Masood, UDDI: Better on the Inside, Infoconomy, 2002 http://www.infoconomy.com

The research of logistics scheduling optimization model based on agile rules and high performance algorithm

Fuzhong Wang [1], Fumao Yu [2], Zuzhi Shen [1]
[1] School of Management, Zhejiang University , Hangzhou 310058, China;
E-mail：wfz212126@zju.edu.cn
[2] School of Management Hangzhou University of Electronic Science and Technology, Hangzhou 310018, China
E-mail：morefast@126.com

Abstract

In order to provide an agile logistics scheduling plan to achieve the optimal level of logistics service and transportation cost and to consider the dynamic characteristics of vehicles as well in the practical scheduling process, agile rules of logistics scheduling and agile scheduling general work logic based on the rules were proposed, and an agile optimization model of logistics scheduling based on the rules was established. By hierarchically decomposing the model with hierarchical control mechanism, a hierarchical solving algorithm with high performance was put forward. The result of example shows the algorithm reduces empty loading rate more than other heuristic algorithms.

Keywords agile rules, logistics, scheduling optimization model, high performance algorithm.

1. Introduction

The core problem of logistics scheduling is the routing arrangement of vehicles. The research on logistics agile scheduling technology can be summarized as vehicle routing and scheduling problems (VRP/VSP) in Operations Research. Because most of them belong to NP-hard problems, we can only seek for heuristic solutions which are the main contents and objectives of many researches[1,2,3]. But, the VRP/VSP are of serious computing complexity, so the scholars have spent much energy constructing high quality heuristic algorithms [4,5,6]. Moreover, except a few types of logistics scheduling problems under special conditions have gotten relatively good algorithms, most of the logistics scheduling problems in practice are still lack of very effective algorithms.

In practical problems, to respond the demands of dynamic and changeful logistics market rapidly, delivering timely and credibly becomes the core of logistics market competition, from this aspect, not only the cost but also the agility of logistics scheduling are the main objectives to be considered. Furthermore, in the practical scheduling process, considering the strong dynamic characteristics of transportation vehicles, it is difficult to achieve the optimal level of logistics service and transportation cost. Therefore, researching on the technology of logistics agile scheduling optimizing to satisfy the above demands is necessary.

As the agile scheduling process is a decision process in substance which runs under certain control mechanisms, the decision process of agile scheduling needs to respond rapidly and effectively to environmental changes. So the authors have researched logistics agile scheduling optimization technology based on hierarchical control mechanism [7,8] and “agile rules” in the process of the research on “logistics agile scheduling decision support systems” which is a sub-subject of the great and key project of science and technology in Zhejiang province named “the development and research of modern logistics system”. The paper puts the emphasis to introduce the construction of scheduling optimization model and its solving algorithm with high performance.

2. The construction of “agile rules” of logistics scheduling

In order to adapt well to the dynamic and changeful logistics market and dynamic characteristics of logistics transportation vehicles, it is always necessary to summarize the “agile rules” of logistics scheduling based on the practical experiences of logistics scheduling practical operators and optimization application researchers to effectively support the agile and rapid producing and implementing of logistics agile scheduling optimization plans. Therefore, we construct several types of “agile rules” of logistics scheduling with close relationship.

1)scheduling priority rules

They are the chief rules to implement logistics scheduling agility on a policy level, which are set in a layer sequence of scheduling policy:

①the first grade of scheduling priority rule: arrange the distribution place with earlier delivery demand first

② the second grade of scheduling priority rule: for distribution places with same delivery time, arrange the distribution place which is far away from the center first

③the third grade of scheduling priority rule: for assured distribution places, arrange the distribution place, which needs vehicles of bigger tonnage first

④the fourth grade of scheduling priority rule: for assured vehicles, arrange the goods with higher weight first.

2）selection rules of scheduling ways

Because the self-provided transportation vehicles are usually at different and dynamic states, such as “usable state”, “used state”, “maintenance state”. Only the self-provided transportation vehicles are at the “usable state”, which can be the so-called “valid resources”. When arranging the loading or distribution

tasks, we should use the description ways of "rules" to set the rule according to the state of "set of valid resource":

The first selection rule: IF the "set of valid resource" is not empty, THEN we use "self-scheduling way" and based on grade one to grade four of "scheduling priority rule" to arrange loading tasks.

The second selection rule: IF the "set of valid resource" is empty, THEN we use "outsourcing way" and based on the first and the second grade of "scheduling priority rule" to arrange loading tasks.

3)fast optimization realization rules

For agile scheduling, an important principle is to supply good and fast services and give attention to the benefit at the same time, the construction of scheduling optimization model and the research on relevant and fast solving algorithms should follow the principle, so we present the following four types of fast optimization realization rules.

①the decomposing rule of distribution network sequence

Based on the second grade of scheduling priority rule, we decompose the distribution network into a sub-network sequence consisted of feasible paths from distribution center to distribution places. At the same time, according to the sequence, we respectively produce network geographical information database of distribution network corresponding to sub-networks based on GIS to support the fast searching for the paths in the process of scheduling optimization.

②the integration rule of geographical information and distribution information

The distribution network geographical information database based on the above sequence structure and the produced distribution demand database based on loading bill, according to the idea of data warehouse, we construct decision support multi-dimension database based on multi-dimension variable tables with geographical information and distribution information produced by the sub-networks, to effectively support the constructing of scheduling optimization model with low complexity and the computing of its fast optimization from the level of data support.

③ hierarchical decomposing rule of scheduling optimization model

Because the scheduling optimization model constructed for the whole distribution network is often a very complex nonlinear mixed-programming model, whose solving is a NP-hard problem, based on the sequence decomposing of distribution network, for the objective with the smaller empty loading rate, according to the first grade of scheduling priority rule and based on the assured time-dimension, we decompose the constructing of scheduling optimization model for total distribution network into the constructing of scheduling optimization model for sub-network, further decompose into the constructing of scheduling optimization model (NP) for paths on the sub-network, and then according to the second grade of scheduling priority rule, for the place dimension, further hierarchically decompose the model into scheduling optimization model $(NP)_h$ for distribution places on the path, which decreases the complexity of the model greatly, so as to effectively support fast optimization computing from the support aspect of low complexity model.

④the combination rule of model driving and data driving of optimization computing process

According to the second grade of priority rule, in the process of hierarchically solving $(NP)_h$, firstly, for the objective of resource vehicles' configuration and optimization, we transform the $(NP)_h$ model into a 0-1 programming model to solve the "set of optimization selection vehicles" for assured distribution places by the way of model driving, secondly, based on "set of optimization selection vehicles", according to the third and the fourth grade of scheduling priority rules, to realize the configuration optimization of goods by the way of "data driving", so as to effectively support fast optimization computing from the support aspect of algorithm.

3. The construction of logistics scheduling general work logic based on "agile rules" and scheduling optimization model

3.1. The construction of logistics agile scheduling general work logic based on "agile rules"

In the process of scheduling optimization, the core demand information of distribution is the distribution place、the category and number of distribution goods and corresponding delivery time. Moreover, each distribution amount of goods is always corresponding to three dimensions: place dimension (distribution place dimension)、goods number dimension and time (delivery time) dimension, so based on the idea of data warehouse, it is very reasonable to design the metadata (distribution number) based on the above three dimensions into multi-dimension variables.

We denotes respectively time dimension、goods number dimension、place dimension and distribution amount as t、i、l, B, so the multi-dimension variable of B can be described as $B(t,i,l)$, whose set of variables makes up of multi-dimension database with geographical information and distribution information.

After B is incised by assured delivery time t_0 according to time dimension, the variable set $B(t_0,i,l)$ for assured delivery time is produced. Based on $B(t_0,i,l)$, after B is incised by assured distribution place i_0 according to place dimension, the variable set $B(t_0, i_0,l)$ for assured delivery time and assured distribution place is produced. We denote distribution network geographical information database with sequence structure as $A(i,j)$, in which j is a feasible path number corresponding to distribution center to distribution place.

According to the "agile rules", it is convenient to construct the general work logic of agile scheduling under the support of data warehouse technology. The figure is as follows:

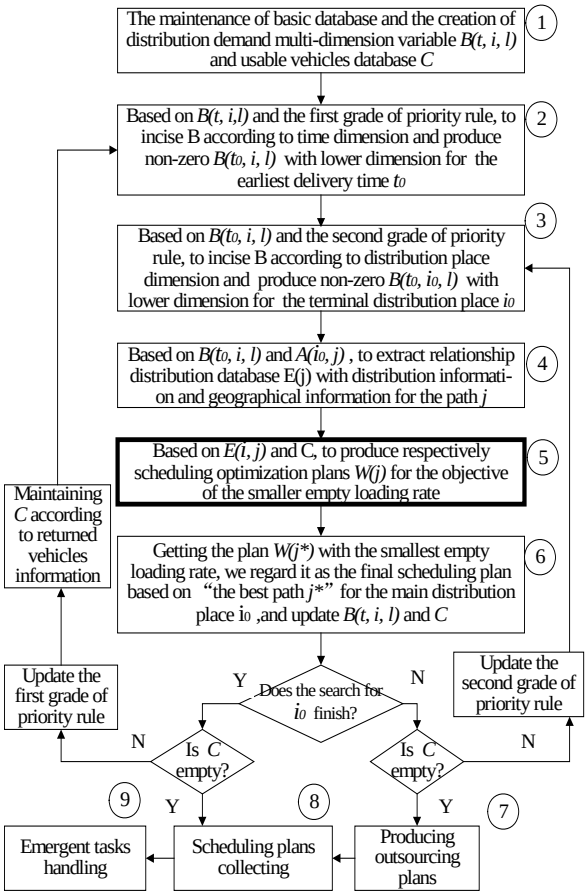


Fig. 1. The general work logic of agile scheduling

Because logistics market demand is dynamic and changeful, the logistics scheduling plans change often, which is the most important element to urge logistics scheduling agility. Therefore, based on data warehouse and the first to the third step of work logic, and after *E(j)* is produced fast in the fourth step, how to produce the scheduling optimization plan *W(j)* for the assured path *j* in the fifth part fast will be a key part to achieve agile scheduling. So constructing the scheduling optimization model and whose objective is to achieve the smaller empty loading rate for the assured delivery time and paths and providing fast and effective optimization algorithm with high performance have become a foundational work to achieve agile scheduling.

3.2. The construction of scheduling optimization model for assured delivery time and paths

We denote set of distribution place as $H=\{h|h=1,2..k\}$, which is of delivery demand at the assured delivery time on the assured path *j*, the *H* is sorted from the small to the big by the distance to distribution center, when $h=k$, it denotes main distribution place i_0 . The constructing and decomposing process of general scheduling optimization model (NP) for assured delivery time and paths is as follows:

1)the definition of decision variables

t_{hf} ----vehicle selection variable, $t_{hf} \in \{0,1\}$, $h \in H = \{1,2...k\}$, $f \in E = \{1,2...F\}$, when $t_{hf} = 1$, it denotes the f^{th} vehicle which is appointed to distribution place *h* is selected, on the contrary, it isn't selected.

x_{hfl} ----the l^{th} goods loading weight when the f^{th} vehicle is appointed to distribution place *h*, $l \in L_h$ (L_h denotes set of distribution goods at distribution place *h*)

2) the definition of scheduling empty loading rate $\eta^{(j)}$ for the assured path *j*

j-----assured path number

$d_h^{(j)}$ ---on the path *j*, the distance between distribution center and distribution place *h*

u_f ----the capacity tonnage of the f^{th} vehicle

then:

$$\eta^{(j)} = 1 - [\sum_{h=1}^{k}\sum_{f=1}^{F}\sum_{l=1}^{L_h} d_h^{(j)} t_{hf} x_{hfl} / \sum_{h=1}^{k}\sum_{f=1}^{F} d_h^{(j)} t_{hf} u_f]$$

3) the construction of the general scheduling optimization model for assured path *j*

$$(NP): \max[\sum_{h=1}^{k}\sum_{f=1}^{F}\sum_{l=1}^{L_h} d_h^{(j)} t_{hf} x_{hfl} / \sum_{h=1}^{k}\sum_{f=1}^{F} d_h^{(j)} t_{hf} u_f]$$

$$S.t. \quad \sum_{f=1}^{F}\sum_{l=1}^{L_h} t_{hf} x_{hfl} = G_h^{(t_0)}, \quad h = 1,2...k$$

$$\sum_{h=1}^{k}\sum_{l=1}^{L_h} t_{hf} x_{hfl} \le u_f , \quad f = 1,2...F$$

$$t_{hf} \in \{0,1\}, \quad h = 1,2...k , \ f = 1,2...F$$

$$x_{hfl} \ge 0 , \ h = 1,2...k , \ f = 1,2...F , \ l = 1,2...L$$

In the model, $G_h^{(t_0)}$ denotes the total distribution demand for delivery time t_0 and distribution place *h*

Obviously, the (NP) model is a nonlinear mixed-programming model, in order to solve a "satisfactory solution" for the (NP) model fast, we further hierarchically decompose the (NP) model into a scheduling optimization model $(NP)_h$ for assured paths and the distribution place *h*.

4) The construction of the scheduling optimization model $(NP)_h$ produced by the decomposing hierarchically based on the (NP) on the path *j* for the assured distribution place

After the distribution optimization at distribution place *h+1*, we denote:

f '-------the vehicle number with the remnant loading weight in the "set of optimization selection vehicles" appointed the distribution place *h+1*

M_{h+1} --- the remnant loading weight of the f'^{th} vehicle (when $h+1=k$, $M_{h+1} = 0$)

$E_h = \{f_1, f_2, ..., f_{M_h}\}$ ----"set of usable vehicles" for *h*.(when $M_{h+1} \neq 0$, $f' \in E_h$, and the biggest loading weight of the f'^{th} vehicle is revised as $u_{f'} = u_{f'} - M_{h+1}$)

So, by the decomposing based on the model (NP), the scheduling optimization model for the assured distribution place *h* is as follows:

$$(NP)_h : \max[\sum_{i=1}^{M_h}\sum_{l=1}^{L_h} d_h^{(j)} t_{hf_i} x_{hf_i^l} / \sum_{i=1}^{M_h} d_h^{(j)} t_{hf_i} u_{hf_i}]$$

$$S.t. \quad \sum_{i=1}^{M_h}\sum_{l=1}^{L_h} t_{hf_i} x_{hf_i^l} = G_h^{(t_0)} \qquad (1)$$

$$x_{hf_i^l} \le t_{hf_i} u_{f_i} \qquad i \in E_h, l \in L_h \qquad (2)$$

$$t_{hf_i} \in \{0,1\} \qquad i \in E_h$$

$$x_{hf_i^l} \ge 0 \qquad i \in E_h, l \in L_h$$

Although the scale of the model $(NP)_h$ decreases greatly, it is still a non-linear mixed-programming. So we develop the approximate algorithm for solving the model $(NP)_h$ by the way of "integrated model- algorithm" based on data driving and model driving.

4. Solving of the model $(NP)_h$ based on data driving and model driving

4.1. The solving process

Because the solving process of the model $(NP)_h$ virtually includes the idea of two-grade optimization, the first grade is to optimize for the vehicle configuration of assured distribution places, the second grade is to optimize for the goods loading of assured distribution vehicles, the concrete and approximate solving process is as follows:

1)Process one: the data driving process

STEP ONE: After the solving for the model $(NP)_{h+1}$, under the restriction of the remnant loading weight M_{h+1} of the f'^{th} vehicle, according to the distribution demand information for the distribution place *h* at delivery time t_0 in the distribution demand database *D*, the f'^{th} vehicle is loaded for the objective of full loading according to the fourth grade of priority rule, and the loading weight $x_{hf'}$ is produced for the f'^{th} vehicle at distribution place *h*, and then distribution demand database *D* and usable vehicles database E_h are revised immediately.

STEP TWO: Computing the total remnant distribution demands for the distribution place *h*

$$F_h = G_h^{(t_0)} - M_{h+1}$$

2) Process two: the producing process of model

The data driving based on the process one transforms the model $(NP)_h$ into a 0-1 programming model $(0-1)_h$, which regards the smaller "total remnant loading weight" as the objective, regards F_h as loading constraints and regards the vehicle selection variable t_{hf_i} as the only decision variable.

$$(0-1)_h \qquad \min \sum_{i=1}^{M_h} t_{hf_i} u_{f_i}$$

$$S.t. \quad \sum_{i=1}^{M_h} t_{hf_i} u_{f_i} \ge F_h$$

$$t_{hf_i} \in \{0,1\} \qquad i \in E_h$$

3) Process three: the approximate solving process

STEP ONE: We use the fast algorithm with the core of the bisection search to solve $(0-1)_h$ and based on the non-complete solution of the best solution $t^*_{hf_i}$ ($i \in E_h$) to produce "set of optimization selection vehicles" E'_h, and then the "usable vehicles" database is updated immediately and set of usable vehicles E_{h-1} is made for the next distribution place *h-1*, at last, the "remnant loading weight" M_h of vehicles with the lowest optimization grade is calculated.

STEP TWO: Based on $t^*_{hf_i}$ ($i \in E_h$), the model $(NP)_h$ will be transformed into the following loading equation with constraints:

$$(P)_h \qquad \sum_{i \in E'_h}\sum_{l \in L_h} x_{hf_i^l} = G_h^{(t_0)}$$

$$S.t. \quad 0 \le x_{hf_i^l} \le u_{f_i} \qquad i \in E'_h, l \in L_h$$

For the objective of full loading based on $(P)_h$ according to the third and the fourth grade of scheduling priority rules and by the way of data driving, the best solution x_{hfl} of the model $(NP)_h$ is produced.

4.2. Testing of the algorithm's advantages

In order to test the approximate solving algorithm with high performance proposed by the paper, we have solved respectively for several examples by the first algorithm proposed by the paper and the second algorithm presented by the reference [9], which is a heuristic algorithm of vehicles optimization scheduling for multi-vehicle transportation used often. It is found that the lines empty loading rate of the first algorithm's solving is 3.8 to 7.5 percent less than that of the second algorithm's solving, and the running speed of the first algorithm improves one times much more than that of the second algorithm. So the first algorithm proposed by the paper is proved to have obvious advantage for the objective of the smaller empty loading rate at least. Considered the length of the paper, we only present one example to prove its advantages.

We denote the distribution center as *O*, and give a group of practical data which is summarized and simplified by a practical case.

Distribution Place no.(h)	distance from *O* to *h* (unit: km.)	The first to the sixth goods distribution demand quantity (unit: tonnage)					
		1	**2**	**3**	**4**	**5**	**6**
1	98	2.7	0	1.0	0	1.6	0
2	132	0	3.7	2.9	0	3.1	0

3	171	1.2	2.5	0	1.8	0	1.3
Usable vehicle no. (l)		**1**	**2**	**3**	**4**	**5**	**6**
tonnage		8	5	4	2	2	1

The first algorithm and the second algorithm solve respectively after optimization, the results are as follows:

According to the second algorithm, the non-zero approximate solution x_{hlk} (namely, the loading weight of goods k loaded by the l^{th} vehicle appointed to place h) is as follows:

x_{361}=1, x_{351} =0.2, x_{354} =1.8, x_{342} =2, x_{332} =0.5,
x_{336}=1.3, x_{232} =2.2, x_{222} =1.5, x_{223} =0.4, x_{225} =3.1,
x_{213} =2.5, x_{111} =2.7, x_{113} =1.0, x_{115} =1.6

So, the lines empty loading rate η is 8.983 percent, the total tonnage kilometer Q calculated by used vehicles tonnage and running distances is 3255 tonnage kilometer.

According to the first algorithm, the non-zero approximate solution x_{hlk} is as follows:

x_{322} =2.5, x_{324} =1.8, x_{326} =0.7, x_{341} =1.2, x_{346} =0.6,
x_{242} =0.2, x_{212} =3.5, x_{213} =1.4, x_{215} =3.1, x_{253} =1.5,
x_{151} =0.5, x_{131} =2.2, x_{133} =1.0, x_{135} =0.8, x_{165} =0.8

So, the lines empty loading rate η is 1.477 percent, the total tonnage kilometer Q calculated by used vehicles tonnage and running distances is 3007 tonnage kilometer.

It is very easy to understand that the lines empty loading rate of scheduling plan solved by the first algorithm is 7.506 percent less than that of scheduling plan presented by the second algorithm, and the total running distances decrease 248 tonnage kilometer.

5. Conclusion

According to the authors' practices in several processes of logistics agile scheduling system design and development, the modeling technology of logistics scheduling optimization models based on the hierarchical control mechanism and "agile rules" and high performance algorithm proposed by the paper play an important role in achieving logistics scheduling agility. Especially in implementing logistics agile scheduling decision support system on the background of Yiwu logistics market, not only the mean empty loading rate has been descended 5 to 8 percent, but also the service level has been improved greatly, it is really to achieve logistics scheduling agility in a high performance. It shows that a type of "knowledge" such as "agile rules" put into the full process of modeling and solving of scheduling optimization is meaningful for solving the high performance problems of scheduling optimization algorithm.

6. References

[1]P.H. Hansen, B. Hegedahl, et al, A heuristic solution to the warehouse location-routing problem. European Journal of Operational Research. July 1994,76（1）: 111-127.

[2]B. Julien, S.L. David, A location based heuristic for general routing problems. Operations Research. July/August 1995,43（4）: 649-660.

[3]D. Tuzun, I.B. Laura, A two-phase tabu search approach to the location routing problem. European Journal of Operational Research. July 1999,116（1）: 87-99.

[4]H. Dennis, F. Richard, et al, A Robust Solution Approach to the Dynamic Vehicle Scheduling Problem. Transportation Science. November 2004, 38（4）: 447-459.

[5]M.C. Ann, S. Martin, Efficient Insertion Heuristics for Vehicle Routing and Scheduling Problems. Transportation Science. August 2004, 38（3）:369-378.

[6]L.B. Alexandre, G.C. Teodor, et al, A Guided Cooperative Search for the Vehicle Routing Problem with Time Windows. IEEE Intelligent Systems. July/August 2005,20（4）: 36-50.

[7] L. Xiao, Z.Z. Shen, The planning research of logistics system based on semi-hierarchical control. China Journal of Highway and Transport, April 2003, 16（2）: 120-123.

[8] L. Xiao, Z.Z. Shen, et al, The simulation and planning research of bus network based semi-hierarchical control mechanism. Journal of Donghua University, April 2004, 30（2）: 56-61.

[9] J. Li, Y.H. Guo, Theory and methods of logistics distribution vehicle optimization scheduling. Beijing: China Logistics Publishing House, 2001, 2-9, 96-98.

A Topic-Based and Distributed Search Engine for Business Intelligence

Yulian Fei[1] ,Anding Zhu[2], Guangmin Wang[3]
College of Computer & Information Engineering, Zhejiang Gongshang University
No.149 Jiaogong Rd. Hangzhou, Zhejiang, China 310035
Email: [1]fyl@mail.zjgsu.edu.cn [2]hz_andrew@yahoo.com.cn [3]gmwang@mail.zjgsu.edu.cn

Abstract

Site-based search engines are always the bases for most web-based business intelligences. Specific business applications need specific search engines. The paper introduces a specifically designed search engine – ToSE (Topic-oriented Search Engine) to achieve business intelligence for the applications of comparative shopping and intelligent negotiation.

With the development of electronic commerce websites, a tremendous amount of merchandise information has been published on Internet. Price plays a very important role for e-business decision-making. Starting from a few of classified candidate e-business sites, ToSE collects price pages and parses price information. Full-text price contexts are then analyzed and transformed into structured data to support decision-making. Because price information on Internet is changeful, ToSE must complete one updating cycle within an acceptable period. An URL spanning tree algorithm is proposed to speed up traversing the hyperlink graph of websites. And a price-sensitive PageRank algorithm is proposed to score the most valuable and believable price information.

Keywords
Topic-oriented, Search Engine, URL Spanning Tree Algorithm, PageRank Algorithm.

1. INTRODUCTION

With the blast increasing of Internet, the information capacity of Internet is growing larger and larger. Nowadays, e-business on Internet is becoming a common business style. A lot of B2B, B2C and C2C e-commerce websites are established and a lot of business information can be obtained from Internet. However, compared to the rapid increasing of e-commerce websites, mining business values deeply on Internet is relatively rare to be concerned with.

Comparison-shopping is one of the applications of business intelligence. Nowadays, lots of goods information is published on Internet. And price takes a very important role of e-business. How to guide the consumers to purchase cleverly? And how to provide the consumers the very information they are looking for? Intelligent negotiation is another application of business intelligence. In order to bring out a reasonable negotiation suggestion, the intelligent negotiation system must collect and analyze sufficient business information.

These high-level business intelligences need a search engine to crawl among the e-commerce sites and collect the useful business information from them. With the rapid increasing of Google, the values of search engines have been receiving increasing attention. The traditional full-text search engine clusters more than thousands of computers to crawl Internet and gathers gigas of web pages. But the refresh cycle of traditional search engines is about several weeks or even several months. It is too long for the business information, since business application needs fresh data. So the business intelligences need a relative little-scale topic-based search engine with a short refresh cycle. On the other hand, according to the specific topic of business intelligence, the amount of candidate web sites concerned with is not very large. The search engine only needs crawl among these candidate web sites. So different from the traditional search engine, the topic-based search engine has its own characteristics.

And another difference between topic-based search engine and traditional full-text search engine is the search depth. A full-text search engine may only provide the relative page links. But a topic-based search engine needs extract the objects from the pages and then organize them into structure contents. So what the search engine replies the visitors are not pages, but more accurate content elements.

Since the traditional full-text search engine cannot provide personalized service, a lot of alternative search strategies are under researching [1]. The research group of Nobuyoshi SATO [2] develops a real-time retrieval system named as Cooperative Search Engine (CSE) to gather fresh information on Internet. CSE uses the Forward Knowledge (FK) to fasten the traversing process. And CSE establishes the change aware index system to reduce the computation of repeated crawling. Wray Butine [3] proposes a distributed topic-based search engine, which runs on a cluster of Linux Machines in a fast local network. And topic models are proposed in his/her system to provide a content analysis of documents. To date, the Internet search and mining group of Microsoft Research Asia [4] pays a great attention to search Internet not only by the structure of hyperlinks but also the contents themselves. Heng Tao Shen [5] develops an efficient document searching system by effectively summarizing and maintaining all documents within the network with different granularity. But due to the different search purposes, specific topic-based search engines might not be suitable for each other.

Furthermore, in order to provide the most valuable information the users ask for, a search engine must score and rank the web pages. PageRank algorithm [6], which is based on the hyperlink citation order of the webpage, might not be very suitable for a topic-based search engine. Taher H. Haveliwala [7] proposes a topic-sensitive PageRank algorithm using the topic of the context. And by using linear combinations of biased PageRank vectors computed by a set of representative topics, this score algorithm is reported can generate more accurate rankings than the traditional rank algorithm.

This paper proposes a topic-based search engine ToSE (Topic-oriented Search Engine) designed for business intelligence.It not only collects information but also processes the contents intelligently.

The rest of this paper is organized as follows. In section 2, the architecture of the scalable and distributed topic-based search engine is introduced. In section 3, the URL spanning tree algorithm is proposed to shorten the update cycle. And in section 4, the price-sensitive PageRank algorithm for the specific business intelligence

applications is proposed too. At last, the search engine is applied to Chinese steel industry. About 65 Chinese steel B2B candidate web sites are crawled. And the search results are analyzed.

2. ARCHITECTURE OF TOSE

First of all, an idea of the search engine architecture is proposed as follow. ToSE adopts a distributed architecture and the architecture block diagram is shown as Fig. 1.

2.1 Architecture of ToSE

In order to shorten the update cycle, ToSE integrates several spiders and processing terminal computers into a distributed system. Because of the delay of the web servers, a spider computer cannot acquire a web page very fast. It is estimated that a spider thread averagely can only download about 4.8 normal size web pages within a minute. But with today's computer hardware, most of the computational and networking resources are wasted while waiting for the replies from web servers.

One solution is using multithread spiders [8]. The CPU and network resources can be switched among multithread programs, so when one spider thread is waiting for replies, the other spiders can take charge of these resources synchronously. Another solution is clustering several spider computers into a system [9]. In fact, both these two solutions are taken into in a practical system.

No matter which solution is selected, the spider threads need be scheduled carefully. A spider scheduler program is introduced to schedule the crawling plan and balance the workload. The tasks are distributed among every spider computer. In order to simplify the scheduler algorithm and reduce the communication among multi-threads, the paper set one spider thread to crawl one website. So the scheduler program only needs to assign websites to every spider thread on every spider computer. To date, the spider computers work behind a local area network, but under the scheduler program, they work together as one. So the output of these distributed computers will be several times of one spider computer. Except the crawling task, the spider has other heavy load jobs such as recognition of target pages, extraction of URL characteristics, page denoising, and extraction of page objects, and so on. The scheduler program needs to assign these heavy load computation jobs to terminal computers when they are waiting for the reply of the server. So the usage of CPU resource is always at a high percentage.

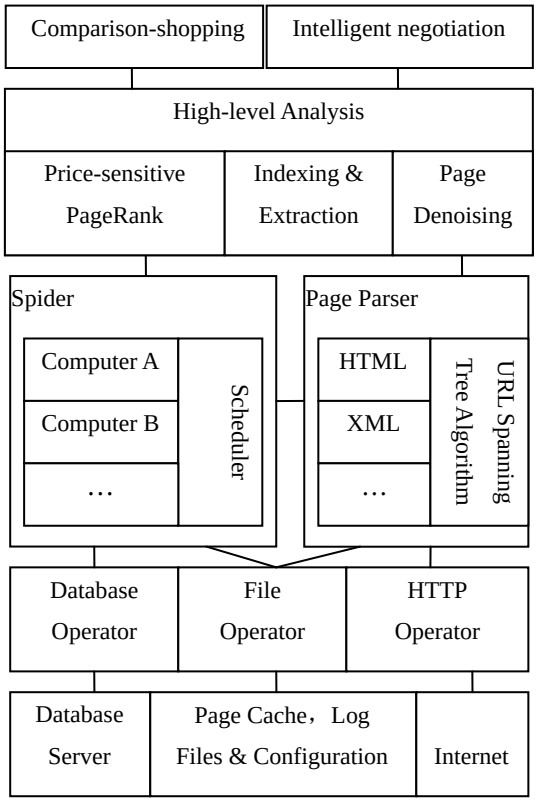


Fig. 1. The Architecture Block Diagram

Considered that only hundred thousands of target pages are correlative to a specific topic and the data transmission rate is very fast within a LAN, a central database server is enough to be used to store the download web pages and the structured information.

In the architecture, a scalable parser platform is implemented to parse different web page styles such as HTML, XML and so on. The page parser is one of the most important parts in the system, because this module extracts the hyperlink structure, the URL characteristics and the page contents. Nowadays, there are many dynamic page technologies such as JSP, PHP and so on. Usually, all these type of pages will be ultimately transformed into HTML type more or less and be sent to the users' computers. However, there are also many differences among different page types. So a scalable parser structure seems to enhance the extensibility of our system. URL spanning tree algorithm module tries to track the hyperlink traversing route and analyze the URL characteristics and finally guide a rapid traversing.

Furthermore, because different types of goods have different keywords and different attributes, Database operator and file operator modules need provide scalable storage structure and configuration information to organize them.

High-level analysis will post-process the downloaded web pages. Price-sensitive PageRank module will score the pages. Page denoising module will remove the unnecessary elements such as ad and copyright information. Extraction module will transform pages into structured data.

2.2 Workflow of ToSE

Since ToSE is a topic-based search engine, it starts from a

classified candidate websites. Every type of goods has presetting candidate websites. The workflow is listed as follow.

Step 1: Start up the scheduler program and assign a candidate website from the seeds list;

Step 2: Download the homepage of this website;

Step 3: Analyze the homepage and extract the first layer of hyperlink;

Step 4: Download each page and check whether this page is the target one;

Step 5: If this page is a target page, write down the URL and send to the URL spanning tree analysis module. Otherwise jump to Step7;

Step 6: The URL spanning tree module analyzes the URL characteristic and generates the target URL generation rules;

Step 7: Analyze each page's hyperlink structure and filter the next layer's URLs under the target URL generation rules;

Step 8: If the URL does not belong to the candidate website, which means that it is another candidate website related to the topic, add this new website to the seeds list;

Step 9: Throwaway non-target URLs and keep target ones;

Step 10: If the URL has been already downloaded and analyzed, skip this one;

Step 11: Download the target URLs and jump to Step 4.

The scheduler program will be notified that more target pages need to be extraction and analysis. The scheduler program will append them into the job list and in the near future, it will assign these jobs to a suitable terminal computer.

Since there are different types of target webpage according to different websites, there is no universal target webpage recognition algorithm. In our architecture, a scalable target webpage recognition structure is adopted. Different recognition algorithms can be integrated in the architecture easily and smoothly.

And on the other side, the frequency of the keywords within the webpage and the structure of the webpage can be used to determine whether the webpage is the target one. Usually, a target webpage might contain more keywords, so a frequency threshold can be set to tell these pages apart. But in practical, this method is not always available. So structure analysis can be taken as an enhance algorithm. Due to the dynamic web page technology, the target web pages of lot of websites are combined by a same template, so the similarity between the webpage and target template can be used to recognize the target webpage.

The topologic of a website's hyperlink is a graph. According to the traversing strategy above, A Depth First Multi-way Spanning Tree Traversing algorithm is taken in practice. With this strategy, the useful target pages will be gathered directly.

3. URL SPANNING TREE ALGORITHM

Nowadays, since most B2B, B2C and C2C web servers are established by dynamic web page technology [10], which means that the ultimate web pages sent to the visitors are not prepared in advance. All the dynamic pages are combined by the server using the data queried from the database and the web page templates at the required time. So the similar target pages always come with similar URLs, this is called the local theory. The search engine can fasten the traversing process by analyzing the target URL characteristics.

On the other side, due to the dynamic page generation technology, a lot of pages might not be cited in other pages. These pages are hidden behind the hyperlink graph. So not all the information a web server contained can be traversed through the hyperlink graph.

Furthermore, an e-commerce web server always contains much more other information pages than price and attributes ones, such as ads, news, copyrights and so on. These kinds of pages are not interested in very much.

Although the topology structure of hyperlink is a graph, the URL arrangement always takes a form of tree structure like directories. According to the local theory, a target price page always comes with another target price page. So the URLs of target pages are similar and much near in the URL vector space. This paper proposes a URL spanning tree algorithm based on URL vector space.

3.1 By Means of Spanning Tree Algorithm Traversing Hyperlink Graph

The appearance of the dynamic URL always takes the format like

http://www.xxx.com/dirxxx/pagexxx.jsp?id=xxx&att=xxx

.

In this format, a dynamic URL always combined with several parts:

- Directory xxx;
- General page entry, and the type of these pages might be jsp or asp and so on;
- Question mark "?";
- Query condition equations such as id number and other attributes and connected by "&".

So these URL subparts can be looked as one element of the URL vector. And the URL vector is denoted as $U(dir, page, id, att1, ...)$ as. Since the elements of the URL vector represent a structure of directory, the URL can be divided into different layers of a tree as shown in Fig. 2.

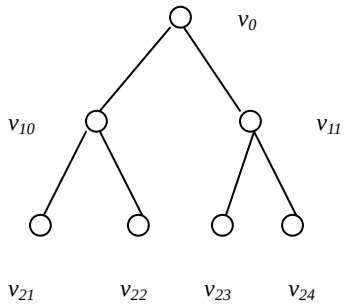


Fig. 2. URL Spanning Tree

The URL is divided into several subparts by the separator "/". And every subparts between "/" are denoted as a node of the graph. So according to the URL directory structure, the spanning tree structure can be got easily. For example, the root node v_0 can be set as www.xxx.com, and the first layer node v_{10} as dirxxx in Fig. 2. Then the hyperlink graph is transformed into a spanning tree.

In order to find the dynamic target URLs, the parent node of every webpage need to be written down. If this page is a target page, the weight of the parent node will be increased. After weight of the parent node exceeds one

threshold, this parent node will be taken as the target URL root node.

After the target URL root node is generated, the Depth First Multi-way Tree Traversing algorithm is selected to traverse the hyperlink graph. Some large-scale websites may have several classified root URLs according to classified topics. So although one root node is generated, the traversing process will not stop tracking the URL characteristics.

Since the structure rebuilding of websites is not often, the directory structure and URL characteristics of the website are not changed within a long time. And in order to collect fresh price pages, the spider needs to visit the candidate websites within a short period, for example, one day. So the extraction and recognition of URL characteristics need to run only once. In the next refreshing period, the spider only needs to traverse directly with the stored URL generating rules. It will fasten the traversing process very much.

3.2 Finding the Hidden Pages out Using URL Query Conditions

Many target pages cannot be found through the hyperlink graph. On some websites the amount of these pages may be very large. A traditional full-text search engine based on hyperlink graph cannot find them out.

Since a URL sent to the server sometimes takes with the query conditions, it is advisable to try to query the server's database with different keywords and conditions combined by the URL rules. The $att1 = ???\&att2 = ???\&...$ elements are filled in and if the page exists, the server will respond. With this brute force searching strategy, lot of target pages behind the hyperlink can be found out.

On many websites, retrieval of price and goods' attribution information is based on querying. There are only few example price pages are listed on current trends board. So practically, the web server provides an entrance to retrieve its background database. And there are usually only few query entrance pages.

After the URL generation rules are created, the query conditions can be guessed. Some page templates may provide query keywords for the visitors to select. So these keywords can be extracted from the page. But because every page template is not same with each other, the extraction of these keywords is somehow difficult. Of course, this job can be completed manually, because according to a specific topic, the amount of candidate websites is not very large and the amount of query page templates is not very large too. And this job will be done only once except the website is rebuilt.

3.3 The Incremental Refreshing Process

Normally, on a website, within a short period, few new price items are appended. So since the crawl process is refreshed every day, not all the possible pages need to be visited, but only the new coming items. These items may not be combined into new pages. And with the query algorithm, new items may be listed with the old items within one page.

In a traditional full-text search engine, the new appended items cannot be found before the full hyperlink graph is traversed completely. And the search engine will regard the pages with partial new items as new pages. So the information stored in the search engine database may contain redundant information. In the ToSE, the items are transformed into structure data. So with this object-based storage strategy, the pages are stored repeatedly.

Incremental refreshing strategy is adopted to collect the new items. If the page-extracting module finds the old items, which have been collected already in the database, it will stop tracking and try the other queries.

So the first traversing of our topic-based search engine may cost lot of time. But the following traversing will cost little of time so that the refreshing period will be reduced greatly.

4. PRICE-SENSITIVE PAGERANK ALGORITHM

Since the amount of information on Internet is very large, a price rank algorithm is needed indeed. Because the price information on Internet may be not creditable and timely, in order to give a reasonable rank to the price information, three aspects as follow will be concerned with.

4.1 The score of the entire website

Although the PageRank algorithm based on citation orders cannot provide topic-sensitive rank scores, the rank score of website, which is provided by a traditional full-text search engine, can be used to represent the importance of the website on Internet. Google scores every website with ten ranks. In the price-sensitive rank algorithm, the Google rank score is taken as the importance score of the website. If the website has a higher Google rank score, it is thought of having a high authority, and will be more probable to be visited. So the information it provided will be much more believable and valuable.

Furthermore, some other factors are also taken into account. Such as member amount, visit frequency. These types of information cannot be gained from the Google rank score. And on a lot of professional websites, the member amount and visit frequency will be published. A website with larger member amount and higher visit frequency might be much more authoritative.

4.2 The attention degree of the webpage

Many professional websites might count the webpage clicks and publish the count on the webpage. This information might be regarded as how many people are interested in this webpage. It is reasonable to think that the more clicked webpage might be more valuable and creditable.

4.3 The temporal information of the webpage

In a business environment, the price information is a type of temporal information. So timely price information will be more valuable than outdated one. And the influence of the outdated price information will decrease rapidly.

Finally, all these three factors are combined into the price-sensitive rank scores by the formula below.

$$Score_{price} = \sum(Weight_{site} \cdot Score_{site} + Weight_{page} \cdot Score_{page} + Weight_{time} \cdot Score_{time})$$

Where

$$\begin{cases} Weight_{site} + Weight_{page} + Weight_{time} = 1 \\ Score_{site} = \sum (Weight_{Google} \cdot Score_{Google} + \\ \quad Weight_{members} \cdot Score_{members} + Weight_{visits} \cdot Score_{visits}) \\ \quad and \quad Weight_{Google} + Weight_{members} + Weight_{visits} = 1 \\ Score_{page} = Score_{clicks} \\ Score_{time} = 10 \cdot e^{-a \cdot elapse + a} \end{cases}$$

The scores of members, visits and clicks are normalized into ten ranks as Google page ranks. And the weights of these factors are related to the specific topic. Normally, these factors are positive correlated.

The scores of site, page and time are normalized into ten ranks too. And the weights of these factors are related to the specific topic too. But generally, these three factors are not correlated.

The score of time is an exponential function of the elapse days after the webpage is published. In the formula, the first day (*elapse* = 1), the score of time equals 10. When the elapse days increase, the score will decrease rapidly. And the decreasing rate is controlled by the parameter of *a*. The parameter of *a* is related to the specific topic.

The selection of these weight parameters will be evaluated and adjusted according to the specific topic to provide more valuable items preferentially.

Because lot of price and goods attributes pages take the form of lists and tables, they are rarely cited by other pages. The price-sensitive PageRank algorithm can provide more believable and creditable rank scores than traditional citation order based rank technology.

5. EXPERIMENTS

In order to test the ToSE, Chinese steel industry is chosen as one application field, because in China, steel trading websites typically adopt dynamic web page technology.

First about 65 steel candidate websites are collected by searching manually through traditional full-text search engines, such as Google and so on. The ToSE is started at 12 p.m. every day and the search result is refreshed every day. In this experiment, three DELL PC computers with 3.0G Pentium CPU, 1G DDR RAM and 100M bps Ethernet network are used. The scheduler program runs on one computer. And every computer runs 4 crawl threads synchronously and every thread is assigned to crawl one website. Other computational jobs are assigned dynamically by scheduler program to an idle computer.

As an overview, table 1 shows the time consuming of every refreshing period. Table 2 shows the pages amount after every refreshing. Table.3 shows the price item amount after every refreshing.

Table 1. Time Consuming of Every Refreshing Period

	1st Day	2nd Day	3rd Day	4th Day
Time Consuming	5.69 Hours	20.38 Minutes	27.17 Minutes	23.63 Minutes

Table 2. Page Amount after Every Refreshing

	1st Day	2nd Day	3rd Day	4th Day
Page Amount	12,783	13,549	13,990	14,471

Table 3. Price Item Amount after Every Refreshing

Period	1st Day	2nd Day	3rd Day	4th Day
Item Amount	61,358	63,035	65,254	68,330

According to Table 1, 2 and 3, it can be found that the first time of refreshing cost a lot of time, but the next time of refreshing cost little time. So ToSE can refresh every day. Other topic searches can be run together on the same platform and be refreshed every day.

6. CONCLUSIONS

This paper reports a topic-based search engine – ToSE for business intelligence. Because of the specific requirement of the business application, a traditional full-text search engine might not be suitable very much. The traversing algorithm is improved by using the structure of URL and a price-sensitive PageRank algorithm is proposed to meet the specific requirement of business intelligence. And practically, ToSE works well under the business environment.

7. REFERENCES

[1] J. Pokorn, "Web Searching and Information Retrieval", Computing In Science & Engineering, JULY/AUGUST 2004, 43-48

[2] N. Sato, M. Uehara, et al, "A Distributed Search Engine for Fresh Information Retrieval", 12th International Workshop on Database and Expert Systems Applications, Sept.2001, 211-216

[3] B. Wray, L. Jaakko, et al, "A Scalable Topic-based Open Source Search Engine", Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI'04), IEEE 2004,228-234

[4] T. Qin, T.Y. Liu, et al, A Study of Relevance Propagation for Web Search, The 28th Annaul International ACM SIGIR Conference (SIGIR'2005), August 2005, 408-415

[5] H.T. She, Y. F. Shu, et al, "Efficient Semantic-Based Content Search in P2P Network", IEEE Transactions on Knowledge and Data Engineering,July 2004,16(7).813-826

[6] L. Page, S. Brin, et al, The PageRank Citation Ranking: Bring Order to the Web, Technical report, Stanford University, Stanford, CA, 1998

[7] T.H. Haveliwala, "Topic-Sensitive PageRank: A Context-Sensitive Ranking Algorithm for Web Search", IEEE Trans. On Knowledge and Data Engineering,August 2003,15(4):. 784-796

[8] H. Chen, M. Chau, et al, CI Spider: a tool for competitive intelligence on the Web, Decision Support Systems, Vol 34, 2002,1–17

[9] M. Kaewmoracharoen, R. Walters, "Enhancing Electronic Highway Design Standards and Specifications", the 2005 Mid-Continent Transportation Research Symposium, Ames, Iowa, August 2005, 1-11

[10] P. G. Kumar, "Active Server Pages: Technology for Creating Dynamic Web Pages and Webenabled Databases", Workshop on Multimedia and Internet Technologies 26th-28th February,AC, 2001

A Forewarning Research for Enterprises' Marketing Environmental Risk on the Base of Entropy Right Markov Chain

Guoqiang Xiong, Hailei Liu
School of Business Administration, Xi'an University of Technology
Xi'an , 710054, P.R. China
Guoqiang Xiong
Northwestern Polytechnical University
Xi'an , 710072, P.R. China
Email: xqg168@163.com

ABSTRACT

Aiming at the multiple uncertainty characteristics involved in marketing environmental risk such as fuzziness, randomness and so on, the entropy concept that is used to measure the uncertainty of system state is introduced into Markov chain to forecast the risk state of enterprises' marketing environment of the coming period. Firstly, the index system that is employed to forewarn the marketing environmental risk is established and fuzzy evaluation is taken to multi-index data of the present time period. Secondly, according to the experiences and the facts of enterprises' marketing environment, the grades of risk states are set to make sure the grades of different periods. Thirdly, state transition probability matrixes of different step-sizes are established and then the Markov chain method that is based on entropy right is applied to solve the forecasting probability of each risk state of the coming time period. Finally, a case is given and proves the validity of this method.

Keywords: Marketing Environment, Risk Forewarn, Markov Chain, Entropy Right, Fuzzy AHP.

1. INTRODUCTION

Under the market economy background, the enterprises' marketing environments are increasingly complex and ever changing. Grasping the changes in marketing environment accurately and then forewarn and control of the potential risk are becoming an important task. But because they arise from numerous uncertain factors, which cannot be anticipated beforehand, the marketing environmental risk has the characteristics of fuzziness, randomness and so on.

In recent years, the quantitative forewarning research of marketing risk increases day by day, and while mainly focusing on marketing organization risk, marketing project decision risk as well as finance risk and so on[1,2,3,4], it pays less attention to marketing environmental risk. Concerning the analytical methods, scholars mainly adopt methods such as analytical hierarchy process (AHP), fuzzy evaluation, artificial neural network and so on [5, 6, 7]. Although these methods can describe fuzziness of marketing risk quantitatively, and also are easy to understand and calculate, they cannot reflect randomness and uncertainty of marketing environment. The essential of risk forewarning is to monitor and manage the random process information, and then the state of the coming period can be deduced. The literatures about the application of Markov chain method in random process are much more [8, 9], but they are primarily related to disaster forecast and hydrology forecast. It can hardly see the forewarning research on marketing environmental risk.

Basing on the above consideration, the entropy concept is introduced into Markov chain to forecast the risk state of enterprises' marketing environment of the coming period. The main process includes: Firstly, fuzzy AHP method is taken to evaluate the risk state quantitatively of the present time period; Secondly, the evaluation value is divided into different grades according to certain ranges; On the basis, the method of entropy right Markov chain is used to forecast the risk state of enterprises' marketing environment of the coming period according to the state transition probability; Finally, a case is given and proves the practicability and validity of this method.

2. THE BASIC THOUGHT ABOUT FOREWARNING OF ENTROPY RIGHT MARKOV CHAIN

2.1 Principle of Markov Chain

Markov chain is an important kind of random process [10], which describes the random phenomenon: the present state is aware, probability distribution of the future state has nothing to do with the past state. The randomness of marketing environmental risk conforms to the supposition. Specifically, if a random process $\{X(n), n=1, 2, \ldots\}$, for arbitrary k non-negative integer $n_1,n_2,\ldots,n_k$ $(0\le n_1<n_2<\ldots<n_k)$, arbitrary $i_1, i_2, \ldots, i_k$, and arbitrary natural number m, conditional probability satisfies:

$$P\{X(n_k+m)=j \mid X(n_1)=i_1, X(n_2)=i_2, \cdots X(n_k)=i_k\} = P\{X(n_k+m)=j \mid X(n_k)=i_k\} \quad (1)$$

Then it is said that $\{X(n), n=1, 2, \ldots\}$ is Markov chain.

Definition[11] $P\{X(n+m)=j \mid X(n)=i\}$ is the probability of the process of that from state i at the moment n to state j at the moment $n+m$ pass through transition of m steps. It is called the transition probability of m steps, record as $p_{ij}(n, n+m)$.

Markov chain which transition probability does not rely on the moment n is called homogeneous Markov chain, record as $p^{(m)}=(p_{ij}^{(m)})$, and $p_{ij}^{(m)}$ denotes the probability from state i to state j pass through transition of m steps. Particularly, when $m=1$, $P\{X(n+1)=j \mid X(n)=i\}$ is the transition probability of one step, record as $p^{(1)}$:

$$P^{(1)} = \begin{pmatrix} p_{11}^{(1)} & p_{12}^{(1)} & \cdots & p_{1n}^{(1)} \\ p_{21}^{(1)} & p_{22}^{(1)} & \cdots & p_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}^{(1)} & p_{n2}^{(1)} & \cdots & p_{nn}^{(1)} \end{pmatrix}. \quad (2)$$

Estimation of state transition probability usually applies frequency approximation. Firstly the data sequence is divided into several kinds of state, record as i_1, i_2, …, i_n. Then the transition probability of the data sequence from state i to state j pass through transition of m steps is

$$P_{ij}^{(m)} = \frac{M_{ij}^{(m)}}{M_i}. \quad (3)$$

Including: $M_{ij}^{(m)}$ is the times of that from state i to state j pass through transition of m steps; M_i is the times of emergence of state i. Let the last state of the sequence is k, when m=1, due to the uncertainty of the transition of the state at the last moment, the last data should be take out when we calculate M_k. The solution principle of each Mi infers from this when m>1.

2.2 Markov Chain on the Base of Entropy Right

The transition process of risk state of marketing environment is uncertain, with Markov chain to forecast the risk state of enterprises' marketing environment in the future is based on the transition probability between states. The state transition probability matrix P_{ij} describes the whole possibility of that the system transfers from state i to other states. Entropy can be used to measure the diversity degree of this possibility. The entropy concept comes from thermodynamics, afterwards is introduced into information theory by Shannon. According to the entropy thought, it can be used to measure the uncertainty of the system state changes. Therefore, when state i transfers to other states, the transition probability is p_{ij} (i=1, 2, ⋯ , n), the entropy[12] which describes the diversity degree of the possibility is

$$\theta_i = -\sum_{j=1}^{n} p_{ij} \ln p_{ij}. \quad (4)$$

Thus, when Markov chain is used to take risk forewarning for marketing environment, firstly we can solve transition probability matrix of different step-sizes and then calculate entropy value of each state in each transition probability matrix; secondly forecast the transition probability of each state of certain time period; thirdly take entropy value of the corresponding state as weight, make weight sum to forecasting probability of the same state; finally the risk state of certain time period can be forecasted. The above process is the method of Markov chain on the base of entropy right.

3. STEPS OF A FOREWARNING METHOD FOR MARKETING ENVIRONMENTAL RISK ON THE BASE OF ENTROPY RIGHT MARKOV CHAIN

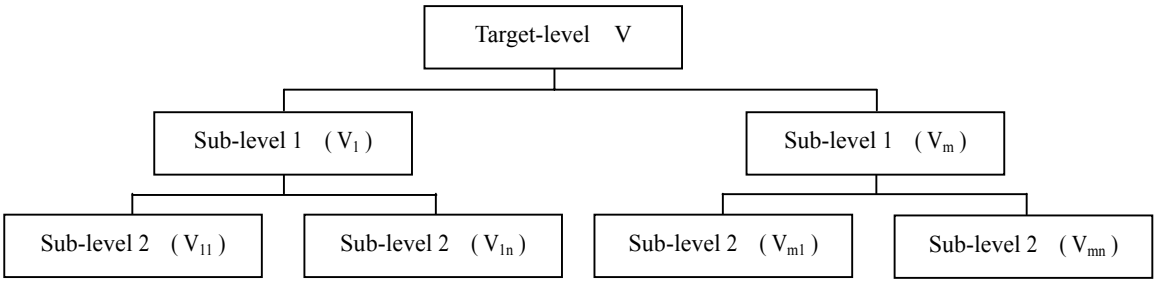


Fig. 1. The hierarchy structural model of the index system.

Basing on the above thought, steps of the method of entropy right Markov chain are as follows:

Step1. Evaluate the marketing environmental risk of the present time period.

Fuzzy AHP can be used to evaluate the risk and the process is:

1) Establish the index system that is used to forewarn enterprises' marketing environmental risk. Let the index set which is used to evaluate marketing environmental risk is V={V_1, V_2, …, V_m}, V_i (i∈[1, m]) is an index of V. V_{ij}={V_{ij1}, V_{ij2}, …, V_{ijm} } is the index set of the jth index in V_i, and V_{ijk} is an index of V_{ij}. Then the hierarchy structural model can be established as that is shown in Fig. 1.

2) Ascertain the weight set of the index. According to AHP [13], the judge matrix of each level can be established through pairwise comparison, and then the weight set of the index can be established: ω_i= (ω_{i1}, ω_{i2}, …, ω_{in}), i∈[1, m] and

$$\sum_{j=1}^{n_i} \omega_{ij} = 1, \omega_{ij} \geq 0, i = 1, 2, \cdots, m.$$

3) Establish fuzzy evaluation matrix. Let fuzzy judgment set is U={U_1, U_2, …, U_n }, and U_j (j=1, 2, …, n) denotes the comments from best to worst, such as five grades comments of "best, better, good, common, bad" or nine grades comments and so on. According to the comments, a mark can be given and then fuzzy evaluation matrix of multi-index of each level can be established:

$$\tilde{R}_i = (r_{ij}).$$

4) Ascertain fuzzy synthetic evaluation vector. Carrying on synthetic calculation to fuzzy matrix, then fuzzy synthetic evaluation vector can be got:

$$\tilde{B} = \omega \cdot R = (\tilde{B}_1, \tilde{B}_2, \cdots, \tilde{B}_n),$$

$$\tilde{B}_i = \vee(\omega_i \wedge r_{ij}), j = 1, 2, \cdots n.$$

Where $\wedge$ denotes the minimum between ω_i and r_{ij}, $\vee$ denotes the maximum of ($\omega_i \wedge r_{ij}$).

5) Solve synthetic evaluation value of marketing environmental risk. Record C= [C_1, C_2… C_n] ^{T} as a mark set. C_j (j=1, 2… n) denotes the mark of the jth grade. Let a hundred-mark system is applied, and then calculate synthetic evaluation value of marketing environmental risk, as follows:

$$D = \tilde{B} \cdot C.$$

When synthetic evaluation value of the present time period is got, we bring it into the original synthetic evaluation value sequence.

Step2. Set the warning limit of different grades for marketing environmental risk.

We need to set different grades to synthetic evaluation value and that is decided by the experiences and the facts of enterprises' marketing environment. For example, enterprises' marketing environmental risk can be divided into five grades: double green light, green light, yellow light, red light, double red light and so on (corresponds to state space I={1, 2, 3, 4, 5} in Markov chain).

Step3. According to the standard of warning limit that is divided by Step 2, the risk state of marketing environment of each time period can be ascertained.

Step4. According to Eq. (3), the risk state of each time period is classified. Then transition probability matrixes of different step-sizes can be got (especially, step-size is m): $p^{(k)}$, k=1, 2, … , m.

Step5. Deduce state transition probability of the future time period.

We can choose the present time period and several time periods before (especially, m time periods in all), then take the state corresponds to certain time period as initial state, link to state transition probability matrixes that is got by Step 4, state transition probability of marketing environmental risk of the future time period can be deduced: $P_i^{(k)}$, i is the state, k is the step-size, k=1, 2, … , m.

Step6. Calculate entropy right.

The entropy value θ_i of different step-sizes Markov chain of the future time period can be calculated according to Eq. (4) and then θ_i is standardized:

$$\lambda_i = \frac{1-\theta_i}{m-\sum_{j=1}^{m}\theta_j}.$$

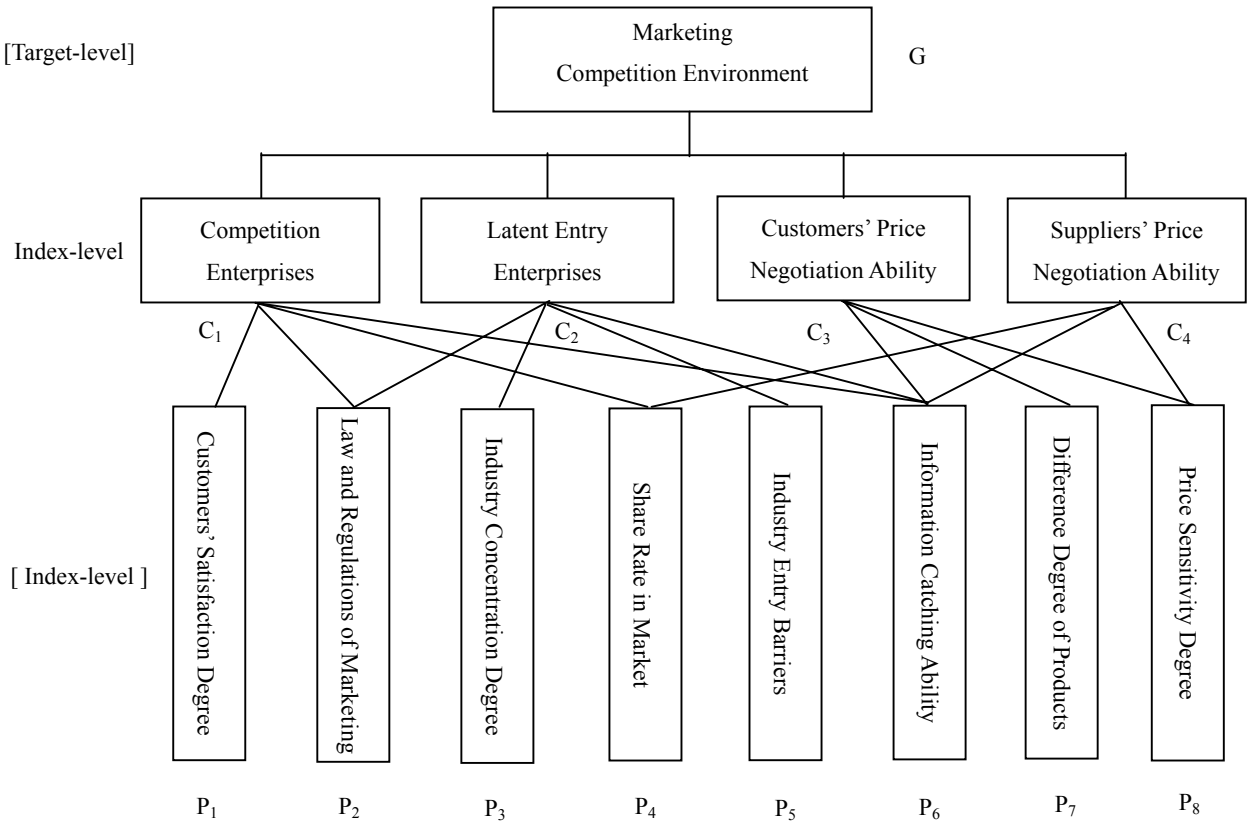


Fig. 2. The hierarchy of the index system for marketing competition environment.

Finally entropy right of different step-sizes Markov chain can be got, and

$$0 \le \lambda_i \le 1, \sum_{i=1}^{m} \lambda_i = 1 .$$

Step7. Forecast the state of the future time period (the risk state of marketing environment in the future).

Firstly we take the weight sum of the transition probability in the same state as the forecasting probability of that state, as follows:

$$P_i = \sum_{j=1}^{m} \lambda_j P_i^{(j)} .$$

Then the state i' which is corresponds to $\max\{P_i, i=1, 2, \cdots, m\}$ is the risk state of enterprises' marketing environment in the future.

4. THE ANALYSIS OF A CASE

The formation of enterprises' marketing environmental risk is a result affected by many kinds of factors. And the factors not only come from outside of enterprises, but also come from inside of enterprises. In this paper, an enterprise's marketing competition environment is taken as a case for simple. According to four important indexes as competition enterprises, latent entry enterprises, customers' price negotiation ability and suppliers' price negotiation ability;

then eight sub-factors are taken to form an evaluation index system. The hierarchy structural model can be established as that is shown in Fig. 2.

Table 1. The table of the risk state for an Enterprise's Marketing Competition Environment

Time (Month)	0209	0210	0211	0212	0301	0302	0303	0304	0305	0306	0307
Evaluation Value	23.5	34.2	67.8	13.5	61.2	35.9	19.8	28.7	43.7	57.3	52.6
State	2	2	4	1	4	2	1	2	3	3	3
Time (Month)	0308	0309	0310	0311	0312	0401	0402	0403	0404	0405	0406
Evaluation Value	73.1	66.2	28.4	61.9	29.3	35.6	54.7	58.8	81.4	90.6	78.6
State	4	4	2	4	2	2	3	3	5	5	4
Time (Month)	0407	0408	0409	0410	0411	0412	0501	0502	0503	0504	0505
Evaluation Value	61.3	19.4	12.5	43.2	58.3	83.7	13.7	85.2	74.7	56.5	92.7
State	4	1	1	3	3	5	1	5	4	3	5
Time (Month)	0506	0507	0508	0509	0510	0511	0512				
Evaluation Value	58.9	38.6	56.9	17.6	40.2	34.6	46.8				
State	3	2	3	1	3	2	3				

Annotate: Take two numbers of the year and the month

1) Evaluate marketing environmental risk of the present time period:

① Ascertain the weight of the index. The essentiality scale method that is quoted by T.L.Saaty is applied, and the judge matrix of each level can be established through pairwise comparison, then the weight of the index can be got:

C_1-P level: w_1= (0.41, 0.41, 0. 09, 0.09);

C_2-P level: w_2= (0.10, 0.11, 0. 26, 0.53);

C_3-P level: w_3= (0.14, 0.14, 0.72);

C_4-P level: w_4= (0.43, 0.14, 0.43);

G-C level: w= (0.54, 0.23, 0.09, 0.14).

② Establish fuzzy evaluation matrix. According to five grades comments of "best, better, good, common, bad", four fuzzy evaluation matrixes can be established in response to four important indexes as competition enterprises, latent entry enterprises, customers' price negotiation ability and suppliers' price negotiation ability.

$$\tilde{R}_1 = \begin{bmatrix} 0 & 0 & 0.3 & 0.3 & 0.4 \\ 0 & 0.1 & 0.5 & 0.3 & 0.1 \\ 0 & 0.2 & 0.2 & 0.4 & 0.2 \\ 0 & 0 & 0.6 & 0.4 & 0 \end{bmatrix};$$

$$\tilde{R}_2 = \begin{bmatrix} 0 & 0.1 & 0.2 & 0.5 & 0.2 \\ 0 & 0 & 0.5 & 0.4 & 0.1 \\ 0 & 0.2 & 0.4 & 0.4 & 0 \\ 0 & 0.2 & 0.3 & 0.5 & 0 \end{bmatrix};$$

$$\tilde{R}_3 = \begin{bmatrix} 0 & 0.2 & 0.7 & 0.1 & 0 \\ 0 & 0 & 0.6 & 0.3 & 0.1 \\ 0 & 0 & 0.1 & 0.7 & 0.2 \end{bmatrix};$$

$$\tilde{R}_4 = \begin{bmatrix} 0 & 0 & 0.6 & 0.3 & 0.1 \\ 0 & 0 & 0.1 & 0.8 & 0.1 \\ 0 & 0.3 & 0.6 & 0.1 & 0 \end{bmatrix}.$$

③ Ascertain fuzzy synthetic evaluation vector. Link to the weight of the index that is ascertained by ①, synthetic calculation is taken to four fuzzy matrixes. Then a fuzzy synthetic evaluation matrix can be got:

$$\tilde{R}_1 = \begin{bmatrix} 0 & 0.09 & 0.33 & 0.25 & 0.33 \\ 0 & 0.18 & 0.27 & 0.45 & 0.10 \\ 0 & 0.12 & 0.12 & 0.59 & 0.17 \\ 0 & 0.27 & 0.38 & 0.27 & 0.08 \end{bmatrix}.$$

Then according to

$$\tilde{B} = \omega_1 \cdot \tilde{R}_1,$$

fuzzy synthetic evaluation vector can be got:

$$\tilde{B} = (0, 0.18, 0.33, 0.25, 0.33),$$

standardize:

$$\tilde{B}' = (0, 0.17, 0.30, 0.23, 0.30).$$

④ Calculate synthetic evaluation value of marketing environmental risk of the present time period. We may take C=(100, 80, 60, 40, 20)^T, and then D=46.8 is got according to $D = \tilde{B} \cdot C$. So evaluation value of marketing competition environmental risk of the present time period is 46.8. Table 1. is the evaluation value sequence of an enterprise by monthly tracking evaluation.

2) Set the warning limit of different grades for marketing environmental risk. According to the facts of the market, the risk state of marketing competition environment is divided into five grades of warning limit and that corresponds to five states in Markov chain, as that is shown in Table 2.:

Table 2. The division of the sisk state

Range	State	Warning Limit
[0, 20]	1	Double Red Light Risk
[20, 40]	2	Red Light Risk
[40, 60]	3	Yellow Light Risk
[60, 80]	4	Green Light Risk
[80, 100]	5	Double Green Light Risk

3) According to the warning limit that is divided by Table 2., the risk state of marketing competition environment of each time period can be ascertained, as that is shown in Table 1.

4) According to Eq. (3), the state transition probability matrixes of different step-sizes Markov chain can be got:

Table 3. The table of the forecasting risk state for an Enterprise's Marketing Environment in January, 2006

Initial Time	Initial State	Step-size	Weight	State 1	State 2	State 3	State 4	State 5
0512	3	1	0.2603	1/11	2/11	4/11	1/11	3/11
0511	2	2	0.1785	2/8	2/8	3/8	1/8	0
0510	3	3	0.2037	2/10	3/10	0	3/10	2/10
0509	1	4	0.1847	1/5	1/5	1/5	0	2/5
0508	3	5	0.1728	2/9	3/9	1/9	3/9	0
	Weight Sum P_i			0.1844	0.2476	0.2177	0.1647	0.1856

$$P(1)=\begin{bmatrix} 1/6 & 1/6 & 2/6 & 1/6 & 1/6 \\ 1/9 & 3/9 & 3/9 & 2/9 & 0 \\ 1/11 & 2/11 & 4/11 & 1/11 & 3/11 \\ 2/8 & 3/8 & 1/8 & 2/8 & 0 \\ 1/5 & 0 & 1/5 & 2/5 & 1/5 \end{bmatrix};$$

$$P(2)=\begin{bmatrix} 0 & 2/6 & 3/6 & 1/6 & 0 \\ 2/8 & 2/8 & 3/8 & 1/8 & 0 \\ 1/11 & 1/11 & 4/11 & 2/11 & 3/11 \\ 3/8 & 2/8 & 0 & 2/8 & 1/8 \\ 0 & 1/5 & 1/5 & 2/5 & 1/5 \end{bmatrix};$$

$$P(3)=\begin{bmatrix} 1/6 & 1/6 & 3/6 & 0 & 1/6 \\ 1/8 & 1/8 & 4/8 & 1/8 & 1/8 \\ 2/10 & 3/10 & 0 & 3/10 & 2/10 \\ 1/8 & 3/8 & 3/8 & 1/8 & 0 \\ 1/5 & 0 & 1/5 & 2/5 & 1/5 \end{bmatrix};$$

$$P(4)=\begin{bmatrix} 1/5 & 1/5 & 1/5 & 0 & 2/5 \\ 0 & 2/8 & 2/8 & 2/8 & 2/8 \\ 0 & 2/10 & 2/10 & 5/10 & 1/10 \\ 1/8 & 3/8 & 4/8 & 0 & 0 \\ 3/5 & 0 & 2/5 & 0 & 0 \end{bmatrix};$$

$$P(5)=\begin{bmatrix} 1/5 & 0 & 2/5 & 1/5 & 1/5 \\ 1/8 & 2/8 & 2/8 & 2/8 & 1/8 \\ 2/9 & 3/9 & 1/9 & 3/9 & 0 \\ 0 & 2/8 & 4/8 & 0 & 2/8 \\ 1/5 & 1/5 & 2/5 & 0 & 1/5 \end{bmatrix}.$$

5) Five states which are correspond to the time periods from August to December in 2005 are chosen as initial state, then state transition probability of marketing competition environmental risk in January, 2006 can be deduced, as that is shown in Table 3.

6) Calculate entropy right. According to Eq. (4), entropy value of each state of different step-sizes Markov chain in Table 3. can be got:

θ_1=1.4681, θ_2=1.3209, θ_3=1.3662, θ_4=1.3322, θ_5=1.3108.

Then θ_i is standardized, so entropy right is

λ= (0.2603, 0.1785, 0.2037, 0.1847, 0.1728).

In Table 3., $\max\{P_i\}$ is 0.2476 and the state is i=2. Therefore, the forecasting risk is red light risk in January, 2006.

5. CONCLUSIONS

Enterprises' marketing environmental risk arises from effect of many kinds of factors, thus it has the multiple uncertainty characteristics such as fuzziness, randomness and so on. Therefore the forewarning analysis is different. Compare to other research methods, the method that is proposed in the paper has two significances: Firstly the analysis technology of Markov chain in random process is combined with information entropy. Different step-sizes entropy right Markov chain is used to forecast the system state, so the useful information that is provided by raw data can be fully utilized and then the precision and rationality of Markov chain method is improved; secondly not only the fuzziness of certain time period can be described in quantitatively, but also the dynamic characteristics of marketing environmental risk can be reflected. Then the research can be deepened.

6. REFERENCES

[1] F. Luo, L. She, "The Weight and Synthetic Evaluation of the Evaluation Index for Enterprises' Organization Management Forewarning System", China Soft Science, No.7, 2000, 115-117.

[2] F.Y. Gao, J.H. Zhao, "The Study of the Index System of Enterprises' Marketing Organization Forewarning Management", Journal of Hebei University of Science and Technology, 1999, 20（2）：83-86.

[3] She Lian, The Theory of Enterprises' Forewarning Management, Shijiazhuang: Hebei Science and Technology Press, 1999.

[4] T. Alder, J.G. Leonard, "Improving Risk Management: Moving from Risk Elimination to Risk Avoidance", Information and Software Technology, No.8, 2004, 231-236.

[5] K.F. Xie, X.F. Luo, Forewarning Management of Market Risk, Shijiazhuang: Hebei Science and Technology Press, 2004.

[6] F.Y. Gao, L. She, "Risk Fuzzy Evaluation and Forewarning Management of Enterprises' Marketing Organization", Journal of Wuhan Transportation University, 2000, 24（2）: 125-128.

[7] K.Y. Chen, L.P. Zeng, "The Marketing Venture Precaution for Enterprises Based on Artificial Nerve Network", Journal of Chongqing University of Posts and Telecommunications, 2004, 63（5）: 139-142.

[8] H.B. Liu, X.S. Chen et al, "Application of Stochastic Process in Disastrous Forecasting", Systems Engineering Theory and Practice, 1996, 9（9）: 19~23.

[9] H.J. Li, Y. Ma, et al, "Markov Chain and Exchange

Rate Forecast of Japanese Yen",

[10] Journal of Qingdao University, 2002, 17（2）: 4-7.
[11] M.H.A. Davis, Markov Models and Optimization, Chapman and Hall, 1993.
[12] Y.L. Lin, Applicable Random Process, Beijing: Tsinghua University Press, 2002.
[13] W.H. Qiu, Management Decision and Practical Entropy Science, Beijing: Mechanical Industry Press, 2002.
[14] S.H. Xu, Principle of AHP, Tianjin: Tianjin University Press, 1986.

Supply Chain Multi-Product and Multi-Objective Bi-level Programming Under Fuzzy and Grey Uncertainty *

Dongbo Liu [1,2]
[1] **Research Institute of Automation, East China University of Science &Technology**
Shanghai, 200237, China
[2] **College of Mechanical and Electronic Engineering, Shanghai Normal University**
Shanghai, 201418, China
Email: liudbyj@163.com
Dao Huang
Research Institute of Automation, East China University of Science &Technology
Shanghai, 200237, China
Email: dhuang@ecust.edu.cn
Yujuan Chen
College of Mechanical and Electronic Engineering, Shanghai Normal University
Shanghai, 201418, China
Email: chenyjfq@163.com

ABSTRACT

The uncertain parameters in supply chain have different distribution styles. Some uncertain parameters containing fuzzy and grey twofold uncertain factors can be described with fuzzy grey variables. The supply chain's planning model considering multi-product, multi-objective, and multi-factor constraints was presented under fuzzy and grey uncertainty. The first-level master model corresponds to optimizing the manufacturer's profit under certain capital constraint. The second-level slave models correspond to simultaneously minimizing the cost for every type of supplied material. The proposed fuzzy grey simulation and grey simulation technology can generate input-output data to approximate the uncertain functions based on the credibility and chance measure of fuzzy grey variables. The hybrid optimal algorithm integrated the fuzzy grey simulation, grey simulation, and random simulation with bi-level genetic algorithm. This algorithm can optimize the uncertain programming problems. One numerical example is to illustrate the effectiveness of the model and algorithm.

Keywords: Supply Chain, Multi-product, Multi-objective, Bi-level Uncertain Programming, Fuzzy Grey Simulation, Genetic Algorithm.

1. INTRODUCTION

Many parameters in supply chain are uncertain. Customer demand, replenishment lead-time and material prices are important uncertain factors affecting production-planning problems. The bullwhip effect and moral hazard make the transferred information in supply chain becoming much more uncertain [1,2]. These uncertain parameters should be taken special considerations in order to make the uncertain programming more realistic.

Generally speaking, uncertain programming includes stochastic programming and fuzzy programming [3,4]. The uncertain variables with less statistic data can be described with a grey variable [5,6] $\otimes \in [a_1, a2]$, and can be handled with grey programming [5]. The basic method of present grey programming is to transform the uncertain grey programming into deterministic programming that can't accurately reflect the real situations and can't effectively solve the complicated grey programming problems.

The uncertain parameters with twofold uncertainty combining grey and fuzziness factors such as customer demand and product revenue can be denoted with a fuzzy grey variable. This paper tries to present bi-level uncertain programming model about the supply chain production-planning problem. The hybrid optimal algorithm integrating the fuzzy grey simulation technology with genetic algorithm (GA) can give the optimal solution for the model.

2. FUZZY GREY VARIABLE

Some uncertain parameters may be fuzzy. Sometimes, it is difficult to give the exact value for the key parameters of a fuzzy variable. We only know their general scope following some kind of distribution. If the key parameters of a fuzzy variable are grey, the fuzzy variable can be called as a fuzzy grey variable.

The references [3,4,7] have studied many properties of random variables and fuzzy variables. On these bases, we give the definition and properties of fuzzy grey variable.

Definition 1. If $\tilde{\xi}$ is the function from a possibility space (Ω, A, Pr) to a fuzzy variable set, $\tilde{\xi}$ is called as a fuzzy grey variable. Where Ω is a non-vacant set, A is the subset (event) defined on Ω, and Pr is the possibility measure defined on A. $\tilde{\xi}$ is characterized by a possibility distribution function μ of the fuzzy vector.

Example 1. Suppose a triangle fuzzy variable $\tilde{\xi} = (a(\otimes)-5, a(\otimes), a(\otimes)+10)$, where $a(\otimes) = (\otimes_1, \otimes_2)$ is a grey variable defined on the possibility space Ω. The grey property of variable $\tilde{\xi}$ is determined by grey variable $a(\otimes)$.

The possibility, necessity, and credibility of a fuzzy event $\{ f(\tilde{\xi}(\otimes)) \geq r \}$ can be defined as follows.

$$\mathrm{Pos}\{f(\tilde{\xi}(\otimes)) \geq r\} = \sup_{u \geq r} \mu(u), \mathrm{Nec}\{f(\tilde{\xi}(\otimes)) \geq r\} = 1 - \sup_{u < r} \mu(u)$$
$$\mathrm{Cr}\{f(\tilde{\xi}(\otimes)) \geq r\} = 0.5[\mathrm{Pos}\{f(\tilde{\xi}(\otimes)) \geq r\} + \mathrm{Nec}\{f(\tilde{\xi}(\otimes)) \geq r\}] \quad (1)$$

Where, μ is the membership function $f(\tilde{\xi}(\otimes))$.

* Partially supported by the Science &Research Foundation of Shanghai Educational Committee (No. 05DZ33)

Definition 2. Suppose $f(\tilde{\xi}(\otimes))$ is n-dimensional fuzzy grey vector defined on possibility space Ω, and $f:\Re^n \to \Re^m$ is a real-valued measurable function. The primitive chance of fuzzy grey event $f(\tilde{\xi}(\otimes)) \le 0$ is a function from [0,1] to [0,1], defined as

$$\mathrm{Ch}\{f(\tilde{\xi}(\otimes)) \le 0\}(\alpha) = \sup\{\beta \mid \Pr\{\theta \in \Omega \mid \mathrm{Cr}\{f(\tilde{\xi}(\otimes(\theta))) \le 0\} \ge \beta\} \ge \alpha\}. \quad (2)$$

The primitive chance represents the fuzzy grey event holds with probability $\mathrm{Ch}\{f(\tilde{\xi}(\otimes)) \le 0\}(\alpha)$ at possibility α.

Remark 1. If the fuzzy grey vector n becomes a grey vector, we have $\mathrm{Ch}\{f(\tilde{\xi}(\otimes)) \le 0\}(\alpha) = \Pr\{f(\otimes(\theta)) \le 0\}$.

Remark 2. If the fuzzy grey vector n becomes a fuzzy vector, then we have

$$\mathrm{Ch}\{f(\tilde{\xi}(\otimes)) \le 0\}(\alpha) = \begin{cases} 1, & \mathrm{Pos}\{f(\tilde{\xi}(\otimes(\theta)) \le 0\} \ge \beta. \\ 0, & \text{otherwise.} \end{cases}$$

3. BI-LEVEL PRODUCTION PLANNING MODEL

Sharing the inventory and production information between the manufacturer and material suppliers can effectively reduce the side effects of uncertain factors [8]. The Bi-level programming programs have been applied in extensive and diverse areas [9]. According to the historical data and demand predication, the manufacturer can establish the first-level master production- planning model, and can further make material order decisions in order to meet the customer demand as much as possible. Suppose one manufacture produces m types of products to distribution centers, and orders n types of materials from p suppliers. The suppliers and manufacturer take the transportation cost for materials and products respectively. The bi-level production-planning model is shown in Fig. 1.

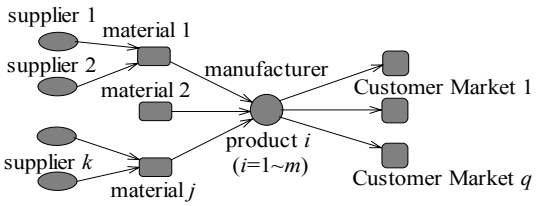


Fig. 1. Bi-level production planning model

Notations for the planning model are as follows.

x_i, x_{ih} and x_{ic} are the production amount, safety holding inventory level, and customer shortage amount for product i. y_j and y_{jm} are the ordered material amount and material safety holding inventory level for material j. y_{jk} and a_{ik} are the supplied material amount and unit material price for material j from supplier k. h_{jm} ,y_{jmL} and y_{jmH} are the manufacturer's unit material safety holding inventory level, the lower and upper bound of safety holding inventory for material j. c_i, r_i, t_i, h_i and s_i are respectively the unit product cost, unit revenue, unit transportation cost, unit inventory cost, unit shortage cost for product i. b_{ji} is the necessary unit coefficient of material j to produce product i. y_{jk0} is the minimal order amount of material j from supplier k. y_{jkH} is the maximal production capacity of material j for supplier k. x_{icH} is the maximal shortage limitation of product i for supplier k. $\tilde{D}_{it}(\otimes)$ and $\tilde{D}_{i(t+1)}(\otimes)$ are the customer demand prediction at prediction period t and t+1 for product i of the manufacturer.

3.1 Manufacturer first-level master model

The manufacturer's main objective function is to make the total cost lower than the maximal usable capital Z_0 under fuzzy and grey uncertainty.

$$\max_{x_i, x_{ih}, x_{ic}, y_{jk}, y_{jm}} \Pr\{\sum_{i=1}^{m} c_i(\xi)x_i + \sum_{i=1}^{m} t_i(\xi)x_i + \sum_{i=1}^{m} h_i x_{ih} + \sum_{i=1}^{m} s_i x_{ic} + \sum_{j=1}^{n}\sum_{k=1}^{p} \tilde{a}_{jk}(\otimes) y_{jk} + \sum_{j=1}^{n} h_{jm} y_{jm} \le Z_0\} \ge \alpha_1. \quad (3)$$

Where $\Pr\{\bullet\}$ denotes the probability measure of the accident in $\{\bullet\}$. α_1 is the given probability level of the objective function. $c_i(\xi)$ and $t_i(\xi)$ are random variables.

The first part is the variable production cost within the manufacturer. It may include the employee's salary, development cost, maintenance cost and so on. The material cost isn't included in this part. The second part is the transportation cost from the manufacturer to all the distribution centers and direct customers. The third part is the inventory holding cost for all products of the manufacturer. The fourth part and sixth part are respectively customer demand shortage cost and material holding cost.

The fifth part is the material cost to produce all the products. Among all the material suppliers, because of the difference of technology, management level, transportation method, and transportation distance between the suppliers and manufacturer, the material price provided by different suppliers usually differs from each other. The parameter $a_{jk}(\xi)$ can be considered as random variables.

Besides of the cost control objective, the manufacturer may have another objective to maximize the probability that the total revenue is larger than predicted revenue π_0. The unit product revenue $\tilde{r}_i(\otimes)$ is considered as a fuzzy grey variable.

$$\max_{x_i} \quad \mathrm{Ch}\{\sum_{i=1}^{m} r_i(\otimes)x_i \ge \pi_0\}(\alpha_2) \ge \beta_2. \quad (4)$$

Where, α_2 and β_2 are the given probability level and the credibility level of the objective function.

The objective functions of manufacturer are typically subject to the following constraints.

The manufacturer's production level for product i should be lower than the production capacity.

$$x_i \le P_i, \quad i = 1,...,m. \quad (5)$$

The production amount of product i should be larger than the customer's minimal demand. The uncertain customer demand can be considered as a fuzzy grey parameter, and can be estimated the demand according to customers' order and history empirical data. ω_{it} and $\omega_{i(t+1)}$ are the estimation weights of the customer demand $\tilde{D}_{it}(\otimes)$ and $\tilde{D}_{i(t+1)}(\otimes)$ at prediction period t and t+1. Then, the total estimated demand constraint from estimation period t to T is described as follow.

$$\mathrm{Ch}\{\sum_{t=0}^{T} \omega_{it}\tilde{D}_{it}(\otimes) \le x_i\}(\gamma_i) \ge \delta_i, x_i \le P_i, \forall i. \quad (6)$$

Generally speaking, we have the following conclusion: $\omega_{it} \ge \omega_{i(t+1)} \ge \omega_{i(t+2)}$. The longer the estimation period is, the worse the accuracy is, and the smaller the weight is.

Certain safety stock level of product i is necessary in order to improve the customer service level.

$$x_{ihL} \le x_{ih} \le x_{ihH}, \ i = 1,...,m. \quad (7)$$

Customer shortage level should be lower than a given upper bound to avoid too much shortage cost.

$$0 \le x_{ic} \le x_{icH},\ i = 1,..., m. \tag{8}$$

Every type of supplied material should maintain certain inventory level.

$$y_{jmL} \le y_j \le x_{jmH},\ j = 1,..., n. \tag{9}$$

3.2 Supplier second-level slave models

The suppliers provide materials according to the optimal decision of the manufacture. However, the suppliers' supply capacity has certain uncertainty because of the uncertainty of their production capacity, material price, and material replenishment lead-time.

Suppose the manufacturer requires *n* types of materials. The total amount of material *j* to produce *m* types of products for the manufacturer is as follow.

$$y_j = \sum_{i=1}^{m} b_{ji} x_i,\ j = 1,2,..., n. \tag{10}$$

Under the supply chain circumstances, the material cost of different suppliers should be taken as the second-level slave objective in order to minimize the manufacturer's total cost. Usually, the material prices from different suppliers are different because of the factors such as production cost, management level, transportation distance, and transportation vehicles.

One type of material may be supplied from more than one supplier. The cost for every kind of material can be minimized with a second-level slave model. Therefore, we have *j* second-level slave models totally.

$$\begin{aligned} &\min_{y_{jk}} Z_j \\ &\text{s.t. } \Pr\{\sum_{k=1}^{p} a_{jk}(\xi) y_{jk} \le Z_j\} \ge \gamma_i, j = 1,.., n. \end{aligned} \tag{11}$$

Where, Z_j is the minimal cost of material *j* under γ_i probability. The parameter $a(\xi)$ is a random variable.

Acquiring enough materials from the suppliers is one of the important factors for the manufacturer to ensure the normal production and to maximize the revenue. The manufacturer can order one type of material from more than one supplier to avoid material shortage hazard. The supplied material amount should equal with the amount required by manufacturer.

$$y_j = \sum_{k=1}^{p} y_{jk} = \sum_{i=1}^{m} b_{ji} x_i, j = 1,..., n. \tag{12}$$

One material supplier may simultaneously take part in several different supply chains. The material supply capacity can be considered as an uncertain grey variable.

$$y_{jk0} \le y_{jk} \le y_{jkH}(\otimes), j = 1,..., n; k = 1,..., p. \tag{13}$$

The supply chain bi-level production planning decision model including one first-level master model and n second-level slave models can be formulated as follow.

$$\begin{cases} \text{lexmin}\{d_1^-, d_2^-\} \\ \text{s.t. } \Pr\{\sum_{i=1}^{m} c_{\mathrm{i}}(\xi) x_i + \sum_{i=1}^{m} t_{\mathrm{i}}(\xi) x_i + \sum_{i=1}^{m} h_{\mathrm{i}} x_{ih} + \sum_{i=1}^{m} s_{\mathrm{i}} x_{ic} + \\ \sum_{j=1}^{n}\sum_{k=1}^{p} a_{\mathrm{jk}}(\xi) y_{jk} + \sum_{j=1}^{n} h_{\mathrm{jm}} y_{jm} \le Z_0\} + d_1^- - d_1^+ = \alpha_1 \\ \mathrm{Ch}\{\sum_{i=1}^{m} \widetilde{r}_i(\otimes) x_i \ge \pi_0\}(\alpha_2) + d_2^- - d_2^+ = \beta_2 \\ \mathrm{Ch}\{\sum_{t=0}^{T} \omega_{it} \widetilde{D}_{it}(\otimes) \le x_i\}(\gamma_i) \ge \delta_i,\ x_i \le P_i,\ \forall i \\ x_{ihL} \le x_{ih} \le x_{ihH},\ 0 \le x_{ic} \le x_{icH},\ \forall i \\ y_{jmL} \le y_{jm} \le y_{jmH},\ \forall j \\ \begin{cases} \min_{y_{jk}} Z_j \\ \text{s.t. } \Pr\{\sum_{k=1}^{p} a_{jk}(\xi) y_{jk} \le Z_j\} \ge \gamma_i, \forall j \\ y_j = \sum_{k=1}^{p} y_{jk} = \sum_{i=1}^{m} b_{ji} x_i,\ \forall j \\ y_{jk0} \le y_{jk} \le y_{jkH}(\otimes),\ \forall j, k. \end{cases} \end{cases} \tag{14}$$

4. FUZZY GREY SIMULATION

Fuzzy grey simulation is a technology to perform sampling experimentation from the fuzzy grey system according to the possibility distribution function of fuzzy grey variable and the whitenization weight function of grey variable. The whitenization weight function is used to denote the degree of preference for different data whited from the grey data. Every sampling from the fuzzy grey simulation can be considered as a process of stochastic whtienization. According to the definition of whitenization weight function [5], the appearing possibility of the grey data with higher whitenization weight value is higher than the grey data with lower whitenization weight value.

Consider a fuzzy grey variable $\widetilde{\xi} = \{a(\otimes), b(\otimes), c(\otimes)\}$, $a(\otimes), b(\otimes)$ and $c(\otimes)$ are grey variables. The whitenization weight function $\phi(x)$ of the grey variable is a nonnegative, continuous, and integrable function. We can make the following formulation.

$$f(x) = \phi(x) / \int_{\tilde{a}}^{\tilde{b}} \phi(x)dx,\ F(x) = \int_{\tilde{a}}^{x} \phi(x)dx / \int_{\tilde{a}}^{\tilde{b}} \phi(x)dx. \tag{15}$$

$f(x)$ and $F(x)$ can be considered as the probability density function and distribution function of grey variable *x*. Based on [3], we know that if $F(x)$ is the probability function of the grey variable *x*, we can first get the random variable μ generated from the equality distribution between 0 and 1, and then can get a random data $\overline{\otimes} = F^{-1}(\mu)$ followed the distribution function $F(x)$.

The function $\phi(x) = 1/(\alpha(x - a_0)^2 + 1)$ is taken as the whitenization weight function of grey variables. The coefficient α denotes the concentration degree of a whitenization value within its scope. The bigger the coefficient α is, the higher the concentration degree of grey data distributes. After inputting the whitenization weight value of grey variable into the fuzzy grey variable, the estimation value of fuzzy grey variable under certain credibility level can be obtained.

Consider one type of fuzzy grey simulation for the constraint containing uncertain factors.

$$U_1 : x \to \mathrm{Ch}\{f(x, \widetilde{\xi}(\otimes)) \le 0\}(\alpha) + d^- - d^+ = \beta$$

The objective of this function is to calculate minimal d^- under the certain credibility β and certain probability α. The steps of fuzzy grey simulation are as follows.

Step 1. Generate the whitenization weight value $\{\otimes_1 ... \otimes_N\}$ from the possibility space according to the whitenization weight function of grey variables.

Step 2. On the basis of the definition of credibility in

Eq.(1), calculate $d_k^- = \mathrm{Cr}\{f(\tilde{\xi}(\otimes_k) \le 0\}, k = 1,...,N,$ by fuzzy simulation.

Step 3. Set *N'* as the integer part of (1-α) *N*.

Step 4. Return the *N'*th largest element in sequence $\{d_1^-,..., d_N^-\}$.

5. BI-LEVEL GENETIC ALGORITHM BASED ON FUZZY GREY SIMULATION

The chance constraint programming problems containing fuzzy, grey and random parameters can be optimized by random simulation [3], fuzzy grey simulation and genetic algorithm. The simulation technology can check the feasibility of optimal results given by GA, and calculate the values of the objective function and constraints. The genetic algorithm can give the final optimal solution.

This bi-level fuzzy grey hybrid-programming model includes one first-level master model and *n* second-level slave models. The first-level master model takes *x* and *y* as decision variables. The second-level slave models take *y* as decision variable, and take *x* as constant parameter. There is one type of common decision variable *y* between the first-level master model and the second-level slave models. It can't be solved with usual genetic algorithm.

A bi-level genetic algorithm based on fuzzy grey simulation and random simulation is presented. Firstly, the first-level maser model separately implements genetic algorithm. Its optimal feasible decision vector x^* is transferred to second-level slave models. Secondly, the second-level slave models implement genetic algorithm respectively according to their own constraints, and return their optimal solutions to the first-level master model. The first-level master model re-optimize again until reaching the given iteration number. The optimal steps are as follows.

Step 1. Initiate the pop size, iteration number, crossover probability P_c, mutation probability P_m for first-level master model and second level slave models.

Step 2. Initiate the pop_size1 chromosomes for first-level master model. The random simulation and fuzzy grey simulation can check the feasibility of chromosomes.

Step 3. The *n* second-level slave models receive the feasible chromosomes of first-level master model and perform GA optimization respectively.

Step 3.1. The *n* second-level slave models respectively initiate their own pop_size2 chromosomes. The random simulation and grey simulation can check the feasibility of their own initial chromosomes.

Step 3.2. The slave models perform crossover and mutation operations. The random simulation and grey simulation can check the feasibility of child chromosomes.

Step 3.3. Calculate the objective values for all the chromosomes of second-level slave models by the random simulation and grey simulation.

Step 3.4. Compute the fitness of all chromosomes of second-level slave models according to their objective values by the rank-based evaluation function.

Step 3.5. Select the chromosomes by spinning the roulette wheel.

Step 3.6. Repeat from step 3.2 to step 3.5 until finishing the given iteration number for the n second-level models. The best chromosome is respectively considered as optimal solution of second level slave models.

Step 4. The best chromosomes of all the second-level slave models are transferred back to the first-level master model. Calculate the objective value of first-level master model. Compute the fitness of chromosomes of first-level master model with the rank-based evaluation function.

Step 5. Select the chromosomes of the first-level master model by spinning the roulette wheel.

Step 6. The chromosomes of first-level master model perform crossover and mutation operations. The feasibility of child chromosomes is checked with fuzzy grey simulation and random simulation.

Step 7. Repeat from step 3 to step 6 until reaching the given iteration number of the first–level master model.

Step 8. Obtain the best chromosome of first–level master model as the optimal solution.

6. NUMERICAL EXAMPLE

One electronic (Shanghai) Ltd. Co. is an international group producing Liquid Crystal Display (LCD) television set. Take its 20", 32" LCD products as example. The main materials include 20" LCD (y_1), 32" LCD (y_2), main electronic board (y_3), and power inverter (y_4).

The unit product cost and unit transportation cost are random variables. $\xi_{c1} \in \mathrm{N}(60,5^2)$, $\xi_{c2} \in \mathrm{N}(80,8^2)$, $\xi_{t1} \in \mathrm{N}(3.2,1.2^2)$, $\xi_{t2} \in \mathrm{N}(4,1.4^2)$. h_1=2.2, h_2 =2.8, s_1=20, s_2=28, h_{1m}=1.5, h_{2m}=1.9, h_{3m}=0.8, h_{4m}=0.6, P_1=P_2=30000, $3000 \le x_{1h} = x_{2h} \le 5000$, $x_{1c} = x_{2c} \le 1500$, $y_{1m} = y_{2m} = y_{3m} = y_{4m}$= 4000, b_{jt}=1, Z_0=110000000, π_0=19000000.

The uncertain customer demand of one month can be predicted. Customer demand and unit product revenue are fuzzy grey variables.

$\widetilde{D}_{1t}(\otimes) = (\otimes_{D1t} - 1800, \otimes_{D1t}, \otimes_{D1t} + 1600)$, $\otimes_{D1t}$=[5000,15000]

$\widetilde{D}_{2t}(\otimes) = (\otimes_{D2t} - 2100, \otimes_{D2t}, \otimes_{D2t} + 1800)$, $\otimes_{D2t}$=[8000,19000]

$\widetilde{r}_1(\otimes) = (\otimes_{r1} - 90, \otimes_{r1}, \otimes_{r1} + 180)$, $\otimes_{r1}$ = [150,340]

$\widetilde{r}_2(\otimes) = (\otimes_{r2} - 180, \otimes_{r2}, \otimes_{r2} + 310)$, $\otimes_{r2}$ = [550,1150]

$$\phi_{1t}(x) = 1/(1 + 0.002(x - 11000)^2)$$

$$\phi_{2t}(x) = 1/(1 + 0.003(x - 15000)^2)$$

$$\phi_{1r}(x) = 1/(1 + 0.008(x - 280)^2)$$

$$\phi_{2r}(x) = 1/(1 + 0.007(x - 820)^2)$$

There are five suppliers capable of supplying the four types of materials. The material prices are random variables. Their supply capabilities are grey variables listed in Table 1. Where Sup.1 in table 1 denotes the material supplier 1.

Table 1. Suppliers' supply capability

	Sup.1	Sup.2	Sup.3	Sup.4	Sup.5
y_1	$\otimes_{y11}$	17000		20000	
y_2	13000	$\otimes_{y22}$		$\otimes_{y24}$	
y_3		32000	$\otimes_{y33}$		$\otimes_{y35}$
y_4	$\otimes_{y41}$		44000		$\otimes_{y45}$

$\xi_{a11} \in \mathrm{N}(850,40^2)$, $\xi_{a12} \in \mathrm{N}(810,60^2)$, $\xi_{a14} \in \mathrm{N}(890,70^2)$, $\xi_{a21} \in \mathrm{N}(3500,70^2)$, $\xi_{a22} \in \mathrm{N}(3600,60^2)$, $\xi_{a24} \in \mathrm{N}(3620,80^2)$, $\xi_{a32} \in \mathrm{N}(380,60^2)$, $\xi_{a33} \in \mathrm{N}(395, 50^2)$, $\xi_{a35} \in \mathrm{N}(410,55^2)$, $\xi_{a41} \in \mathrm{N}(215, 40^2)$, $\xi_{a43} \in \mathrm{N}(230, 55^2)$, $\xi_{a45} \in \mathrm{N}(240,60^2)$

$$\phi_{11}(x) = 1/(1 + 0.004(x - 14500)^2),$$

$\otimes_{y11} \in [10000,19000]$.

$\phi_{22}(x) = 1/(1 + 0.0045(x - 12000)^2)$. $\otimes_{y22} \in$

[9000,17000]

$\phi_{24}(x) = 1/(1+0.0055(x-21000)^2)$,

$\otimes_{y24} \in [19000,23000]$.

$\phi_{33}(x) = 1/(1+0.006(x-39000)^2)$, $\otimes_{y33} \in [32000,42000]$

$\phi_{35}(x) = 1/(1+0.007(x-46000)^2)$, $\otimes_{y35} \in [44000,53000]$

$\phi_{41}(x) = 1/(1+0.006(x-32000)^2)$, $\otimes_{y41} \in [22000, 36000]$

$\phi_{45}(x) = 1/(1+0.008(x-42000)^2)$, $\otimes_{y45} \in [36000,45000]$

The bi-level production-planning model is as follow.

$$\begin{cases} \text{lexmin}\{d_1^-, d_2^-\} \\ \text{s.t.}\Pr\{(\xi_{c1}+\xi_{t1})x_1 + (\xi_{c2}+\xi_{t2})x_2 + 2.2x_{1h} + 2.8x_{2h} \\ +20x_{1c} + 28x_{2c} + \xi_{a11}y_{11} + \xi_{a12}y_{12} + \xi_{a14}y_{14} + \\ \xi_{a21}y_{21} + \xi_{a22}y_{22} + \xi_{a24}y_{24} + \xi_{a32}y_{32} + \xi_{a33}y_{33} + \\ \xi_{a35}y_{35} + \xi_{a41}y_{41} + \xi_{a43}y_{43} + \xi_{a45}y_{45} + 1.5y_{1m} + \\ 1.9y_{2m} + 0.8y_{3m} + 0.6y_{4m} \le Z_0\} + d_1^- - d_1^+ = \alpha_1 \\ \text{Ch}\{\pi_0 - (\tilde{r}_1(\otimes)x_1 + \tilde{r}_2(\otimes)x_2) \le 0\}(\alpha_2) + d_2^- - d_2^+ = \beta_2 \\ Ch\{\tilde{D}_{1t}(\otimes) \le x_1\}(0.7) \ge 0.6,\ x_1 \le P_1 \\ Ch\{\tilde{D}_{2t}(\otimes) \le x_2\}(0.7) \ge 0.6, \\ x_2 \le P_2,\ 3000 \le x_{1h} \le 5000, \\ 3000 \le x_{2h} \le 5000,\ 0 \le x_{1c} \le 1500,\ 0 \le x_{2c} \le 1500 \\ 4000 = y_{1m}, 4000 = y_{2m}, 4000 = y_{3m}, 4000 = y_{4m}. \end{cases} \quad (16)$$

$$\begin{cases} \min\limits_{y_{11}, y_{12}, y_{14}} Z_1 \\ s.t.\ \Pr\{\xi_{a11}y_{11} + \xi_{a12}y_{12} + \xi_{a14}y_{14} \le Z_1\} \ge 0.98 \\ y_{11} + y_{12} + y_{14} = x_1,\ 1500 \le y_{11} \le \otimes_{y11}, \\ 1500 \le y_{12} \le 17000,\ 1500 \le y_{14} \le 20000. \end{cases} \quad (17)$$

$$\begin{cases} \min\limits_{y_{21}, y_{22}, y_{24}} Z_2 \\ s.t.\ \Pr\{\xi_{a21}y_{21} + \xi_{a22}y_{22} + \xi_{a24}y_{24} \le Z_2\} \ge 0.98 \\ y_{21} + y_{22} + y_{24} = x_2,\ 1500 \le y_{21} \le 13000, \\ 1500 \le y_{22} \le \otimes_{y22},\ 1500 \le y_{24} \le \otimes_{y24}. \end{cases} \quad (18)$$

$$\begin{cases} \min\limits_{y_{32}, y_{33}, y_{35}} Z_3 \\ s.t.\ \Pr\{\xi_{a32}y_{32} + \xi_{a33}y_{33} + \xi_{a35}y_{35} \le Z_3\} \ge 0.98 \\ y_{32} + y_{33} + y_{35} = x_1 + x_2,\ 2000 \le y_{32} \le 32000, \\ 2000 \le y_{33} \le \otimes_{y33},\ 2000 \le y_{35} \le \otimes_{y35}. \end{cases} \quad (19)$$

$$\begin{cases} \min\limits_{y_{41}, y_{43}, y_{45}} Z_4 \\ s.t.\ \Pr\{\xi_{a41}y_{41} + \xi_{a43}y_{43} + \xi_{a45}y_{45} \le Z_4\} \ge 0.98 \\ y_{41} + y_{43} + y_{45} = x_1 + x_2,\ 2000 \le y_{41} \le \otimes_{y41}, \\ 2000 \le y_{43} \le 44000,\ 2000 \le y_{45} \le \otimes_{y45}. \end{cases} \quad (20)$$

In order to solve this model, suppose the selected pop size is 30. The crossover probability is 0.8, and the mutation probability is 0.1. The fuzzy grey simulation, the random simulation and grey simulation are preformed 1000 times. The genetic algorithm runs 200 iterations for second-level slave models and run 200 iterations for first-level model.

Firstly, we suppose $\alpha_1=\alpha_2=\beta_2=0.8$. The Stackelberg- Nash equilibrium solution $d_1^- = d_2^- = 0.0$ can be obtained. The total cost can be controlled under 11000 ten thousands, and the revenue can be expected more than 1900 ten thousands under the 0.8 probability. $(x_1,x_2,x_{1h},x_{2h},x_{1c},x_{2c})$= (12256, 19782, 4486, 4153, 1160, 1077). y_{11}=2166, y_{12}=8382, y_{14}=1702; y_{21}=11604, y_{22}=3018, y_{24}=5145; y_{32}=25181, y_{33}=3097, y_{35}=3760; y_{41}=3529, y_{43}=23366, y_{45}=5145.

Then, we suppose $\alpha_1=\alpha_2=\beta_2=0.9$. The Stackelberg-Nash equilibrium solution is d_1^- =0.0 and d_2^- =0.0797. The total cost can be completely controlled under 11000 ten thousands with 0.9 probability. However, the maximal credibility measure under 0.9 probability for the revenue being more than 1900 ten thousands is only 0.8203, lower than the given 0.9 credibility. $(x_1, x_2, x_{1h}, x_{2h}, x_{1c}, x_{2c})$=(13713, 19759, 3579, 4465, 219, 1354). y_{11}=3145, y_{12}=7566, y_{14}=3003; y_{21}=10522, y_{22}=7084, y_{24}=2133; y_{32}=25192, y_{33}=3863, y_{35}=4416; y_{41}=3801, y_{43}= 25595, y_{45}=4072.

When α and β change from 0.8 to 0.9, we can find that the first objective can be met, the second objective gradually becomes unattainable, and the product amounts increase. In other words, with the increase of chance measure and probability level, the decision-maker becomes much more conservative, the total cost increases and the revenue decreases. This conclusion is also consistent with the realistic facts.

7. CONCLUSIONS

The uncertain factors can be described as random variables, grey variables and fuzzy grey variables according to their different situations. The grey simulation and fuzzy grey simulation can approximate the uncertain functions. The presented bi-level fuzzy grey hybrid production-planning model can be optimized with bi-level genetic algorithm. The uncertain programming considering the grey variables and fuzzy grey variables is more applicable than the stochastic programming. It develops the stochastic programming and is much more extensive. The feasibility of this model and algorithm is explained with a numerical example.

8. REFERENCES

[1] L. L Hau, V. Padmanabhan, et al "Information Distortion in A Supply Chain: the Bullwhip Effect", Management science, 50(Supplement), 2004, pp.1875-1886.

[2] J.A. Mirrlees, "The Theory of Moral Hazard and Unobservable Behavior: Part 1", Review of Economic Studies, 1999, 166(226): 3-21.

[3] B. Liu, Theory and practice of uncertain programming, Heidelberg: Physica-Verlag, 2002.

[4] M. Gen, B. Liu, "Evolution Program for Deterministic and Stochastic Optimizations", European Journal of Operational Research, 1996, 194(3): 618-625.

[5] S.F. Liu, Y.G. Dang, et al Grey System Theory with Application, Beijing: Science Press, 2004.

[6] B. Li, X.S. Gu, "Flow Shop Scheduling Based on Admixture Programming", Journal of East University of Science and Technology (Natural Science Edition), 2006, 32(1): 98-102.

[7] J. Zhou, B. Liu, "Analysis and Algorithms of Bifuzzy Systems", Fuzziness and Knowledge-Based Systems, June 2004, 12(3): 357-376.

[8] H.L. Lee, K.C. So, et al "The Value of Information Sharing in A Two-Level Supply Chain", Management Science, 2000, 46(5): 626-643.

[9] J.H. Ryu, V. Dua, et al Pistikopoulos, "A Bilevel Programming Framework for Enterprise-wide Process Networks under Uncertainty", Computers & Chemical Engineering, 2004, 128(6/7): 1121-1139.

A Knowledge Management Model for Middle And Small Enterprises *

Xingsen Li[1], Ying Liu[2], Jun Li[1], Yong Shi[2], Yuejin Zhang[1]
[1] Management School of Graduate University of Chinese Academy of Sciences, Chinese Academy of Sciences Research Center on Data Technology and Knowledge Economy, Beijing, 100080,China
[2]Chinese Academy of Sciences Research Center on Data Technology and Knowledge Economy, Graduate University of Chinese Academy of Sciences. 100080, Beijing, China
Email: Lijun@huatai-serv.com, lixingsen@126.com

ABSTRACT

Chinese middle and small enterprises are not able to undertake big projects due to lack of resources. They have to hire "able person" on short-term to fulfill important tasks. In order to find a solution for companies to utilize the free resources and skilled persons in society, extension theory and the principles of distributed computing were studied; by comparing supercomputer system with able person we realize that distributed computing theory can be used similarly to solve able person problem. Therefore, a novel knowledge management model is proposed. This model consists of project disassembling, task issuing, bidding, signing, monitoring and assembly testing etc. It makes good use of the available human resources in society and makes knowledge management a workflow. The results of application in a real company show that it can improve the management level of complicated projects with low resource cost of the middle and small enterprises.

Keywords: Distributed Computing, Middle and Small Enterprises, Knowledge Management, Extension Theory, Able Person.

1. INTRODUCTION

In China, middle and small enterprises often lack resources, such as able persons, management rules, high tech tools, standard processes, etc. Therefore, they are not able to undertake big projects by themselves. They have to rely on hiring "able person" to fulfill short-term projects. However, "able person" is not reliable and the companies that rely on "able person" often ignore the function of their company system for the sake of overemphasizing the influence of able person. It will raise a lot of serious problems and become an obstacle of a company's development.

In recent years, many companies have been aware of the weakness of "able person" culture and its negative impacts on a company's development. However, this problem is very difficult to solve because companies are often lack of system management knowledge and skilled persons.

In order to solve the above problems in middle and small enterprises in China, we first analyze the disadvantages of "able person" system, and then compare supercomputer system with able person in section 2. Section 3 explores the approach to solve the problems in supercomputer system. In section 4, we present a novel knowledge management model called DKMS based on the principles of distributed computing theory. Furthermore, we test the model in a real company and present the experiment results in section 5. Finally, Section 6 summarizes the paper with some remarks on the novel model.

2. ABLE PERSON SYSTEM VS. SUPERCOMPUTER SYSTEM

There are many disadvantages of "able person" system in middle and small enterprises.

Firstly, it is very costly to hire "able persons" for short-term projects. Their salary is 3-10 times of that of the regular employees.

Secondly, there are no enough able persons available for middle and small enterprises in society. It often takes a long time for a headhunting company to hunt the right able persons for their projects.

Thirdly, if the owner of a company always relies on "able person" to make important decisions, he has to allow him to break the company's rules. In this way, actually, it is the "able person" who controls the management system which will lead to fractions, or project delays, etc.

Last but not least, "able person" company ignores teamwork and training of junior managers. In a word, relying on "able person" is harmful for a company's long-term development.

All these problems are contradiction problems for the enterprises. Extension Theory, which was established in 1976 by Prof. Wen Cai in China, is a discipline which studies the extensibility of things, the laws and methods of exploitation and the innovation to solve all kinds of contradiction problems in real world with formalized models [1]. Extension theory establishes matter-element, affair-element and relation-element to describe matter, affair and relation. From the view of matter-element analysis in extension theory, a matter can be divided into two parts: imaginary part and real part, from the view point of material nature of matters, or soft part and hard part from the view point of system nature, which is called the conjugate nature of matter-elements. Every matter is the entity of the real part and the imaginary part. It's saying that real part is the base and imaginary part is what we used [2]. The soft part or imaginary part of one thing, such as theory, ideas, methods, and characters can be a reference for others or even be used to solve other contradiction problems.

All the characteristics of the "able person" system are similar to that of a supercomputer. So the ideas in solutions for supercomputer can be used to solve the above "able person" problems. Table 1 summarizes the characteristics of supercomputer and "able person".

* This paper has been partially supported by Key Project #70531040, National Natural Science Foundation of China; #70472074, #70501030, National Natural Science Foundation of China; 973 Project #2004CB720103, Ministry of Science and Technology, China. Corresponding author: Jun Li

Table 1. Comparison of supercomputer and able person

No	supercomputer	able person	contrast
1	Powerful, manage many tasks	Powerful, manage many business	same
2	Loading everything to a supercomputer to process is not feasible, huge cost for computation, storage, and communication	Taking multiple tasks is not feasible, large cost for communication, limitation of a person's ability	almost the same
3	Supercomputer is all in one system, any hardware or software error can interrupt the computer's processes	Critical customer list is under control of able person, any unexpected behavior can interrupt the company's business	same
4	Need careful maintenance	Need careful maintenance	same
5	Large data sets would potentially be unable to run in-core or would take a tremendous amount of time	Large project or complicated problem can't be solved on time by one or two persons	almost the same
6	High price and good environments	High salary and good work environments	same

We can see that "able person problem" is similar with supercomputer problem in some aspects. Therefore, we can solve the "able person" problem by exploring the concepts of distributed computing.

3. PRINCIPLES OF DISTRIBUTED COMPUTING SYSTEM

The principle of distributed computing is to solve a large problem by assigning small partitions of a problem to multiple computers and then combining the solutions from the sub problems [3]. Distributed computing system integrates networking, communication, computation and information to provide a virtual platform for computation and data management. Any device (e.g. personal computer，supercomputer, PDA, etc.), which has computation capability, could serve as a participating node in a distributed system; Internet or any wired/wireless network could serve as the connections between the geographically distributed nodes. Distributed computing techniques are in strong demand in many fields, such as business, government, research, science and entertainment.

Distributed computing has been investigated for many years in computer science domain. Many techniques and software have been very well developed so that the infrastructure can guarantee the availability, reliability and performance. Recent projects have been designed to use hundreds of thousands of volunteering computers all over the world to collaboratively work on many large projects, which is impossible for any one computer or person to solve in a reasonable amount of time. In addition to the benefit in terms of speed, strong reliability is also an important aspect of distributed system. With a single computer system, any unexpected event may bring the system to its knees. If a workload is distributed across multiple computers, each computer is relatively more independent and its local hardware problems won't affect the hardware of other computers (see figure 1).

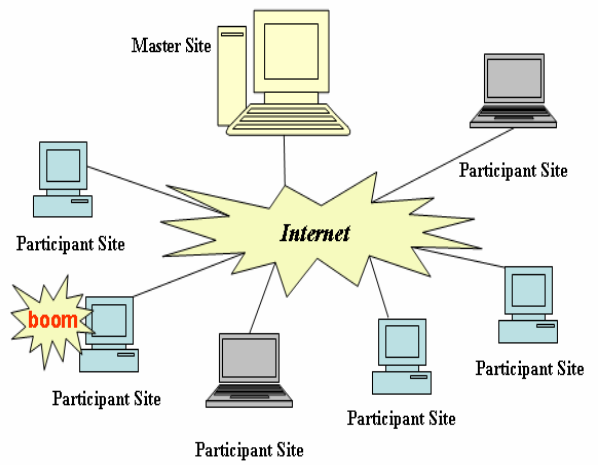


Fig. 1. Distributed computing environment

Figure 1 shows a typical distributed computing environment. A distributed computing system consists of three components: network connection, participating site and master site. The Internet, local area networks, and wireless networks are good examples of network connections. Each computer can serve as a participating site, performing the computation tasks using its local resources, such as CPU, memory, disk, and contributing its results to the master site through the network connection. A master site, which could be one of participating sites, is the coordinator of all the participants in this system.

Enabled by Internet, distributed computing can utilize all the available computing resources, services and data collections. Using multiple processors not only accelerates the computation, but also enables the use of more memory so that a larger dataset can be handled in the main memory attached to the processors [4]. Theoretically speaking, distributed computing can bring infinite CPU power, memory and storage. In our example in Figure 1, if enough candidate users sign up, the connected computers can surpass the fastest supercomputer four times at a fraction of

the supercomputer's cost. No project could resist the concept of more power for less money [5]. Then how does it work? Here is a distributed computing model.

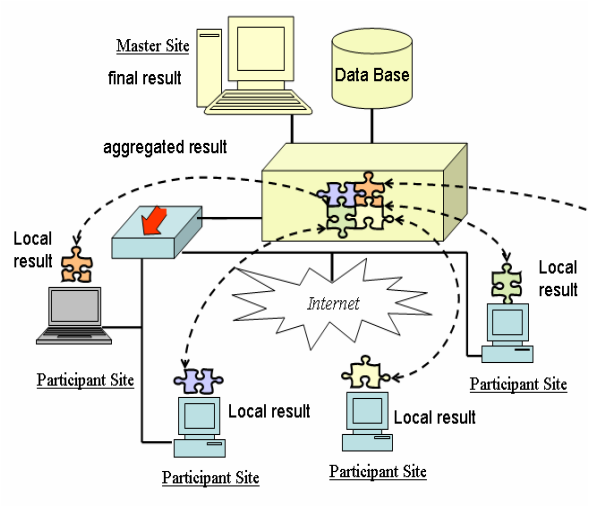


Fig. 2. Distribute computing model

Figure 2 shows a typical distributed computing model. Each participating computer performs the same computation tasks on its local data using its local resources, such as CPU, memory, disk. Then, they send the results obtained from their local data to the master server to be aggregated [6].

Table 2. DKMS vs. Distributed computing system

No.	Distributed computing	DKMS
1	Computation load is balanced on multiple computers. If one computer crashes, other computers are unaffected.	Tasks are balanced, rather than being controlled by a single person.
2	Share all the available computing resources, such as computer's CPU, memory, disk, no matter with the physical locations	Find all the available able persons to work for middle and small Enterprises
3	Each computer works on a sub-problem independently.	Each participant works on a sub-task independently.
4	Effective management software is required.	Effective management software is required
5	Security is important.	Security authorization is required.
6	Participating site's local results are sent to the master server to be aggregated	Participant's results need to be checked and aggregated
7	Network connection between participating site and master site is required.	Network connection between participants and companies is required.

4.2. Demand of knowledge management system

Knowledge management system also must be combined with corporation strategy [7], employee and operation execution. It usually consists of the following five steps:
(1) Define the goal of a given project;
(2) Disassemble the project into sub-tasks and assign them to different servers;
(3) Authorize qualified able persons;
(4) Monitor the progress;
(5) Collect the results and share knowledge.
Our proposed system provides three levels of management as following [8,9]:

1) Strategic management level

According to the development prospects and inside & outside competitive elements, company will set goals, choose executing approaches and divide the goals into separate working plans, get real time feedbacks and optimize strategy accordingly.

2) Operation management level

Perform real-time monitoring and adjust working plans based on the performance.

3) Business operation level

Execute specific business operations. Database is set up to store various experience or knowledge learned in the process. So each employee can reach needed knowledge conveniently.

Data communication (data transfer) between computers may happen if one computer needs to access data on another computer's local memory or disk. Synchronization may be incurred during the computation, for example, when a global variable is being updated.

Distributed computing also has disadvantages: A master computer has to distribute partitions of data and collect the results from participating sites. Collecting Internet users and having them sign up for a distributed computing program is not easy: organizations need to attract users with offers and incentives, from the screensaver to cash payments [5]. So is the "able person".

4. DISTRIBUTED KNOWLEDGE MANAGEMENT MODEL

By investigating the principles of distribute computing, we propose a distributed knowledge management system model (called DKMS). It is driven by the needs of middle and small enterprises to reduce costs and share services. The goal of our proposed model is to define the appropriated business and management platform, thus, to achieve end-to end management in support of business operations.

4.1. Functions from characters of distributed computing

The functions of DKMS from characters of distributed computing are listed in table 2:

In a word, to solve the problems of able person system, the key is to set up a system for company, employees, managers, and all the participants should work under the system, and the system should support company's strategy and train the employees.

4.3. Distributed knowledge management system model

By abstracting the basic ideas of distributed computing system and studying the knowledge management system, a lot of functions are designed, then, we propose a "simple" and "effective" knowledge management system model (DKMS) for the middle and small enterprise, showing in figure 3:

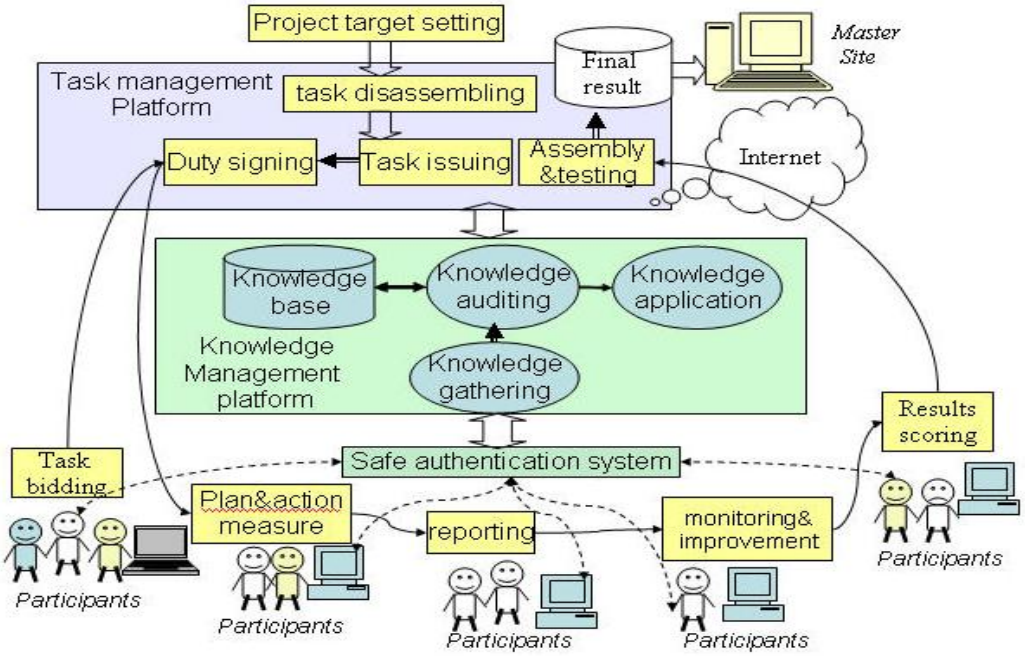


Fig. 3. Distributed knowledge management system model (DKMS)

Figure 3 is a task-sharing platform with knowledge management, which is closely combined with company's operation management.

It's working flow is as follows:

1) Project target setting: set up a yearly operation/budget plan or a project target, then divide it into separate business units with collaboration of participants, and distribute the tasks over the Internet.

2) Participants bid for the tasks. The company chooses the most suitable ones and then sign up duty contracts with the participants.

3) Participants who signed up the contracts propose working plans, including key measures, budget, time schedule, and operation objectives, action measures.

4) Participants send their reports to the performance monitoring system periodically after the task is partly finished. The reports should reflect facts, and give improvement proposals to ensure the success.

5) Score each participant's performance including his or her knowledge contribution. Then assemble the results from the participants into the final result.

6) Accumulate management and business knowledge learned in all these processes into a database. Make sure to save the useful knowledge into a knowledge base, so that the participants and employees can apply the knowledge to the management systems of their future projects [10, 11].

The hardware configuration of DKMS is shown in figure 4. The participants can be any low cost computing devices, such as personal PC, server, laptop, etc. Thus, the platform of DKMS is affordable by most small business companies. The model's network is displayed in figure 4:

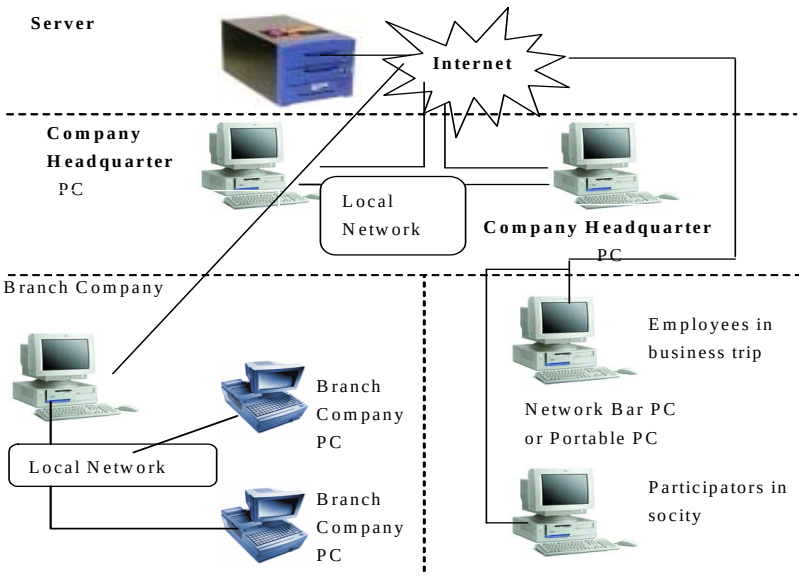


Fig. 4. DKMS environment

Most important, the distributed computing concept provides a challenge in cooperation: resource usage and administration must bridge technological, political and social boundaries. The distributed model will provide an incentive to the individuals to contribute to the success of the group.

5. CASE STUDY

In order to evaluate the effectiveness of DKMS, DKMS model has been developed into software and applied in companies. One of them is a famous national-wide engineering machine agent company. It has six subsidiaries in six provinces, acting as an agent of Korean Daewoo, Japanese Shinko. It sells grabs and other engineering machines, yearly sales value over 300 million Yuan. Its business includes product distribution, maintaining service, parts supply and technical support.

Although the performance of this company is acceptably good, it has serious problems: the performance gap between the experienced salesman and the novice is too wide, from more than 10 machines per month to zero per month. All the critical resources including the VIP customers, the banks, and the suppliers are controlled by a small group of people. By applying our proposed distributed knowledge management model, the company built an achievement-oriented culture, emphasizing the performance improvement. A database is set up to store more and more knowledge including sales channels, customer relationships, suppliers, etc. With the help of DKMS, beginners can start to complete their sales quota very soon and more talents are attracted to join in. Actually, its business obtained 7% improvement in sales in the first year of using DKMS.

6. CONCLUSIONS

In this paper, we propose a distributed knowledge management system (DKMS) for middle and small enterprises. It is a resource-sharing platform, where all employees of a company and all the available “able persons” have chances to contribute their capabilities to the enterprises with a proper compensation. On the company side, this knowledge management model is also the best carrier to implement business management by IT technology. The companies can extract knowledge from their daily business, and then store it in databases for their future reference. It makes international management experience and knowledge available for Chinese middle and small enterprises. It is a systematic platform that does not rely on any one or any small group of people. By integrating the contributions from many “able persons”, an able person system can be converted into a sharing system step by step. Eventually, a small company may accomplish large-scale projects by using DKMS.

7. REFERENCES

[1] W. Cai, C.Y. Yang, et al, A New Cross Discipline —Extenics, Science Foundation In China, 2005,13(1): 55-61.

[2] W. Cai, Extension theory and its application, Chinese Science Bulletin，1999, 44(17): 1538-1548.

[3] http://www.distributedcomputing.info/.

[4] J.W. Li, Y. Liu, and et al. Parallel Data Mining Algorithms for Association Rules and Clustering, Handbook of Parallel Computing: Models, Algorithms and Applications. Sanguthevar Rajasekaran and John Reif, ed., CRC Press, 2006.

[5] http://library.thinkquest.org/C007645/english/.

[6] B. Park, H. Kargupta, Distributed Data Mining: Algorithms, Systems, and Applications, Data Mining Handbook, Editor: Nong Ye, 2002.

[7] C.H. Lin, S.M. Tseng, Ridging the implementation gaps in the knowledge management system for enhancing corporate performance, Expert Systems with Applications, July 2005, 29(1): 163-173.

[8] M. du. Plessis, Drivers of knowledge management in the corporate environment, International Journal of Information Management, June 2005, 25(3): 193-202.

[9] Maija-Leena, Huotari, et al, Trust in Knowledge Management and Systems in Organizations, Idea Group Publishing, 2004, Information Processing & Management, May 2005, 41(3): 720-722.

[10] Dianne J. Hall, David Paradice, Philosophical foundations for a learning-oriented knowledge

management system for decision support, Decision Support Systems, May 2005, 39(3): 445-461.

[11] Mohammed Quaddus, J. Xu, Adoption and diffusion of knowledge management systems: field studies of factors and variables, Knowledge-Based Systems, April 2005,18(2-3): 107-115.

E-commerce Security Need Evaluation BP

Siqing Liu
Strategic Planning College, Chongqing Technology and Business University
Chongqing, 400067, P.R.China,
Email: Lsqbrian@126.com

ABSTRACT

Firstly, on the basis of e-commerce security characteristics, a clear hierarchy of e-commerce company needs was proposed, and the relationships of security needs and other needs were further analyzed; secondly, on the basis of hierarchy analysis of needs, the e-commerce company security ranking was provided and the acceptant security area was analyzed; thirdly, on the basis of security ranking, the cost and return of e-commerce security are analyzed; at last, based on above, the dynamic relationships of individual perceived security, value tropism and security needs, and the e-commerce security evaluation BP are formed, what's more each steps of analysis are figured.

Keywords: E-commerce security, Security needs, Perceived security, Security evaluation BP.

1. INTRODUCTION

There are probable three kinds of e-commerce in china, including e-commerce R&D, pure e-commerce company and common applied company. There are no green fields for them. There is no one product or practice that provides 100% comprehensive security-in-a-box for their entire enterprise networks.

2002 Computer Security Institute survey of 503 security personnel in U.S. corporations and government shows that 80% of respondents had detected breaches of computer security within last 12 months and suffered financial loss as a result, and only 44% were willing or able to quantify loss, which totaled $456 million in aggregate, and 40% reported attacks from outside the organization, 40% experienced denial of service attacks, and 85% detected virus attacks [1]. So many security problems and risks, must an in-depth understanding of basic security problems and relevant e-commerce solutions be provided, and help industry professionals implement the most advanced security problems.

As e-commerce security becomes of paramount importance, many different level measures are provided from the technology solutions, organizational policies and procedures, laws and industry standards, such as encryption, firewalls, security tools, security management, network security protocols, access control, virtual private network, authentification, intrusion detection, tunneling, proxy/agent system, Rainbow Series Library, Common Criteria for Information Technology Security Evaluation, etc..

Being different from other measures, the paper puts forward a new method of e-commerce security research. On the basis of hierarchy of e-commerce company needs and security ranking, the paper analyzes the acceptant security area, the cost and return of e-commerce security and the dynamic relationships of individual perceived security, value tropism and security needs, and the e-commerce security evaluation BP. What's more, each step of analysis is figured.

2. E-COMMERCE SECURITY CHARACTERIST_ ICS

The e-commerce includes three aspects of problems: no one should be able to view a person's or an organization's data without authorization; all authorized users should feel confident that the data presented to them is accurate and not improperly modified; authorized users should be able to access the data they need, when they need it [2].

E-commerce security can be broken up into six requirements, or tenets. All of the tenets are equally important for helping to ensure the confidentiality, integrity, and availability of data. The tenets are listed as follows [2]:

Identification is concerned with user names and how users identify themselves to a computer system; Authentication is concerned with passwords, smart cards, biometrics, and so forth. Authentication is how users demonstrate to the system that they are who they claim to be. Access control (also called authorization) is concerned with access and privileges granted to users so that they may perform certain functions on a computer system; Confidentiality is concerned with encryption. Confidentiality mechanisms help ensure that only authorized people can see data stored on or traveling across the network; Integrity is concerned with checksums and digital signatures. Integrity mechanisms help ensure that data is not garbled, lost, or changed when traveling across the network; Non-repudiation is a means of providing proof of data transmission or receipt so that the occurrence of a transaction cannot later be denied.

Another very important aspect of security is auditing. Audit logs may give the only indication that a security breach has occurred. Or, if the breach is discovered some other way, proper audit settings generate an audit log that can help administrators pinpoint the location and the perpetrator of the breach.

They are many reasons for the requirements such as malicious code, Hacking and cyber vandalism, Credit card fraud/theft, Spoofing, Denial of service attacks, Sniffing, Insider jobs, etc.

3. E-COMMERCE SECURITY NEEDS

Need is a physiological or psychological deficiency that makes certain outcomes appear attractive. As we know Maslow's Hierarchy of Needs [3]: physiological needs (warmth, shelter, and food), security Needs (protection from danger), social Needs (love, friendship, comradeship), ego needs (self respect, personal worth, autonomy), self-actualization needs (full potential). In the special condition some need is primary or secondary. We can consider Internet daily life and work as online economic activities, of which the inputs or outputs meet different needs; the content degree is just the contrast of need inputs and outputs. The different companies in e-commerce world may have many needs —— hierarchy of needs: Resource

needs (network
equipment, HR), security needs (fraud, hacker), development needs (scale, profit), social needs (commonweal, local development) and harmonious needs (harmonious employee, environment community and society, value actualization).

Fig. 1. Hierarchy of E-commerce Company Needs

From the above Fig.1., we can get that only the condition that the essential resource needs and security needs can be met, e-commerce companies can develop well and implement their social needs and harmonious needs, but the relations between hierarchy of needs are dialectic: even the e-commerce companies are in the top level——harmonious needs level, the resource needs and security needs exist. For every e-commerce company, the resource and security needs are basic and important.

Since security needs for e-commerce companies are so important, how can these companies evaluate their security needs? Just like the macroscopical criteria: Rainbow Series Library, Common Criteria for Information Technology Security Evaluation, and China Information Classified Security Protection. Based on the security needs of the e-commerce companies, the e-commerce security can be classified into security deadline, low security, moderate security, high security and security absoluteness. The five-security ranking just shows the degree of uncertainty and risk. The security deadline means catastrophic loss, e.g. crashing loss of all data, and beyond retrieve; low security means fatal loss, e.g. part data loss, and part being beyond retrieve; the moderate security means distinctive loss, e.g. part data loss, and beyond retrieval, but restoration from other equipment; the high security means low loss, e.g. important data loss but retrieval; the security absoluteness means lower loss, e.g. unimportant data loss. There are no green fields for every e-commerce company, and generally, the e-commerce companies can be in the moderate and high degree level. But the security need satisfaction degree and other needs satisfaction degree can be reverse: if other needs of an e-commerce company can be met well, the low security needs may be ok; if reverse, the high security is needed. (Fig. 2.)

4. E-COMMERCE SECURITY: COST

E-commerce security administration can be divided into five steps: security assessment, security policy, implementation, security organization, security audit. So the e-commerce security cost also can be divided into the same five steps, in which are there many increasing charges such as human resource cost, hardware cost, software cost, information dealing cost, and opportunity cost and also the diminishing charges such as risk loss cost. The relations of increasing and diminishing charges can be depicted in the below Fig. 3., and the optimization cost can be got.

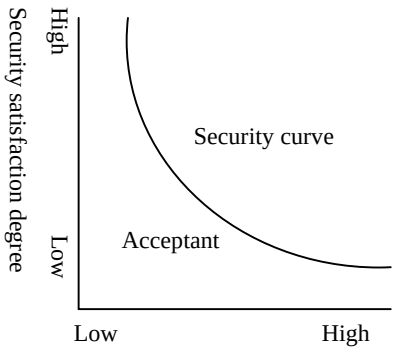


Fig. 2. Acceptant Security of E-commerce Company

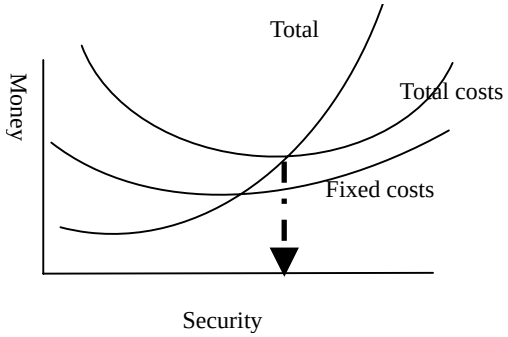


Fig. 3. the optimization cost choice of e-commerce security

5. E-COMMERCE SECURITY: RETURN

In 1972, K. J Arrow, the Nobel Prize winner, said that most economic decisions are made in considerable uncertain conditions, and the uncertainty is of the economic costs, and thus uncertainty reducing is a return. So e-commerce security problems can be settled and the risk reduced, the e-commerce should be efficient. The efficiency composes of recessive and obvious aspects such as smooth dealing, loss deduction, core competition and so on. If the All the security cost can be divided into the fixed and variable, the cost-return chart can be shown in the below Fig. 4. .

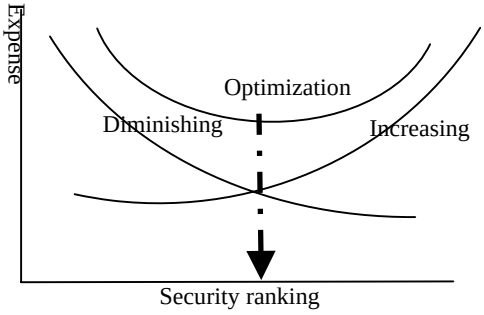


Fig. 4. Return Choice of E-commerce Security

6. E-COMMERCE APPERCEIVING SECURITY AND VALUE TROPISM.

Reid thinks that "perceive" is to obtain knowledge of through the senses; to receive impressions from by means of the bodily organs; to take cognizance of the existence, character, or identity of, by means of the senses; to see, hear, or feel; as, to perceive a distant ship; to perceive a discord. And Locke thinks that "perceive" is to take intellectual cognizance of; to apprehend by the mind; to be convinced of by direct intuition; to note; to remark; to discern; to see; to understand.

From the definitions of "perceive", the perceived security should obtain security knowledge through the senses, take intellectual cognizance of security, and to apprehend security by the mind. But each perceiving is under the direction of value tropism, and each individual has different value tropism, such as ease of use: the more security measures that are added, the more difficult a site is to use, and the slower it becomes; the desire of individuals to act anonymously. What's more, a person' security apperceiving may be got from his proactive concept, technology experience and subjective feeling. So Fig.5. the perceived security, value tropism and e-commerce security ranking can be got.

The perceived security and value tropism from the employees of e-commerce companies and e-commerce security are tied up. To some extent, the perceived security and value tropism are main obstacles on online translations and create negative influence on the e-commerce security. The dynamic relationships between the perceived security, value tropism and e-commerce security ranking are dipicted in the round. So under the direction of individual value tropism, the gap size of security apperceiving and security needs can be used to evaluate training level and risk idea.

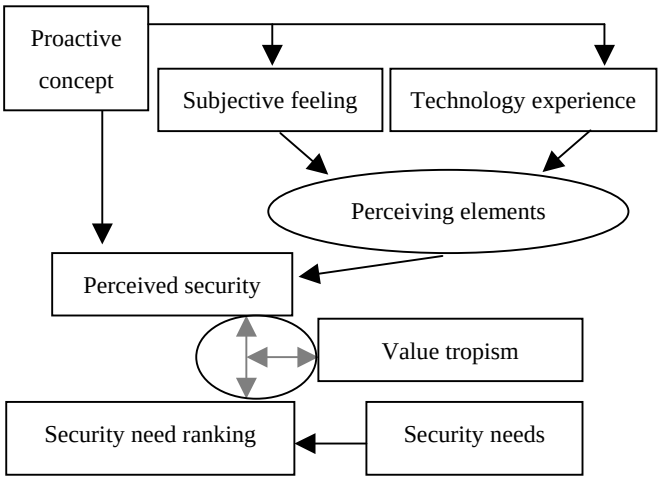


Fig. 5. The perceived security, value tropism and e-commerce security ranking [4, 5]

7. ANN MODEL FOR E-COMMERCE SECURITY NEED EVALUATION

A neural network is a massively parallel system comprised of many highly interconnected, interacting processing elements (also called nodes or neurons) based on neurobiological models of the brain (Dayhoff, 1990). A major task for a neural network is to learn and maintain a model of the world consistent enough to achieve the goals of the application of interest (Haykin, 1994). These systems act as non-linear, non-parametric function estimators that learn to map inputs to outputs on a non-linear, multidimensional surface to fit general nonlinear, multivariate functions (Zurada, 1992).

Neural networks exhibit many advantageous properties for solving complex problems. Back-Propagated Delta Rule Networks (BP) is a development from the simple Delta rule in which extra hidden layers (layers additional to the input and output layers, not connected externally) are added. The network topology is constrained to be feedforward: i.e. loop-free - generally connections are allowed from the input layer to the first (and possibly only) hidden layer; from the first hidden layer to the second, and from the last hidden layer to the output layer.

For e-commerce securtiy, the input layer provides the information security need, business flow securtiy need, administration securtiy need; other needs such as resource needs, development needs, etc.; the cost and return of security. The hidden layers including perceived security, security ranking and security value tropism, learns to recode (or to provide a representation for) the inputs. The output layer should be the e-commerce security need evaluation.
How to train the BP Networks? Forward pass: the outputs are calculated and the error at the output units calculated; Backward pass: The output unit error is used to alter weights on the output units. Then the error at the hidden nodes is calculated (by back-propagating the error at the output units through the weights), and the weights on the hidden nodes altered using these values. For each data pair to be learned a forward pass and backwards pass is performed. This is repeated over and over again until the error is at a low enough level (or we give up). So we can get the e-commerce security evaluation [6].

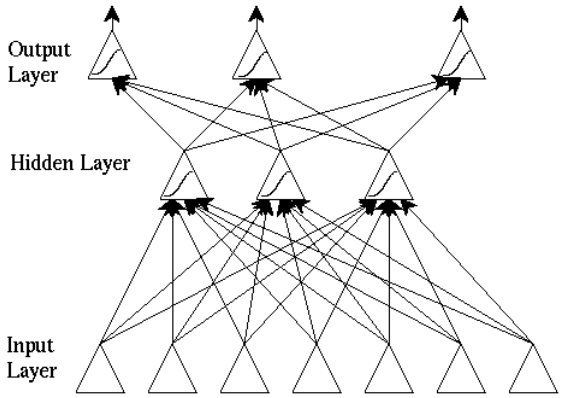


Fig. 6. Security needs ranking BP network

8. CONCLUSION

The paper provides a new method to make research and develop e-commerce security, and give readers thoughtful relationship models including hierarchy of e-commerce company needs, the acceptant security of e-commerce company under the security ranking from security deadline, low security, moderate security and high security to security absoluteness, the cost and return model of e-commerce security, and the relationships between the the perceived security, value tropism and e-commerce security ranking.

The paper quickly pinpoint e-commerce security requirements, cost and return, and the importance of the curity evaluation can be got. However, more study needs to be developed, such as the the individual perceiving quantification, the strategic research of e-commerce security, etc.

9. REFERENCES

[1] C. Guercio, Traver, "E-commerce business, technology, society", Second edition Kenneth C. Laudon, Copyright © 2004 Pearson Education, Inc.

[2] Huitt, W. (2004), "Security Administration", http://www.microsoft.com/technet/itsolutions/cits/mo/smf/smfsecad.mspx, Service Management Functions Updated: April 18, 2005

[3] Valdosta, "GA: Valdosta State University. Maslow's hierarchy of needs. Educational Psychology individual perceived security and value tropism. And the se_

Interactive", Retrieved from, http://chiron.valdosta.e_du/whuitt/col/regsys/maslow.html.

[4] X.Y. Gao, J. Zou, "Risk Analysis and Management of E- Commerce Security", JOURNAL O F WU T (IN FORMA T ION &MANA GEM EN T EN GIN EER IN G) Aug. 2005 Vo l. 27 No. 4.

[5] M.X. Wei, "Research on Influence of Perceived Security and Privacy upon the Trust in E-commerce Chinese Journal of Management", Vol. 2 No. 1, Jan. 2005. © 1995-2006 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

[6] Prof. Leslie Smith, "An Introduction to Neural Networks. Centre for Cognitive and Computational Neuroscience", Department of Computing and Mathematics University of Stirling, Last major update: 25 October 1996. http://www.cs.stir.ac.uk/~lss/NNIn_tro/InvSlides.html#what.

XOEM+OWL Based Production Semantic Retrieval on the Multi Heterogeneous Information*

Chengfeng Jian, Meiyu Zhang
Software College at ZheJiang University of Technology,
HangZhou, ZheJiang 310027, P.R. China
Email: jiancf@zjut.edu.cn , zmy@zjut.edu.cn

ABSTRACT

Aim at the STEP and SGML product information, the XOEM+OWL model is put forward for the uniform classification of product information. And then the correspondent mapping between STEP SchemaGraph and OWL SchemaGraph are build as Cos(sc,oc) to get the pattern matching degree for the semantic retrieval on the product information. At last a search system is presented which uses the method of the multi path-layers faced process and can resolve the identification and representation of production information.

Keywords: VO, STEP, Information Retrieval, XOEM+OWL, SchemaGraph.

1. INTRODUCTION

In the Virtual Organization (VO), knowledge sharing and exchange of product information over the Internet becomes a focus of research, which is different from virtual enterprise. Because of the characteristics of the autonomy and heterogeneity among the distributed enterprises, the sharing and exchange of product semantic information should meet the requirement of dynamic and agility [1]. With the development of the semantic web and semantic grid, it is possible to realize the on-line reconstruction on the product semantic information. And then the uniform product semantic representation is the prerequisite for the realization of the on-line information reconstruction over the Internet.

In this paper, at first in order to realize the identification and description of multi heterogeneous information such as STEP, SGML, multimedia, a general syntax description base on the XML [2] is represented. And with XML based syntax standards, we can realize the semantic description with OWL [3] for the interoperability and integration of Web services. According to the uniform classification of product information, aim at the STEP and SGML product information, the XOEM+OWL is put forward. And then the correspondent mapping between STEP SchemaGraph and OWL SchemaGraph are build as Cos(sc,oc). So we can get the pattern matching degree for the semantic retrieval on the product information. At last a search system is presented which uses the method of the multi path-layers faced process and can resolve the identification and representation of production information including STEP and SGML.

* This project is supported by the National Natural Science Foundation of China (No. 50305033) and the Natural Science Foundation of Zhejiang Province (No. X104631).

2. THE DESCRIPTION OF MULTI HETEROGENEOUS PRODUCT INFORMATION

The representation of product information has many ways. Generally it exists the two kinds standard representation for the product information. One is STEP for product data and another is SGML for product document. And besides, the Non-text mode such as multimedia contains much related product information. So we should resolve the identification so as to realize the capability of searching the multi heterogeneous product information.

1) The identification and description of product data: STEP->XML.

As the representation of the product data, the description language of STEP standard is the EXPRESS language. So the key technology is the pattern matching between EXPRESS and XML. Because the XML has much defect in the representation of the data structures and constraints, we can formulate the rules and definitions by means of the XML extensibility. According to the rules and definitions we can realize the pattern matching [4].

2) The identification and description of database information: Database->XML.

Generally it is the standardization information in the database, so we can extract its structure and convert into XML DTD. The question lies in the status of DTD.

3) The identification and description of non-text mode information: non-text mode ->text mode base on XML.

The non-text mode information has much ways such as image, video and so on. It can't be represented by the XML directly. So we should formulate the general description method by means of DTD and ensure the structure is the same with other representation of XML

(The structure is ignored temporary)

4) The identification and description of product document: SGML/HTML->XML.

The SGML standard defines the formal data model of documents and other data types and data abstractions. The XML is a subclass of the SGML. So we can easily convert the SGML into XML by means of the syntax parsing.

5) Dynamic building the OWL semantic association link among the XML based heterogeneous product information.

On the above basis we can use the semantic association link to build the semantic relation among the heterogeneous product information over the web.

3. XOEM+OWL: THE KNOWLEDGE REPRESENTATION AND EXCHANGE OF EXPRESS-XML-OWL

3.1 EXPRESS-XML-OWL

Because the XML is the subclass of SGML, the product

document represented by SGML can directly be identified in the Internet. We should realize the XML representation of EXPRESS [5].

EXPRESS is a powerful language for the definition of the data structure, it provides for defining the types of object and their properties and the constraints are far richer than those provided by the XML. On the other hand, XML is far richer than EXPRESS in the data transmission and identification. XML is easily interpretable by humans but EXPRESS not. OWL is also powerful for the semantic description and syntax express uses XML.

In order to realize the representation, according to the OO conception, Table1 shows as follows:

Table 1. EXPRESS-XML-OWL

OO	EXPRESS	XML	OWL
Object	Entity	Element type	Class
Object instance	Entity instance	Element	Instances
Object property	Entity attribute	Element	ObjectProperty /DatatypeProperty
OO declaration	Entity Schema	DTD	Ontologies
OO relationship	Complex Constraint express	hiberarchy	Complex Constraint express

3.2 Schema Graph

According to the XOEM [6], we can get the XML-based STEP representation. So we get the graph pattern matching between the XML-based STEP graph and OWL graph [7].

Table 2. XML Schema Graph

XML-based STEP	XOEM+OWL SchemaGraph representation
ComplexType	Node
Elementary XML Data Type Element	Node and an Edge between complexType node and this node with name “hasElement”
ComplexType XML Data Type Element	Edge
SimpleType	Node
Values defined for simple types	Node and edge between simpleType and this node with name “hasValue”
Elements	Nodes

Table 3. OWL Schema Graph

OWL	XOEM+OWL SchemaGraph representation
Class	Node
Property with basic datatypes as range (Attribute)	Node with edge joining it to the class with name “hasProperty”
Property with other class as range (Attribute)	Edge between the two class nodes
Instance	Node with edge joining it to the class with name “hasInstance”
Class – subclass relationship	Edge between class node to subclass node with name “hasSubClass”

3.3 Mapping STEP SchemaGraph and OWL SchemaGraph

Every concept from STEP SchemaGraph is compared against concepts from the OWL SchemaGraph[7]. The function listed in Table 4 calculates the match score (Cos) between a STEP SchemaGraph concept and OWL SchemaGraph.

Table 4. Mapping function

FUNCTION	Mapping
INPUTS	sc, oc Є W
OUTPUT	mi = (sci, ocj, Cos) where, Cos is the Match degree calculated for the mapping sci and ocj (Cos Є [0,1])

The Cos(sc,oc) is composed of two different measures Element Level Match (ElementMatch) and Semantic level match (OWLMatch). ElememntMatch provides the linguistic similarity of two concepts whereas OWLMatch takes care of semantic structural similarity. The Cos(sc,oc) is calculated as the weighted average of ElementMatch and OWLMatch as shown in Equation 1.

Equation 1. Cos(sc,oc)

$$Cos(sc, oc) = \frac{w_{sc} * ElementMat\quad ch + w_{oc} * OWLMatch}{w_{sc} + w_{oc}}$$

$$where\ ,\quad (0 \le w_{sc} \le 1)\quad (0 \le w_{oc} \le 1)$$

Weights w_{sc} and w_{oc} indicate the contribution of Element level match and Semantic level match respectively in the total match score. If two concepts have matching structure then it is more appropriate to give more weights to OWLMatch. Based on these conditions the values of w_{sc} and w_{oc} are changed as shown in Table 5.

Table 5. Match value

w_{sc}	w_{oc}	Condition
1	0	Only STEP syntax level representation
0.1	0.9	OWLMatch > 0.9, ElementMatch < 0.9
0.2	0.8	OWLMatch > 0.75, ElementMatch < 0.75
0.3	0.7	OWLMatch > 0.65, ElementMatch < 0.65
0.5	0.5	OWLMatch < 0.5, ElementMatch < 0.5 and XML-based STEP syntax level representation
0	1	All STEP semantic level representation

4. THE SEARCH METHOD OF MULTI PATH-LAYERS FACED THE PROCESS

4.1 The XML template definition of product information

We can define the uniform semantic template according to the product information classification for multi VOUs. Each product information record is divided into two parts: product_description and product_associations. Product_description contains the inner information and product_association contains the outer related information.

The ways of each product information record is divided into four parts: Product element, Product type element, Production factory element, and Location element. Product element mainly contains the detail information about product such as product code, product name, technical parameter, condition, price and so on. Product type element mainly contains the recursive structure of product type such as every son of product type, the son's flag of product type, the code of product type and so on. Production factory element mainly contains the detail information about factory such as the factory code, the main products about the factory, the flag of VO and so on. Location element mainly contains the recursive structure of the location about the product or factory.

```
<!DOCTYPE ProductInformation>[
    <!ELEMENT ProductItem(Product_Description, Product_Associations ) >
    <!ATTLIST ProductItem   Id ID  #REQUIRED
    Nm CDATA #REQUIRED ref IDREFS #IMPLIED >
        <!ELEMENT Product_Description
          (Product_BusinessInfo, Produc_TechInfo ) >
        ... <!– Product Business Info. and Product Technology Info. -- >
        <!ELEMENT Product_Associations
          (Inherit_elements, Path_elements ) >
        <!-- context-related-element : Inherit and path-- >
        <!ELEMENT Inherit _elements
          (#PCDATA|empth| Partsitem )* >
        <!-- context-related-element :parts name-- >
        <!ATTLIST Inherit _elements Id
          ID  #REQUIRED  Nm CDATA  #REQUIRED ref IDREFS  #IMPLIED >
        ...    <!– XML Link parts data-- >
        <!ELEMENT Path _elements
          (#PCDATA|empth| Pathsitem )* >
        <!-- context-related-element :paths name-- >
        <!ATTLIST Path _elements Id
          ID  #REQUIRED  Nm CDATA  #REQUIRED ref IDREFS  #IMPLIED >
        ...    <!– XML Link the structure of Product data-- >
          <!ELEMENT Pathsitem
          (MemberItem,ProductTypeItem, LocationItem...) >
          <!– VB Member cell, Product type cell, Location cell... -- >
    <!ELEMENT MemberItem
    (Member_Description, Member_Associations ) >
    <!ATTLIST MemberItem Id
      ID  #REQUIRED  Nm CDATA  #REQUIRED ref IDREFS  #IMPLIED >
    ...
    <!ELEMENT ProductTypeItem
    (ProductType_Description, ProductType _Associations ) >
    <!ATTLIST ProductTypeItem Id
      ID  #REQUIRED  Nm CDATA  #REQUIRED ref IDREFS  #IMPLIED >
    ...
    <!ELEMENT LocationItem
    (Location_Description, Location _Associations ) >
    <!ATTLIST LocationItem Id
      ID  #REQUIRED  Nm CDATA  #REQUIRED ref IDREFS  #IMPLIED >
    ...
>]
```

Fig. 1. the XML template for the syntax layer

4.2 Multi path-layers information retrieval with multi VOUs based XML

Information extraction and combination by means of the OWL-XML register repository accommodates the possibility of multi path-layers information retrieval, meantime, the uniform classification with product information accommodates to the way of multi path-layers information retrieval. Information extraction and combination takes place during the process of the information retrieval, the result of information retrieval is showed as the specified style. XML is the bridge with both of them.

On the basic ways of Product type, Production factory, Location and Product itself, by means of the identification of OWL, we can realize the intercross semantic search, tend to exactness step by step and recursion process in information search process. Fig.2 describes the search process.

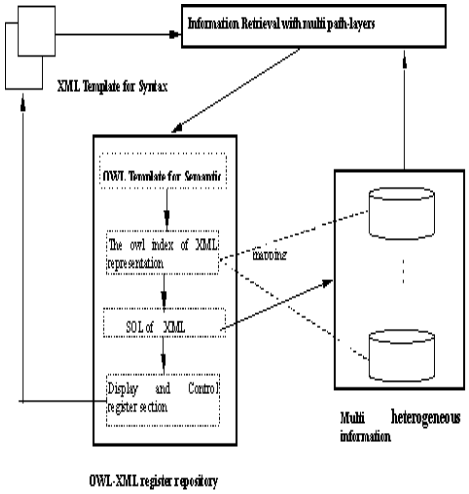


Fig. 2. Multi path_layers search

5. REFERENCES

[1] S. Agarwal, S. Handschuh, et al "Surfing the Service Web", Proceedings of the 2nd International Semantic Web Conference (ISWC2003), 2003.

[2] W3C. Extensible Markup Language (XML) 1.0. http://www.w3.org/TR/1998/REC-xml-19980210.

[3] W3C. OWL Web Ontology Language. http://www.w3.org/TR/2004/REC-owl-features-20040210/.

[4] C.F. Jian, J.R. Tan, "description and identification of XML-based STEP production data", Journal of Computer-Aided Design & Computer Graphics, 2001, 13(11): 983-990.

[5] ISO 10303 Industrial automation systems and integration –Product data representation and exchange Part11: Description methods: The EXPRESS language reference manual.

[6] C.F. Jian, J.R. Tan, "The Uniform Information Representation For On-line Reconstruction Of Virtual Enterprise Base On XML", Proceedings of the Third International Asia-Pacific Web Conference, Xi'an, China, 27-29, October 2000, 292-293.

[7] A. Ankolekar, M. Burstein, et al, "DAML-S: Web service Description for the Semantic Web", Proceedings of the 1st International Semantic Web Conference (ISWC2002), 2002.

An Approach for Business Knowledge Expression and Exchange

Xinyi BU, Xiaoyun WANG
School of Management, Hangzhou Dianzi University,
Hangzhou,Zhejiang,P.R.China 310018
Email:bxy002@126.com

ABSTRACT

In the supply chain management, a lot of information needs to communication between partner enterprises. Synonymous isomerism knowledge expression and computer-based understandable information is the critical problem to information automatic exchange. This paper constructed a business knowledge expressed framework based on the semantic. In the framework, we analyzed the business knowledge with the semantic method, described the semantic components of business knowledge and their relationship with RDFS (Resource Description Framework Schema), expressed component in detail with XML (extensible Markup Language). The business knowledge expressed by this framework can be identified and understood in the semantic level by computer. The research shows that it's a feasible way to realize the complex knowledge expression and exchange combining the semantic analysis with knowledge expression of RDFS and XML description.

Keywords: business knowledge，semantic model，supply chain management，RDF Schema，ontology interchange language.

1. INTRODUCTION

Supply chain management is one of the most advanced management models in modern enterprise operational management. A good supply chain management will greatly contribute to advantages of the partner enterprises in market competition. However, supply chain management poses a practical difficulty in the real world. As supply chain management involves a big number of sectors and enterprises spreading in various regions, it becomes necessary to provide smooth real-time exchanges and to share business knowledge among these enterprises. But, the real world is rather complex: different enterprises, even different departments within an enterprise, use varying knowledge formats resulted from varying application platforms and data bases; what's more, even the same knowledge may be described in different ways due to different perspectives and focuses. These complexities hinder knowledge interchanges in the supply chain, and become the major reason why it is difficult to transfer and share business knowledge in the current supply chain network.

As a business knowledge interchange standard and practical technology developed in the 20th century, EDI, to some extent and in a certain scope, gives us a way to solve the problem of business knowledge interchange in supply chain management. However, the high cost and complexity in the use of EDI retards its further popularization.

With the development of Internet technology, it is expected to produce a cost-efficient interchange technology for business knowledge by making use of such development. The emergence of XML makes it possible to turn this expectation into reality. XML is capable of describing data property in a rich and accurate way. By relying on this capability, we can use XML to transform business knowledge into a computer-recognizable data and construct a computer-understandable business vocabulary, thus, realizing the automatic identification of and interchange among business knowledge in the supply chain. This business knowledge interchange model is illustrated in Fig.1.

However, rather than basing on the semantic analysis and understanding of business knowledge, the business knowledge interchange as represented in the above model works merely through matching of knowledge's data properties, thus, deficient of semantic transparency and resulting low-level intelligence. Such interchange model uses low-reuse and inflexible core components, incapable of understanding data's deep meaning without mastering business knowledge's own semantics, and of introducing interchanges among isomerous business knowledge. Therefore, if the abundant business knowledge in the supply chain management can be described in a uniform way and modeled without platform's restriction from the semantic perspective, it will not only facilitate knowledge matching and interchanges among enterprises, but also play an important role in enriching business knowledge resources of the sharing enterprise, improving enterprises' responses to market conditions and lowering business operational costs.

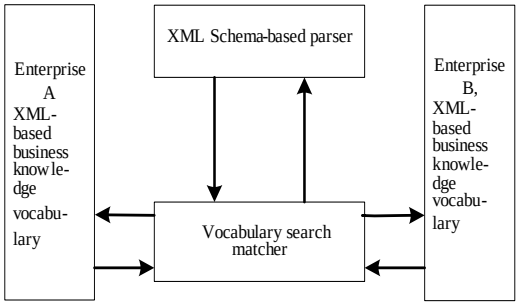


Fig. 1. XML Schema based business knowledge interchange

RDF based on XML solves the problem of describing knowledge from the grammatical and partially semantic perspectives, while the concept of ontology and its relative technologies offer exchanges among different sectors and varying application systems represented by the mutually understandable knowledge. Enhanced by the concept of ontology, RDF will have a more transparent semantic description of business knowledge.

2. RDF AND RDF SCHEMA

2.1 RDF

In 1999, W3C issued RDF, a metadata specification based on XML, which aims to provide a fundamental structure for metadata network applications and to make it possible to exchange and treat data by various application programs via the network.

RDF is based on the denomination property and property value model [1,2]. The basic RDF data model is composed of three pairs of object types:

1) Resource. Resource refers to the object represented and described by RDF. The business knowledge in the supply chain described by RDF also belongs to resource, including product or parts profiles embedded in web pages, procurement document and product releases, etc. Resources are usually marked with unified resource identifiers (URI). Any object possesses one URI.

2) Property. Property refers to the description of resource characteristics or relations, with specific meanings, i.e. property value, which covers appearance, feature, nature and relationship with other resources. Each resource has its own property. Every resource may have one or more different property values. Taking the supply chain management as an example, enterprise's procurement bill includes different values such as product name, quantity, specification and unit price, etc.

3) Statement. Statement when a specific resource is designated by one denomination property and its values, an RDF statement is generated. This statement consists of three parts: subject, predicate and object. The object (i.e., property value) may be a resource, which may share several values and is able to nest other resources. Therefore, RDF's data model has powerful capability to thoroughly describe any complicated business knowledge in the supply chain management.

In RDF's basic data model, elliptical node indicates resource; arc represents denomination property and rectangular node means property value, which is illustrated in Fig. 2.

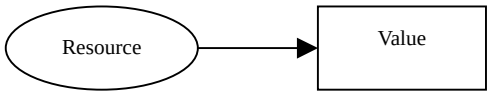


Fig. 2. RDF's basic data model

For example, "Ora Lasdila is the creator of the resource http：//www.w3.org/Home/Lassila" can be represented according to RDF basic data model as follows in Fig .3:

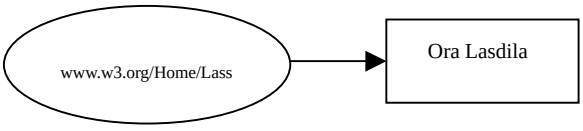


Fig. 3 .An example of RDF's basic data model

In constructing an RDF model, the frequently occurring resource knowledge set can be represented by a RDF capacitor. RDF defines three types of capacitor objects: Bag, resource's unordered list, used to describe multi-value property whose values may appear in random order and repeat as well; Sequence, resource's ordered list, used to describe multi-value property whose values must appear in fixed order but repetitions are allowed; Alternative, a list used to describe a property which might be replaced by several property (single) values, such as the alternative components in one enterprise's order when a supplier fails to deliver.

2.2 RDF Schema

RDF Schema (RDFS) defines object's categorization features or category system within certain fields that states the name of property category set defined by such field, the validity of the property of the designated category and the standard of the value restriction [3]. It can be said that to understand a specific RDF Schema means to understand the semantics of each attribute in the definition. In addition, as RDF model is built on RDF data model, computer can calculate the property type and its values via grammatical parse if even certain model is not understood in the application. Furthermore, in the RDF model schema, user can define extended model if necessary. RDF Schema defines the adopted term and element types, accurately linking every resource property and definition model. So, these features explain why RDF has good expansibility and openness.

3. ONTOLOGY AND OIL

By far there has been no universal definition for ontology. Among the most quoted is TR Gruber's definition, i.e. "An ontology is a formal, explicit specification of a shared conceptualization." A "conceptualization" refers to an abstract model of some phenomenon in the world, which identifies the relevant concepts of that phenomenon. "Explicit" means that the type of concepts used and the constraints on their use are explicitly defined. "Formal" refers to the fact that the ontology should be machine readable, as opposed to natural language. "Shared" means that ontology is the private property of an individual rather than of a unit, which is universally accepted at least within a research group. Intuitionally speaking, ontology is an entity resulted from the ontology analysis and modeling of a certain field applications, i.e., abstracting a certain field in the real world into a set of relationships of concepts [4]. These abstracting results will help to accurately define content meaning, efficiently communicate and improve system's mutual operation, reuse and sharing performance.

The representation and modeling, by using ontology technology, of the large quantity of business knowledge in the supply chain management will not only facilitate the exchanges among upper and down stream enterprises, but contribute to the mutual operations of enterprises' knowledge systems as well.

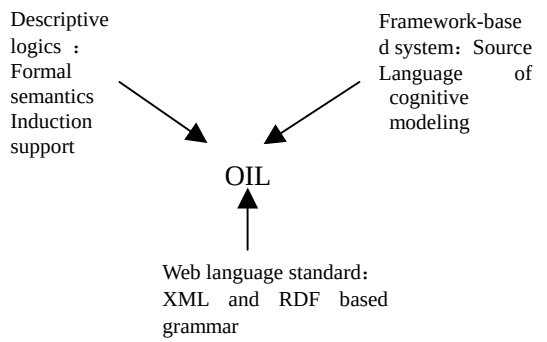


Fig. 4. OIL's relation with its three key technical foundations

There are several ontology languages such as CycL, KIF and Ontolingua, which are not yet widely applied to describe Web resources. Recently, W3C organization

recommended OIL（Ontology interchange Language） as a ontology description language for Web semantics. OIL integrates three kinds of technologies: framework description system provides the source language for modeling; descriptive logics provides formal semantics and induction; markup language standard（XML，RDF）provides language elements[5].

OIL is designed to provide a knowledge semantic representation, which is accessible to machines on the Web. OIL mainly integrates the achievements from three research areas to offer a semantic Web a universal purpose object markup language. The three major technical foundations' relationship with OIL is illustrated in Fig. 4.

As illustrated in Figure 4, OIL has RDFS as its key component. Because, on the Web, ontology language must link the existing Web language standard, otherwise cannot acquire wide applications. RDF Schema lays a good foundation for ontology definition. Extended EDFS can realize OIL specification.

4. SEMANTIC INTERCHANGW MODEL OF BUSINESS KNOWLEDGE

4.1 A RDFS and ontology based interchange model framework for business knowledge

To identify, understand and exchange business knowledge from the semantic level, it is necessary to construct a semantic parsing relation based on the concrete business knowledge. Through analysis and abstraction, form logical ontology clusters for business knowledge, and finally establishes a complete conjunction system chart about business operational knowledge. Use RDF Schema to identify and relate various knowledge elements among individual realities and in each individual reality. Use XML to realize these logical descriptions and form an ontology database for business knowledge elements. In the light of this idea, the interchange model for business knowledge in supply chain management shall be a three-tired model, which is illustrated in Fig. 5.

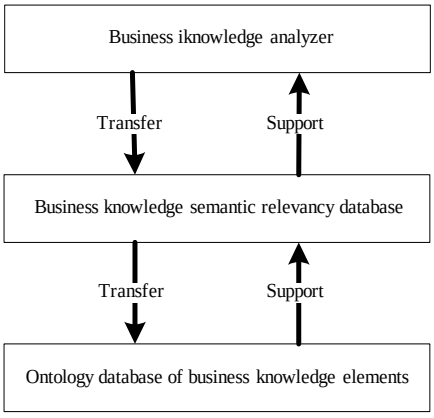


Fig. 5. Interchange model for business knowledge in supply chain management

In which, the function of business knowledge analyzer is, according to the received business knowledge and with the support of business knowledge semantic relevancy database, to have the semantic analysis of business knowledge, and to determine business background and action requirements and purposes so as to decide the subsequent actions (such as knowledge format change isomerous knowledge exchange and reconstruction).

The function of business knowledge semantic relevancy database is to store and manage the relevancy system chart of business operational knowledge, and to map and reconstruct semantic elements and support the work of business knowledge analyzer according to the signals from business knowledge analyzer.

The function of ontology database of business knowledge elements is to store and manage business knowledge elements, and to dynamically construct business knowledge blocks and to realize the computer understanding of business knowledge according to the mapping and reconstruction signals from business knowledge semantic relevant database.

4.2 RDFS and OIL based semantic description and modeling of business knowledge

In Section 4.1 we can see the basis of business knowledge interchange is the determination of the semantic-logic relevancy mapping relation of business knowledge and the modeling of business knowledge element ontology. To determine the semantic-logic relevancy mapping relation of business knowledge, it is required to have a detailed analysis, abstraction and categorization of the concrete business knowledge in supply chain management. Business knowledge element ontology is, according to the categorized and abstracted business knowledge unit (including property, action and knowledge transmission method, etc.)[6], accurately described by RDF and realized by XML or OIL.

The actual differences in the knowledge representations of varying enterprise's knowledge systems show different RDF Schemas in concrete modeling. For the purpose of semantic integrity constraint, the concept of ontology will be used to abstract every semantic element in modeling, restore the business meaning of business knowledge itself, and use RDF Schema to represent the knowledge-based semantic-logical connection [7,8]. When business knowledge element ontology is formed, OIL is embedded into RDF while the node business knowledge in the supply chain is semantically modeled. RDF Schemas of different enterprise are bound together according to enterprises' rules. The variables contained in the knowledge are used to link up the interface parameters of these enterprises. The Schema discrepancies of these enterprises are screened. Thus, as for the visited enterprises, different files are identifiable and interchangeable.

4.3 An example of modeling representation

In the display supply chain is Enterprise A, who is the supplier of Enterprises B and C. Enterprise A is to visit the Inventory knowledge of Enterprise B and C. The table 1 indicates the inventories of the above three enterprises: The above table shows observable differences among the database structures of the three enterprises. These isomerous data are defined by OIL, described in the corresponding extended RDF Schema and screened by IDL. Take the data structure of Enterprise A as an example, its OIL definition and RDF Schema description are given, while the RDF description of the concrete data in the defined data structure example is not repeated here.

1） OIL definition

< ontology-definition>

…

slot definition
…
<class-def>
<class name=”STRING”/>
<documentation>strings</documentation>
</class-def>
<class-def>
<class name=”Display name”/>

Table 1. Knowledge description formats

Enterprise A	Enterprise B	Enterprise C
<Product>	<Parts>	<Component>
<Name></ Name >	< Parts Name ></ Parts Name >	< Component Name ></ Component Name >
<Category></ Category >	< Category ></ Category >	< Category ></ Category >
<Size></Size>	<Price></Price>	<Unit Price></ Unit Price >
<Inventory></ Inventory >	<Model></Model>	<Specification></Specification>
< Unit Price ></ Unit Price >	<Supplier’s Inventory >	< Supplier’s Inventory >
	< Supplier’s Inventory >	< Supplier’s Inventory >
<Color></Color>	<Explanation></ Explanation >	<Notes></Notes>
</ Product >	</ Parts >	< Component >

<documentation>The name of all displays</documentation>
</class-def>
<class-def>
<class name=”CRT Display”/>
<documentation>The class of all displays</documentation>
</class-def>
<class-def>
<class name=”LCD Display”/>

<documentation>The class of all displays</documentation>
</class-def>
<class-def>
<class name=”Type”/>
<documentation>The class of all types of displays</documentation>
</class-def>
<class-def>
<class name=”Size”/>
<documentation>The class of all sizes</documentation>

</class-def>
…
<!class definition of stocks, Price, Color>
…
<class-def>
<class name=”Display”/>
<documentation>The class of all displays</documentation>
<subclass-of>
<class name=”CRT Display”/>
</subclass-of>
<subclass-of>
<class name=”LCD Display”/>
</subclass-of>
<slot-constraint>
<slot name=”Display Name”>
<value-type>
<class name=”STRING”>
<value-type>
</slot-constraint>
<slot-constraint>
<slot name=”TYPE”>
<value-type>
<class name=”STRING”>
<value-type>
</slot-constraint>
…
</class-def>
…
others class definition(Memory, Main Board etc.)
…
</ontology-definition>

2) Described corresponding extended RDFS

<rdfs: class rdf: ID=”Display”>

<rdf: type rdf:resource=”http://www.ontoknowledge.org/oil/rdfschema/2000/11/10-oil-standard#DefinedClass”>
<rdfs: Subclass Of rdf: resource=”#CRT Display”>
<oil: has Property Restriction>
<oil: Value Type>
<oil: on Property rdf: resource=”Display Name”>
<oil: on Property rdf: resource=”Type”>
…
</oil: Value Type>
</oil: has Property Restriction>
</rdfs: class>

5 . CONCLUSION

With RDFS and ontology technology as core modeling technologies, the research prospects and development trend of business knowledge semantic modeling and interchange model in the backdrop of supply chain management and their applications are promising. Meanwhile, RDF specification and ontology technology themselves are undergoing rapid developments, which also stimulate the research progress of business knowledge semantic modeling and interchange model. This means a benign interaction. Better methods and technologies are expected to perfect business knowledge semantic modeling and interchange model applied in supply chain management.

REFERENCES

[1]http://www.w3.org/TR/2000/CR-rdf-schema-20000327,“ Resource Description Framework (RDF) Schema Specification 1.0”.
[2] http://www.w3.org/TR/rdf-concepts/,“Resource Description Framework (RDF): Concepts and Abstract

Syntax”.
[3] http://www.w3.org/TR/2004/REC-rdf-schema-200402.
[4] J. Ma, “Ontology-based XML Schema Generation Method”, Computer Engineering and Applications, Nov. 2003, 57-62.
[5] X.J. Panl, “Semantic Web Markup Language and Ontology System”, Computer Engineering, Apr. 2003, 45-48.
[6] X.M. Chen, “Ontology-based Collaboration of Heterogeneous Enterprise Knowledge”, Computer Engineering and Applications, Jan. 2003, 23-26
[7] http://www.w3.org/TR/2004/REC-rdf-mt-20040210/, “RDF Semantics”.
[8] W.M. Zhang, Semantic Knowledge Modeling and Application, Beijing, Publishing House of Electronics Industry, 2002.

Tiered Prediction Mechanism in Collaborative E-Commerce

Changyou ZHANG
1) School of Computer Science & Technology,
Beijing Institute of Technology, Beijing, P.R.C, 100081
2）Computer Dept., Shijiazhuang railway institute,
Shijiazhuang, P.R.C, 050043
Email: chyv@vip.sina.com
And
Yuanda CAO, Minghua YANG, Zhitao GUAN
School of Computer Science & Technology,
Beijing Institute of Technology, Beijing, P.R.C, 100081
And
Liqun ZHANG
School of Economy Management,
Tsinghua University, Beijing, P.R.C, 100084

ABSTRACT

Many medium and small enterprises have implemented their collaborative e-commerce on ASP (Application Service Provider) platform. On this platform, the enterprises can group spontaneously into some alliances according to their business requirements. The entities in the e-commerce platform can be divided into four levels: host, enterprise, enterprise alliance and global ASP, according to their positions in the architecture. Many surveys show that the aforethought malicious intrusion with the purpose of stealing business secrets begins at a single host and then breaks into other hosts on the same level. To prevent such intrusion, a tiered intrusion prediction mechanism for B2B e-commerce platform was proposed in this paper. The four prediction levels of this mechanism were defined according to the scope of the disaster. Finally, this intrusion prediction system was implemented with C/C++ programming language on Linux platform.

Keywords: ASP, Collaborative E-Commerce, Tiered Intrusion Prediction.

1. INTRODUCTION

Many enterprises regard implementing e-commerce as an important means to enhance their own competitive ability. As a business strategy, e-commerce use electronic technology to capture, store, manipulate, analyze and visualize the diverse sets of trade-referenced data required for electronic communications among enterprises. Entities in e-commerce include enterprises, costumers, suppliers, business enterprises, governments and financial institutes and so on. For the sake of mutual assistance, they group into dynamic alliances to improve their competitive abilities [1]. Those allied enterprises put steady-going trust in each other and have frequent co-operations.

* National Natural Science Foundation of China under Grant No. 60563002 and Scientific Research Program of the Higher Education Institution of XinJiang under Grant No support this work. XJEDU2004I03.

B2B e-commerce based on ASP has developed rapidly since it meets the demands of the medium and small enterprises. The ASP providers lease their applications and related services to the e-commerce enterprises. The enterprises need not pay more attention to establish and maintain the e-commerce system, so they can merely focus on their own businesses. These enterprises can save a large amount of human power, equipments, and capital and obtain professional services.

In many surveys, people find that security is a vital factor that restricts the development of the e-commerce on line. The paper [2] classes the e-commerce security issues into network security and computer security. They consider that the vulnerable spots in both cases are at the endpoints – the customer's computer/network and the business' servers/network. To enhance the QoS and protect the e-commerce site, some researchers [3] raised a method to against DDoS attacks with improved D-WARD detection system. Security auditing mechanism provided by the paper [4] proposed a grey prediction module, which can detect the menaces.

In this paper, a tiered intrusion prediction mechanism was presented for B2B e-commerce platform. The four prediction levels of this mechanism were defined according to the scope of the disaster. The paper is organized as follows. Section 2 describes the architecture and the features of the B2B collaborative e-Commerce based on ASP. Section 3 presents the methodology of the prediction mechanism. Section 4 describes the implementation of a prototype system. Section 5 concludes our work and proposes the further researches.

2. ARCHITECTURE OF THE E-COMMERCE PLATFORM

The application platform provided by ASP is the core of the B2B collaborative e-Commerce based on ASP. Producers, suppliers, dealers, service providers, banks, logistics service agents are located logically around ASP application platform. Generally, as a super member, a producer creates an enterprises alliance that consists of a series of suppliers, dealers, service providers, logistics service agents, maybe a banks, etc. An enterprise can dynamically join several

different alliances and, similarly, a producer may create several alliances. Therefore, the ASP knits a mash between all the enterprises in the collaborative e-commerce platform. Enterprises alliance is a community, which composes a super member enterprise and several common member enterprises. These members of the alliance exchange trade information and develop their cooperation according to their contracts or agreements. The super enterprise of an alliance is the founder of the alliance. It stipulates the collaborate manners and is charged with the necessary administrative responsibility. The common members are the other enterprises. In accordance with their business types, the members can be classified as dealer, service provider, supplier, and so on.

The traffics between the members in the same alliance are in a very large quantity, and the traffics are rare between two enterprises respectively belong to different alliances. Moreover, the contents in the former traffic are highly related. Therefore, it is assumed that if a hacker breaks in a computer, namely target host, in order to steal business secrets, the most possible next target must be one of the hosts which communicate most frequently with the victim computer.

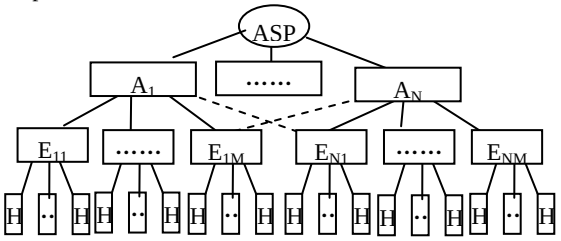


Fig. 1. Hierarchical structure of the e-commerce platform

3. METHODOLOGY

3.1. Definitions and symbols

(1) ***Host H*:** a host is the smallest entity in the e-commence system related to this paper; including all the equipments that involved in the e-commerce activities through network. These equipments may be desktop computers, servers, routers, palmtop devices, and so on. The symbol, H_K, stands for host *K*, and *K* is a positive integer.

(2) ***Enterprise E:*** an enterprise is a set of hosts. It can be formalized as $E = \{Hi \mid i = 1,2,\mathrm{K}\ ,M\}$. E_M stands for the enterprise *M*, and *M* is a positive integer.

(3) ***Enterprise Alliance EA*:** an enterprise alliance is a collection of enterprises and can be formalized as $\mathrm{EA} = \{Ei \mid i = 1,2,\mathrm{K}\ ,N\}$. E_{AN} stands for the enterprise alliance *N*, and *N* is a positive integer. In some cases, one enterprise is allowed to join in two different enterprise alliances. That is to say, $E_X \subseteq A_X$ and $E_x \subseteq A_y$.

(4) ***Global G*:** The global is the top-level entity in this paper. It is a set that contains all the enterprise alliances (*EA*s). Similar to the former entities, global can be formalized as $G = \{EAi \mid i = 1,2,\mathrm{K}\ ,W\}$.

According to the above definitions, we get the topology of the B2B e-commerce platform based on ASP (Figure 1). That is a hierarchical structure, but not a tree. E_{1M} and E_{N1} are connected to A_1 and A_N by a broken line respectively, for they belong to these two enterprise alliances.

Then, four corresponding prediction levels and four type of policy are defined as follows.

***(1) Host prediction PRED$_H$*:** This message will be created by a host and sent to the security manage center, when the host affirms that it is under threatening. A $PRED_H$ mainly consists of a host ID and a threat type description.

***(2) Enterprise prediction PRED$_H$*:** This message will be created by the security manage center of an enterprise and sent to the security manage center of the enterprise alliance when a given quantity of the same kind of $PRED_H$ from different hosts are received. A $PRED_E$ mainly consists of an enterprise ID, a threat type ID and an emergency response.

***(3) Alliance prediction PRED$_A$*:** This message will be created by the security manage center of an alliance and sent to the global security manage center when a given quantity of the same kind of $PRED_E$ from different enterprises are received. A $PRED_A$ mainly consists of an alliance ID, a threat type ID and an emergency response.

***(4) Global prediction PRED$_G$*:** This message will be created by the global security manage center and broadcasted to the all enterprise alliance security manage centers when a given quantity of the same level of $PRED_A$ from different alliances are received. A $PRED_G$ mainly consists of a target alliance ID list, a threat type ID, a threatening alliance ID list and a recommended response.

(5) Host prediction policy POLICY$_H$*:** It is a series of given safety precaution items to protect the host when the host creates a $PRED_H$. The host will adopt corresponding ***POLICY$_H$ items according to the encountering threat type.

***(6) Enterprise prediction policy POLICY$_E$*:** It is a series of given safety precaution items to protect the enterprise when the enterprise security center creates a $PRED_E$. The center will adopt corresponding $POLICY_E$ items according to the encountering threat type.

***(7) Alliance prediction policy POLICY$_A$*:** It is a series of given safety precaution items to protect the enterprise alliance when the alliance security center creates a $PRED_A$. The center will adopt corresponding $POLICY_A$ item according to the encountering threat type.

***(8) Global prediction policy POLICY$_G$*:** It is a series of given safety precaution items to protect the global ASP platform when the global security center creates a $PRED_G$. The center will adopt corresponding $POLICY_G$ items according to the encountering threat type. In this case, the global center can force all the alliance to take some necessary steps.

3.2. Hypothesis

This prediction method fits for the attack behavior with distinct purpose to a specific colony of business parties. On the assumption that a host H_{AD} was broken in, the probability that the hacker choose a host from the same enterprise as the next target is markedly larger than that from another enterprise. In this case, we say this probability is a conditional one, and describe it as $P(H_{\mathrm{NEXT}} \in E_A \mid H_{\mathrm{AD}})$. If we suppose H_{AD} belong to E_A, E_A is a member of A_A, and A_A is a portion of G_A, i.e. $H_{AD} \in E_A$, $E_A \in A_A$, $A_A \in G_A$, the following formula come into existence.

$$P(H_{NEXT} \in E_A \mid H_{AD}) > P(H_{NEXT} \notin E_A \mid H_{AD})$$
$$P(H_{NEXT} \in A_A \mid H_{AD}) > P(H_{NEXT} \notin A_A \mid H_{AD})$$
$$P(H_{NEXT} \in G_A \mid H_{AD}) > P(H_{NEXT} \notin G_A \mid H_{AD})$$

Based on the above formula, we suppose H_X is a random choice of host, E_X is a random choice of enterprise, A_X is a random choice of alliance, H_{NEXT} is the next attack target, $\overline{[\mathrm{H}]}$ is the set contains all the hosts under attacking, $\overline{[\mathrm{E}]}$ is the set contains all the enterprises under attacking, $\overline{[\mathrm{A}]}$ is the set contains all the enterprise alliances under attacking. Analyzing into the policy, we find the intrusion is a gradual

diffusing process. Therefore, a prediction mechanism based on attacking scope is proposed in this paper. This mechanism is efficient in the case of attacking scope increasing by degree.

Step1: A host, H_{AD}, find intrusion, then implements $POLICY_H$, creates and send $PRED_H$ to enterprise security manage center;
Step2: The enterprise security manage center receives the $PRED_H$, recalculates $COUNT_H$＋＝1; if $COUNT_H \leqslant V_H$, implements $POLICY_E$, then repeat Step2; else, creates and sends $PRED_E$ to the enterprise alliance security manage center and implements $POLICY_E$;
Step3: The enterprise security manage center receives the $PRED_E$, recalculates $COUNT_E$＋＝1; if $COUNT_E \leqslant V_E$, implements $POLICY_A$, repeat Step3; Else, creates and sends $PRED_A$ to global security manage center, implements $POLICY_A$;
Step4: The global security manage center receives the $PRED_A$, recalculates $COUNT_A$＋＝1; if $COUNT_A \leqslant V_A$, implements $POLICY_G$, then repeat Step4; else, creates and sends $PRED_G$ to the given super administrator and implements $POLICY_G$;
Step5: Algorithm finishes.

Fig. 2. Prediction algorithm

3. 3. Prediction algorithm

According to the above section, the prediction algorithm shows in Figure 2.

In Figure 2, $COUNT_H$ stands for the count of broken host. $COUNT_E$ stands for the count of broken enterprise. $COUNT_A$ stands for the count of broken host. V_H is a top threshold related to $COUNT_H$. V_E is a top threshold related to $COUNT_E$. V_A is a top threshold related to $COUNT_A$. The other symbols have the same mean as the above context.

4. THE REALIZATION OF THE PREDICATION SYSTEM

To be more stabile and reliable, the software was developed on the LINUX platform. In this project, we program the main process use C/C++, construct the user interface use GTK/GDK, and store data in database MYSQL. In the programming, some other technologies are involved in, such as multi-thread, SOCKET, and so on.

In this paper, every policy at each level is a list of items. Accordingly, we define a list node for each policy item in our program. Then the whole policy can be defined as a dynamic list. Limited by the length of the paper, we narrate the main algorithms as follow sections.

4.1 Node define of the policy

First, the policy files of the prediction system were created according to the defined policy data. Then, these files will be store in the hard disc. The program initiates the policy list at the beginning of the startup through loading the policy file. The prediction type ID is the primary index of the policy list. The list node will be formed from one line, which is a record standing for an item, in the policy file. Each line was divided into three segments by the symbol of '!'. The prediction type ID is laid in the first field, then, the top threshold is laid in the second field, and a recommended response was laid in the last field.

The data structure of the policy list node is described in Figure 3.

preID	topThreshold	num	hostList

Figure 3. **Define of the policy list node**

4.2. Program of Prediction

At the initialization, the program opens the policy file and loads the policy items one by one. After this process, the policy list will be established.

When the enterprise security center received a piece of prediction message, the program searches the policy list according to the prediction type ID. If it finds a node, which matches this ID, the value of num, which is the third member of the node, adds one. At the same time, the source of the message is filled in the hostList field. Here, if $num \geq topThreshold$, program alarms, then, creates and send recommended suggestion and hostList to the given entities according to the policy.

5. CONCLUSIONS AND FUTURE WORK

With the development of B2B e-commerce based on ASP, security becomes one of the vital issues that urge the ASP suppliers to implement appropriate countermeasures. In this paper, we consider the features that the enterprises group into alliances and form a tiered architecture, and we regards the amount of the entities that are broken-in through at the same level as the prediction parameter. Then we raise four prediction levels and realize the prediction prototype system.

We specify the algorithm from the view of the application entities in the e-commerce platform. We have not adopted unified entity identifier in the program. This is more flexible, but lacks uniformity. The intrusion detection methods adopted by hosts are not involved in this paper. We merely abstract the information from the messages created by the local IDS to form our prediction messages. This mechanism can be used in other distributed systems by some revisal.

6. REFERENCES

[1] S. Li, J. Yao, X. F. Shen, et al., "ASP Resolution of Regional Virtual Enterprise for Small and Middle Enterprises". Computer Integrated Manufacturing Systems. Vol.10, 2004.5,590-593

[2] R. C. Marchany, J. G. Tront, "E-Commerce Security Issues". Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS-35.02), Jan 2002,2500- 2508

[3] J. Kang, Z. Zhang, J. B. Ju, "Protect E-commerce against DDoS Attacks with Improved D-WARD Detection System". e-Technology, e-Commerce and e-Service, 2005. IEEE '05. Proceedings. The 2005 IEEE International Conference on 29 March - 1 April 2005, 100-105.

[4] C. Y. Zhang, C. H. Piao, et al., "The security audit mechanism in the enterprise e-commerce platform". Advances In E-Engineering And Digital Enterprise Technology-I, Proceedings, 2004.9, 499-506

Research and Realization of Medicine Materials Sales BI System

Hong Xue, Zaiwen Liu, Xiaoming Shun
(College of Information Engineering, Beijing Technology and Business University, Beijing 100037, China)
Email:hongxue6@yahoo.com.cn

ABSTRACT

A management information system based on transaction database was built on the basis of the data of medicine raw materials sales. The system provided a platform which can help users collect and accumulate historical data of enterprises. The business intelligence system of B/S framework, using Data Warehouse and Web technologies, was developed. The system separated the data used for decision and analysis from the transaction database. And a data warehouse was built. It was subject-oriented, integrated, time-varying and nonvolatile. On the basis of the data warehouse, we built multidimensional data analysis making use of Mondrian OLAP engine based on Java technology，which provided decision-makers with a great deal of statistical and analytic data and charts. Data Mining module detected the purchasing laws of customers by analyzing relativity of products bought by customers，thereby channeling sales, enhancing the intention of sales and increasing sale. The system provides effective support for decision-making of managers in medicine materials sales.

Keywords: Business Intelligence(BI)，Data Warehouse，On-Line Analytical Processing，Data Mining，Association Analysis.

1. INTRODUCTION

The many corporations which sell medicine materials are characterized by no high informational level, which makes medicine raw materials sales the obstacle of the whole medicine industry operation in China. BI system can analyse and mine helpful information from a great lot of daily operation data, which provides effective support for decision-making of managers in medicine materials sales.

2. SYSTEM COLLECTIVITY DESIGN

The system adopted B/S framework.It was composed mainly of system operating page layout, application server and Data warehouse server.It included four function modules — sale information management, decision support, online order and system management.Programme language adopted Java language. Interface codes used JSP.Database adoptsed MySQL Database. OLAP engine adopted Mondrian OLAP engine based on Java technology.

3. REALIZATION OF SALE INFORMATION MANAGE FUNCTION

3.1 Establishment of database connect pool

JDBC is a technology of calling on database and has the advantages of simpleness and easy use. Web application program was developed making use of JDBC technology, which made management of database connect resource low efficiency. Therefore Database connect pool was built for efficient management of database connect resource.

```
class DBConnectionPool {
    ……
    public DBConnectionPool(String URL,String
            user,String password,int maxConn){
        this.URL = URL;
        this.user = user;
        this.password = password;
        this.maxConn = maxConn;
    }
    public synchronized void
          freeConnection(Connection con){
        freeConnections.addElement(con);
        connNumb--;
        notifyAll();
    }
    public synchronized Connection
        getConnection() throws SQLException{
      Connection con = null;
        ……
    }
    private Connection newConnection() throws
          SQLException {
        Connection con =
          DriverManager.getConnection(URL,
          user, password);
        return con;
    }
}
```

3.2 Realization of the system exhibiting layer

The system exhibiting layer was realized using Struts framework.

1） Controller

In Struts, ActionServlet is a controller.ALL control logics are collocated by Struts-config.xml document.
Part codes of Struts-config.xml documentary:

```
<!--Collocating various kinds of ActionForm in products
information manage module-->
<struts-config>
<form-beans>
    <form-bean name="ManageProductForm"
  type="com.ebusiness.presentation.salesmanage.
                ManageProductForm"/>
    <form-bean name="AddProductForm"
  type="com.ebusiness.presentation.salesmanage.
                AddProductForm"/>
    <form-bean name="EditProductForm"
```

```
      type="com.ebusiness.presentation.salesmanage.
                        EditProductForm"/>
        </form-beans>
  <action-mappings>
       <!--Collocating various kinds of Action in
           products information manage module-->
      <!-- manage product! -->
      <action      path          =
                "/salesmanage/product/manage"
                   type          =
  "com.ebusiness.presentation.salesmanage.
                     ManageProductAction"
                   Name    = "ManageProductForm"
                   scope   = "request"
                   validate = "false"
                   input     = ""
    >
    <forward name="manage"
      path="/salesmanage/product/manage.jsp" />
    </action>
   </action-mappings>
  <message-resources
parameter="com.ebusiness.presentation.salesmanage.
                     applicationSalesmanage"/>
  </struts-config>
```

2） Views

The views were completed by JSP making page layout. Struts provides abundant JSP label storeroom — Html，Bean，Logic and Template,which is propitious to part behaving logic and program logic.

3） Model

The models existed in the form of a or several java beans that can be divided up into three kinds—Action Form(FormBean)、Action(ActionBean)、JavaBean or EJB.

Acting ActionBean codes in the module of products information manage:

```
public class ManageProductAction extends Action {
   public ActionForward execute(ActionMapping
    mapping, ActionForm form,HttpServletRequest
          request, HttpServletResponse response)
          throws Exception {
    //Example of seting up kinds
     ManageProductForm manageproductform =
                  (ManageProductForm)form;
     ProductManager manager = new
                          ProductManager();
    // Inquiring about condition of obtaining page
    layout — id parameter
      Long id = manageproductform.getId ();
      String action = manageproductform.getAction();
  Long productId =
                manageproductform.getProductId();
     if(productId ==null&&request.
            getParameter("productId")!= null){
        productId = new
          Long(request.getParameter("productId"));
     }
     //To carry out inquiring about means and deleting
means
     List productList = null;
     if(action.equals("query")){
       if(id!=null&&id.compareTo(new
                          Long("-1"))!=0){
          productList=(List)manager.
           getProductByStoredAmount(id);
    }else{
      productList = (List)manager.getAllproduct();
    }
  }
  if(action.equals("del")){
    manager.deleteProduct(productId);
    productList = (List)manager.getAllproduct();
  }
   //To return results to page layout
   request.setAttribute("productList",productList);
  return (mapping.findForward("manage"));
  // To return results to page layout named by manage
  }
}
```

4. DESIGN OF DATA WAREHOUSE MODEL

Snowflake model was adopted in the system for reducing redundance, easy maintenance and saving storage space to OLAP analyse[1,2]. The order form fact table was center. Dimension tables were linked to the fact table by main key. The layers of dimensions were divided according to system request (Fig. 1.).

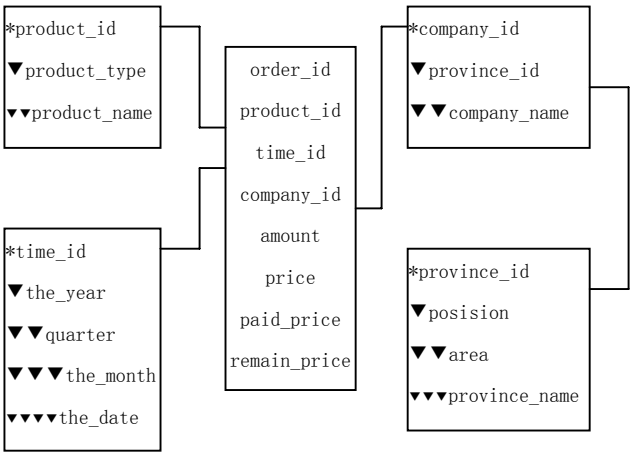


Fig. 1. Structure of system snowflake model

5. ESTABLISHMENT OF DATA CUBE

For OLAP analyse, data cube was established first[2].The system adopted Mondrian OLAP engine.To establish data cube adopted manner of XML script.The information about cube,dimensionality and measurement was compiled to XML documents according to XML criterion.When OLAP engine was started,the system analysed automatically many dimensionalities data according to the script. XML script of establishing sale data cube:

```
  <Schema name="DrugSales">
  <!—establishing data cube-->
  <Cube name="Sales"><!—cube name-->
    <Table name="orders" /><!—fact table name-->
    <!—establishing product dimensionality-->
    <Dimension name="Product"
            foreignKey="PRODUCT_ID">
    <Hierarchy hasAll="true"
              primaryKey="PRODUCT_ID"
```

```
                    primaryKeyTable="product">
          <Table name="product"/>
          <Level name="typeName" table="product"
              column="PRODUCT_TYPE"
 uniqueMembers="false" />
          <Level name="productName"
              table="product"
                  column="PRODUCT_NAME"
   uniqueMembers="true" />
     </Hierarchy>
     </Dimension>
   ……
 <!—establishing selling amount measurement-->
   <Measure name="amount"
      column="AMOUNT" aggregator="sum"
formatString="#,###.00"/>
 <!-- establishing sales measurement-->
  <Measure name="paidPrice"
     column="PAID_PRICE" aggregator="sum"
 formatString="#,###.00"/>
   <Measure name="productCount"
              column="product_id"
               aggregator="distinct count"
 formatString="#,###"/>
   <Measure name="customerCount"
              column="company_id"
               aggregator="distinct count"
  formatString="#,###"/>

 </Cube>
 </Schema>
```

6. OLAP EXHIBITING

Mondrian OLAP engine was realized by the label storeroom of JPivot. Taglib based on Java technology in JPivot used Web Component Framework (WCF) to exaggerate Web UI groupwares based on XML/XSLT.

First, correlative information of connecting database and URI of XML script were collocated in JSP label parameters.Then MDX inquiring sentence was compiled according to the format of MDX language in label system.If there are many MDX enquiries, many label systems were compiled.

```
 <jp:mondrianQuery id="query01"
          jdbcDriver="com.mysql.jdbc.Driver"
 jdbcUrl="jdbc:mysql://localhost:3306/ebusiness"
              jdbcUser="root" jdbcPassword="sa"
 catalogUri="/decsmanage/queries/DrugSales.xml">
 select
   {[Measures].[amount], [Measures].[paidPrice]}
                                on columns,
   {([Product].[all products])} ON rows
 from Sales
 where ([Time].[year].[2004])
 </jp:mondrianQuery>
```

When OLAP user interface was exaggerated, JPivot used the label of wcf to transfer various figure groupwares, such as slice, cutting piece, drilling and circumrotating.For example, <wcf:renderref="table01" xslUri="/WEB-INF/jpivot/table/mdxslicer.xsl" xslCache="true"/> transfered slice groupware so the corresponding groupware was showed on user interface.

7. REALIZATION OF DATA MINING ARITHMETIC

Association analysis function module in which Apriori arithmetic was core was developed in the system, which was used for mining various kinds of interesting relations among manufactures bought by customers and helping corporation sale supervisors to set down sale strategies [3,4].

Association analysis was carried through based on sale information picked (Table 1), least support degree(S=40%) and least belief degree(C=40%). DA={vitamin B1 saltpetre B，vitamin B1 salt C，vitamin B6 D，vitamin B2 E}.

Table 1. Affair Database

TID	Merchandise list
1	B
2	B,D
3	C
4	B,D
5	B,C,D
……	……

The arithmetic was divided up into two steps according to the characters of Apriori—connection and pruning.

（1）Connection: Candidate k-term volume C_k was produced when frequent k-1-term volume was connected with itself —$L_{k-1} \infty L_{k-1}$.The elements of L_{k-1} were attachable.

（2）Pruning: According to the characters of Apriori , Candidate terms that were in (k-1) subclass of Candidate k-term volume C_k and not in L_{k-1} were deleted ,which compressed C_k.To search database decided the count of each candidate term in C_k.Then all candidate terms that met the count of least support degree came under frequent k-term volume L_k.

Partial codes:

```
private FastVector SelfConnect(FastVector La){
      int i,j;
      FastVector Cb = new FastVector();
      FastVector pieces = BreakOne(La);
       FastVector temp ,swp,add ;
      Item item ;
       for(i=0 ; i<La.size() ; i++){
           for(j=0 ; j<pieces.size() ; j++) {
              temp = (FastVector)La.elementAt(i) ;
              item = (Item)pieces.elementAt(j);
              if( !IsContain(temp,item)) {
                    add = new FastVector();
                    add.appendElements(temp);
                    add.addElement( item ) ;
                    if( !IsContain2(Cb,add) )
                          Cb.addElement( add ) ;
                  }
              }
           }
```

The results suggest that corporation forwardly sells vitamin B2 and vitamin B1 saltpetre to clients of ordering vitamin B1 salt，which can enhance the possibility that clients order simultaneously and greatly increase economy benefit.

8. CONCLUSIONS

The new business intelligence system, based on data warehouse and Web technologies, was developed with the sale data of medicine materials. Basal OLAP analysis was achieved via MDX enquiries. Managers can more know and discover sale knowledge by the association analysis function. It applied data warehouse, online analysis processing and association analysis technologies to practical system.The system provides effective support for decision-making of managers in medicine materials sales.

9. REFERENCES

[1] W.H.Inmon, “Building the Data Warehouse”, Third Edition, Beijing :China Machine Press,2004.

[2] J.M.Chen, “Data Warehouse and Data Mining Techniques”, Beijing: Publishing House Electronics Industry,2003.

[3] M.L.Gonzales, “IBM Data Warehousing With IBM Business Intelligence Tools”, Beijing: Publishing House Electronics Industry, 2004.

[4] X. Li, “Mining association rules on a huge number of items”, Proc of the 8th Joint International Computer Conference, 2002, 11:587-593.

Optimization for Supply Chain Configuration Design

Lihong Yuan
School of Information Technology, Southern Yangtze University,
No. 1800, Lihudadao, Wuxi, Jiangsu, China 214122

ABSTRACT

The design of a supply chain network as an integrated system with several tiers of suppliers is a difficult task. It consists of making strategic decisions on the facility location, stocking location, production policy, production capacity, distribution and transportation modes. This research develops a hybrid optimization approach to address the Supply Chain Configuration Design problem. The new approach combines simulation, mixed integer programming and Particle Swarm Optimization (PSO). The PSO provides a mechanism to optimize qualitative and policy variables. The mixed integer-programming model reduces computing efforts by manipulating quantitative variables. Finally simulation is used to evaluate performance of each supply chain configuration with non-linear, complex relationships and under more realistic assumptions. The approach is designed to be robust and could handle the large scale of the real world problems.

Keywords: Supply China Management, Optimization, Mixed Integer Program, Genetic Algorithm

1. INTRODUCTION

Managing a supply chain is very different from managing one site. Activities at the various sites have complex interrelationships. Design and optimization of supply chain configuration is a problem at the highest level, the strategic level. Supply chain configuration design consists of deciding on the facility location, stocking location, production policy (maketo-stock or make-to-order), production capacity (quantity and flexibility), assignment of distribution resources and transportation modes while imposing standards on the operational units for performance excellence. Therefore, the aim of supply chain configuration optimization is to find the best or the near best alternative configuration with which the supply chain can achieve a high level of performance. Usually, there are two categories of configuration decisions on supply chain design.

1. Structural decisions: Location, capacity, distribution channel.
2. Coordination decisions: Supplier selection, partnership, inventory ownership, sharing information about sales, demand forecast, production plan, and inventory.

All supply chain design decisions affect each other and must take this fact into consideration. Location decision has long-term impact. It is very expensive to shut down a facility or to move it to a different location. This decision also has direct effects on production, inventory and transportation costs. Those in turn have significant impact on supply chain performance, in terms of the service level, since a good distribution network can increase responsiveness.

Macroeconomic, political, strategic, technological, infrastructure, competitive, logistical and operational factors influence network design decisions in supply chain. Many other concerns need to be taken into account if the system is a global supply chain. Many kinds of company resources in a supply chain are duplicated. When collaborating or being integrated, companies can eliminate redundancies. Companies also improve the efficiencies through integration since even if a company has available resources to perform a particular task, another company in the supply chain may be better suited to perform that task. Determining who in the supply chain should perform a particular function is a part of supply chain configuration design.

2. Optimization of Supply Chain Configuration

This section presents a new general optimization approach developed for the Supply Chain Configuration Design. This approach enables designers to integrate varieties of strategic decisions encountered in supply chain development simultaneously, including qualitative and quantitative. In particular, the following seven types of decisions will be considered concurrently:

1. Make or Buy decisions (or Outsourcing decisions).
2. Partner selection.
3. Production planning policy at each stage.
4. Transportation mode at each distribution channel.
5. Location decisions.
6. Capacity decisions.
7. Production and service allocation.

Many of these problems have been studied separately for years but they have never been placed in one common framework and been considered at the same time. Optimal solutions for each individual decision do not necessarily create a significant improvement in overall supply chain performance, especially when they are limited to one or two stages of the supply chain. Overall supply chain efficiency is jointly determined by the structural configuration as well as system parameters, such as leadtime, responsiveness, quality, delivery reliability etc for each partner. Some companies acquire their partners to enhance integration and control. Others find it more profitable not to. More and more companies are outsourcing in an attempt to increase their competitiveness and focus on their core competencies. They must consequently rely more on their partners for product development, quality, productivity and technology. The challenge is to

determine which products, activities, and functions should be outsourced and which partners should be selected. Some strategic decisions in supply chain design are quantitative but most of the times they are of qualitative and policy nature. For instance, in the above list, the first five categories of decision variables are qualitative. Except for location decisions that could be encoded as binary variables and solved by using integer programming, other qualitative variables could not be expressed mathematically and therefore they require an optimization approach that can deal with policy variables.

3. Mixed Integer Program (MIP) Model

Given a decision vector that specifies values of decision variables in Category 2, a MIP model is constructed to obtain the optimal values of decision variables in Category 1. More specifically, the MIP model's results determine locations of plants/warehouses selected for each stage; capacities of each plant/warehouse; and define plants/warehouses that serve plants/warehouse in downstream stages, given the information about companies serving at each stage; production policy and transportation mode carried out by each company as provided by the optimization algorithm such as PSO employed in this paper. The verbal formulation of the MIP model can be stated as follows. The objective is to minimize the overall system-wide cost that includes fixed investment cost, variable operating cost, transportation cost, pipeline inventory cost, material/component inventory carrying cost and finished product/part inventory carrying cost, subject to customer demand satisfaction requirement, facility capacity constraint and conservation in material flows at each plant/warehouse. The MIP model is mathematically expressed by:

$$\begin{aligned} Min\ \ Total\_Cost = & \sum_{i\in N}\left[\sum_{l\in L_{C_i}} FixCost_{il}\cdot X_{il}\right] + \sum_{i\in N}\left[\sum_{l\in L_{C_i}} VarCost_{il}\cdot Capa_{il}\right] \\ & + \sum_{i\in N}\sum_{l\in L_{C_i}}\left[\sum_{j\in N_{C_i}^{-}}\sum_{k\in L_{C_j}} TransCost_{ijk}\cdot Y_{ijk}\right] \\ & + \sum_{i\in N}\sum_{l\in L_{C_i}}\left[\sum_{j\in N_{C_i}^{-}}\sum_{k\in L_{C_j}} InvCost_{il}^{1}\cdot TranLeadtime_{ijk}\cdot (Y_{ijk}/365)\right] \\ & + \sum_{i\in N}\left[\sum_{l\in L_{C_i}} InvCost_{il}^{2}\cdot Z_{\alpha}\cdot SupplyLTDeviation_{il}\cdot (Capa_{il}/365)\right] \\ & + \sum_{i\in N^{push}}\left[\sum_{l\in L_{C_i}} InvCost_{il}^{2}\cdot Z_{\alpha}\cdot \sqrt{ProdLeadtime_{il}}\cdot DemandDev_{il}\cdot (Capa_{il}/365)\right] \end{aligned} \quad (1)$$

Subject to: $Capa_{il} = Demand_{il}$

$$\forall i \in N_{Retailer}, l \in L_{C_i} \quad (2)$$

$$\sum_{j\in N_{C_i}^{+}}\sum_{k\in L_{C_j}} Y_{iljk} = Capa_{il}$$

$$\forall i \in N, l \in L_{C_i} \quad (4)$$

$$X_{il} = 0 \quad or \quad 1$$

$$\forall i \in N, l \in L_{C_i} \quad (5)$$

$$M_{upperLimit}\cdot X_{il} \geq Capa_{il}$$

$$\forall i \in N, l \in L_{C_i} \quad (6)$$

$$Capa_{il} \geq M_{LowerLimit_i}\cdot X_{il}$$

$$\forall i \in N, l \in L_{C_i} \quad (7)$$

Where

N	Set of stages in the supply chain
C_i	Company that fulfills stage I; specified by a position vector in PSO
L_{C_i}	Set of locations of company C_i
$N_{C_i}^{-}$, $N_{C_i}^{+}$	Set of down stream stages and up stream stages of stage *i* respectively.
$N_{Retailer}$	Set of retailers
N^{Push}	Set of companies that implement Push production planning policy
X_{il}	Binary variables, $X_{il}=1$ if location of l of company C_i opened at stage I; $X_{il}=0$ otherwise.
Y_{iljk}	Material flow from facility l of company C_i to facility k of company C_k
Z_{α}	Safety stock factor, depending on pre-defined customer service level . For example $Z_{95\%}=2$, $Z_{99.7\%}=3$.
$Demand_{il}$	Demand at location I of retailer C_i
β_{ji}	Ration of parts from stage C_j to produce one product at C_i
$DemandDev_{il}$	Demand deviation at location l of company C_i
$FixCost_{il}$	Demand deviation at location l of company C_i
$SupplyLTDeviation_{il}$	Supply leadtime deviation at location l of company C_i

$\Pr oLeadtime_{il}$	Average production leadtime at location l of company C_i
$InvCost^1$, $InvCost^2$	Unit holding cost for pipeline in inventory and carry over inventory
$M_{UpperLimit}, M_{LowerLimit}$	Upper bound and lower bound for facilities of company C_i
$TransCost_{iljk}$, $TransLeadtime_{iljk}$	Unit cost and average leadtime for transportation from location l of company C_i to location k of company C_j.

After the MIP model is solved and all parameters of the supply chain configuration are determined, the model generator developed in the previous chapter builds a simulation model for the system. The supply chain simulator is invoked to run the simulation model. The simulation model returns the overall long run system-wide cost and customer service level of the supply chain. In this paper, we employ Particle Swarm Optimization to solve the MIP model.

4. Particle Swarm Optimization

Particle Swarm Optimization (PSO) algorithm is a population-based optimization technique originally introduced by Kennedy and Eberhart in 1995 ([6]). A PSO system simulates the knowledge evolvement of a social organism, in which individuals (particles) representing the candidate solutions to the problem at hand fly through a multidimensional search space to find out the optima or sub-optima. The particle evaluate its position to a goal (fitness) at every iteration, and particles in a local neighborhood share memories of their "best" positions, and then use those memories to adjust their own velocities, and thus subsequent positions.

In the original PSO with M individuals, each individual is treated as a volume-less particle in the *D*-dimensional space, with the position vector and velocity vector of particle *i* at *k*th iteration represented as $X_i(k) = (X_{i1}(k), X_{i2}(k), \cdots, X_{iD}(k))$ and $V_i(t) = (V_{i1}(k), V_{i2}(k), \cdots, V_{iD}(k))$. The particles move according to the following equations:

$$V_{ij}(k+1) = w \cdot V_{ij}(k) + c_1 \cdot r_1 (P_{ij}(k) - X_{ij}(k)) + c_2 \cdot r_2 \cdot (P_{gj} - X_{ij}(k)) \quad (8)$$

$$X_{ij}(k+1) = X_{ij}(k) + V_{ij}(k+1)$$

$$i = 1,2,\cdots M \; ; \; j = 1,2 \cdots, D \quad (9)$$

Where c_1 and c_2 are called acceleration coefficient. Vector $P_i = (P_{i1}, P_{i2}, \cdots, P_{iD})$ is the best previous position (the position giving the best fitness value) of particle *i* called *personal best position*, and vector $P_g = (P_{g1}, P_{g2}, \cdots, P_{gD})$ is the position of the best particle among all the particles in the population and called global best position. The parameters r_1 and r_2 are random numbers distributed uniformly in (0,1), that is $r_1, r_2 \sim U(0,1)$. Generally, the value of V_{id} is restricted in the interval $[-V_{\max}, V_{\max}]$. Parameter w is intertia weight introduced to accelerate the convergence speed of algorithm [7].

4.1 Case Study

This section demonstrates and evaluates the performance of the proposed approach by applying it to a specific case study. The objective of the case is to design a supply chain for one product line that requires 11 manufacturing/ assembling stages (#0 to #10) as shown in Figure 6. Stage #11 and stage #12 are the distributor and the retailer respectively. The major manufacturer M in the supply chain has capability of performing most stages (#1, #3, #6, #7, #8, #9 and #10). The manufacturer M could carry out stages #1, #3 and #6 by itself or outsource them to outside vendors. If outsourcing is chosen, the manufacturer has to decide which vendors will be its partners. Similarly, supplier selection must be made at stages #0, #2, #4 and #5.

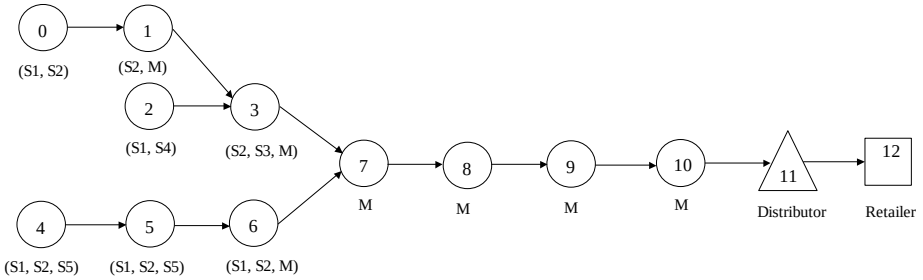


Fig. 1.

In Fig. 1, brackets list potential suppliers for the corresponding stages. For example, either supplier S1 or S2 could be selected for stage #0. However, they are different from each other in production cost, production leadtime, variance of production leadtime etc. Each company has several potential sites for consideration. They have to decide how many plants/warehouses are needed and where to locate them based on associated costs. The customer service levels at retail stores are required not to be less than 95%. The number of

alternatives for this test problem is tremendously large. It has 213 different combinations in terms of transportation modes; 211 different combinations in terms of production policy; 23.34 different combinations in terms of supplier/company selection (stages 0, 1 and 2 have 2 options; stages 3, 4, 5 and 6 have 3 options); 612 different combinations in terms of location (assume that 2 locations out of 4 are selected at each stage). The number of combinations is much higher than 213.211.23.34.612 since we have not yet taken into account service allocation possibilities (which plant serves which plants/warehouses).

To apply the optimization approach, decision variables are divided into two categories as mentioned in the previous section. The PSO undertakes decisions related to supplier/company selection, production policy selection and transportation mode selection at each stage. Each stage requires 3 genes representing these 3 decisions. The Pull production policy is encoded as 0; Push as 1; fast and expensive transportation mode as 0; slow and less reliable transportation mode as 1. PSO is executed with 20 particles for 50 iterations. The inertia weight is varying from 0.9 to 0.4 and c1, c2 are both set to 2.0.

A MIP model incorporates decisions of category 1 that relate to location selection, facility capacity and distribution decision. A MIP model is composed of 332 variables, including 56 binary variables, and 214 constraints.

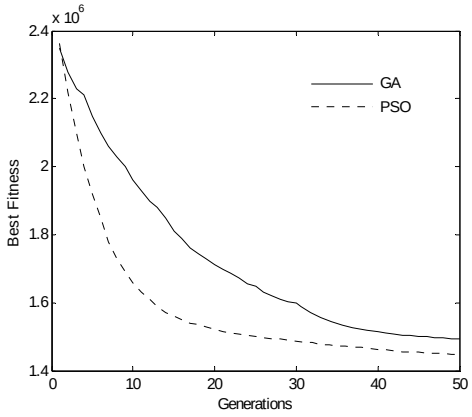


Fig. 2.

For each specific configuration, a simulation model is generated and is run for 700 days. The first 100 days is considered as the warm-up period and outputs in that period are discarded. Simulation outputs are the overall system-wide cost and the customer service level. We also test Genetic algorithm for performance comparison. Figure 2 presents results by PSO and GA after 50 generations. It can be seen that PSO can find the optima more quickly than GA.

5 CONCLUSIONS

A new optimization approach for Supply Chain Configuration Design problem is introduced and implemented. The proposed hybrid approach combines PSO, Mixed Integer Programming and Simulation into one framework. This research attempts to consider many strategic decisions simultaneously. Those decisions include Make or Buy decision (or Outsourcing decision), Partner selection, Production planning policy Inventory stock at each stage, Transportation mode at each channel, Location decision, Capacity decision and Production and service allocation. Computational results on a medium-size problem showed that this approach could result in efficient solutions.

Reference

[1] S. Andradottir, 1998. Chapter 9: Simulation Optimization. In Handbook of Simulation, ed. J. Banks. New York: John Wiley and Sons.

[2] B.C. Arntzen, G. G. Brown, T. P. Harrison, and L. Trafton. 1995. Global Supply Chain Management at Digital Equipment Corporation. Interfaces 25, 69-93.

[3] F. Azadivar, 1999. Simulation Optimization Methodologies. In Proceeding of the 1999 Winter Simulation Conference, ed. P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, 93-100.

[4] F. Azadivar, 1992. A Tutorial on Simulation Optimization. In Proceeding of the 1992 Winter Simulation Conference, ed. Robert C. Crain, James R. Wilson, James J. Swain, David Goldsman, 198-204.

[5] F. Azadivar, G. Tompkins, 1999. Simulation optimization with qualitative variables and structural model changes: A Genetic Algorithm approach. European Journal of Operational Research 113: 169-182.

[6 J. Kennedy, R. C. Eberhart, "Particle Swarm Optimization," Proc. IEEE Int'l Conference on Neural Networks, IV. Piscataway, NJ: IEEE Service Center, 1995, pp. 1942-1948.

[7] Y. Shi, R. Eberhart, "Empirical Study of Particle Swarm optimization," Proc. of Congress on Evolutionary Computation, 1999, 1945-1950

[8] S. Bagchi, S. J. Buckley et al, 1998. Experience Using The IBM Supply Chain Simulator. In Proceedings of the 1998 Winter Simulation Conference, ed. D.J. Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan, 1387-1994.

Partition Technology and Application in Large Scale Population Information System *

Zhaoxia Lu, Jie Xi
Software Center, Northeastern University
Shenyang, Liaoning 110004, China
Email: luzx@neusoft.com
Email: xijie777@hotmail.com
And
Jian Wang
College of Information Science of Engineering, Northeastern University
Shenyang, Liaoning 110004, China
Email: wangjian@mail.neu.edu.cn

ABSTRACT

This paper proposes a solution to improve the efficiency of querying and processing large-scale data using partition technologies. The database physical design, partitioning principle and strategy are presented. With a Nation-wide Population Information System as a case study, three partitioning strategies are rigorously analyzed to fulfill the requirements of the real-world system. According to each partitioning strategy, a method of partitioning tables and index is also presented. The whole solution is validated in a simulation environment. The simulation results demonstrate the practical usage of the solution proposed. Moreover, the experience from this project provides guidelines of applying partition technology for other similar systems that have querying requirement on large-scale data.

Keywords: Partition Technology, Partitioned Tables, Optimization, Large Scale Population System.

1. INTRODUCTION

Mass data processing is in high demand and critical for business. The traditional database methodologies cannot meet the requirements, especially in some super large-scale database like Country-Wide Population Information System. The management of big tables must be simplified to improve the response of querying and DML command. The high availability and efficiency of data backup, recovering and migrating should also be considered. Thus, it is a challenging issue to use data stored in large tables and the way to physically store the data.

2. PHYSICAL DESIGN OF DATABASE AND PARTITION TECHNOLOGY

Data storage is based on the physical design of database. Physical design of database should be conducted through compound of logical design and specific database management platform. Physical design principals are as following:

- Reasonable design of table space to decrease I/O collision of data files.
- Characteristic design for table spaces and tables to take best advantages of the db's advanced management features.
- Partition design for super large-scale tables and indexes. Partitioned table and index table can be used to take data apart, improve system performance and decrease I/O.
- Design for size and location of on-line log files and archive files.
- The definition of system hardware and mirrors architecture. Appropriate file systems, naked devices and RAID mirrors are chosen to meet the reliability requirements.

Scalability and manageability of large-scale tables and those with different access schemas can be improved by using partitioning technology. Optimizing program directly uses specific partition because it knows ranges and base data of partition. Performance is improved by limiting data retrieved during the querying process.

There are several partitioning methods: Range partitioning, Hash partitioning, List partitioning, Composite range-hash partitioning, and Composite range-list partitioning. Indexes, as well as tables, can be partitioned. Local indexes (created using LOCAL attribute) that are partitioned in the same way as the associated table. Global indexes (created using the GLOBAL attribute) that can be partitioned in any way.

Partitioned tables allow data to be broken down into smaller, more manageable pieces called partitions, or even sub-partitions [1]. Indexes can be partitioned in similar way. Each partition is stored in its own segment, and can be managed individually. Individual partitions can function independently from each other, thus they provide a structure that can facilitate the tuning for improved availability and performance [2, 3].

3. PARTITIONED TABLE DESIGN PRINCIPALS AND APPLICATION IN LARGE SCALE POPULATION SYSTEM

The design of partitioned tables follows the principles devised in: First, when table size is over 1.5G or 2G, or tables contain over 10 million records in OLTP system, the tables should be considered to be partitioned. Second, partitioning can be considered if querying function only focuses on some limited scale data. This excludes the irrelevant data access and thus reduces unnecessary queries to further improve response [4]. Third, in some cases, data maintenance needs to partition tables. In some circumstance tables need to be cleaned and delete records periodically [5, 6]. Deleting mass records is an expensive operation. Some read-only data can be stored in read-only table spaces through partitioning [7]. Fourth, parallel DML, for example

* Supported by National 863 High-tech Program (2004AA1Z2370)

parallel update, often needs partitioning. Fifth, partitioning is often considered when we require that some table records are highly available [8, 9].

In China countrywide population system, records in some table reach the billion-level and query in the large table is a challenging issue. To our knowledge, there are no other solutions to solve this problem efficiently. The performance is very important in the design of database objects and physical storage. After consideration of the data scale and data processing specific functionality, we provided the following partitioning design advices:

The first solution is double partitioned by the first two district digits (province code) and the last two digits of ID card (province_id, nn). This method brings convenience for querying and data will be under the million levels. However this solutions has its disadvantages, that is first too many partitions over 3000 incurs extra problem of the maintenance; second, data not distributed evenly into partitions and different access frequency in partitions will result in hot-spot of disk. These scenarios can incur performance bottleneck.

The second solution is partitioned by birth date (YMMDD: Y means the last digit of years; MM means month; DD means date) and there are total 3660 partitions. From the population characteristic, this partition solution has the following advantages: data evenly distributed into these partitions; data capacity is under 0.4 million (1.4 billion population in China and average 0.38 million every partition); data distributed into partitions also makes it convenient for physical distribution and decrease the probability of hot spot.

The third solution is partitioned by birthday (MMDD) or double partitioned by birthday (YYYY, MM). But this design is rarely used and it has disadvantages as well.

Basic information querying is mostly based on citizen's name and ID number. Two partition indexes are created for better performance: Local Non-Prefix Partition Index on IDs and Global Partition Index on names. Partition in local index is mapped to corresponding one in the basic information table. Partition in global index is mapped to multi-partitions in basic information table.

First of all, global index of basic information partition (JBXX_HM_XQ_INDX) by names is created and partitions will be created then by ASCII name. According to the Chinese name's specific character, family name with large numbers will be partitioned in more fine grained level which makes it possible to have even distribution of these partitions. 3658 partitions are evenly distributed into 30 index table spaces. This policy will dramatically improve querying based on names. Secondly, local non-prefix partition index by ID number (JBXX_XM_SSXQ_PARTED) is created. Like the prior policy 3660 partitions are distributed into 30-partition index table space.

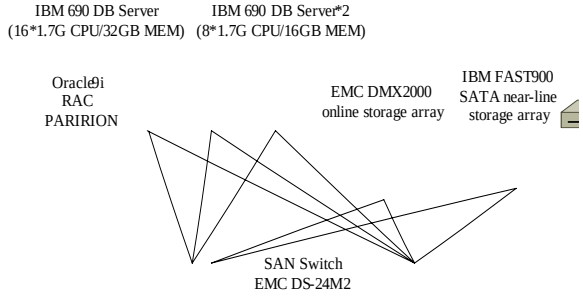
4. COMPARISON OF PARTITION PRINCIPALS AND EFFICIENCY VALIDATION

Experiments on the above two solutions are carried out to compare the efficiency of them in term of performance. The testing tables, executing statistic, query process and steps are provided as followed:

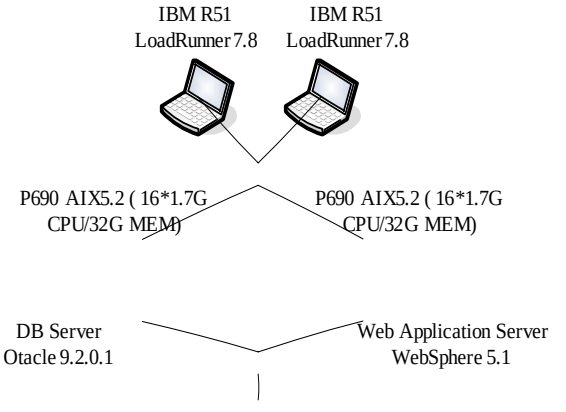
- **Partitioned by administrative district**. Table JBXX_LOAD_3000_NEW is created and partitioned by 3000 ranges. We produce testing data for 300 ranges and about 400,000 records for each range. The maximum in one range is about 7,280,000 records and the minimum is about 240,000. The total number of records is 112,452,795. Three threads are executed in parallel and finished one range of data processing in 13 seconds averagely.
- **Partitioned by birth date**. Table JBXX_LOAD_3660_NEW is created and partitioned by birthday (YMMDD). There are 3660 ranges and testing data created for 650 ranges. Each range has about 400,000 records. The total number of testing records is 246,988,157. If three threads are executed in parallel and finished one range of data processing in 4 seconds. If one thread is executed and finished in 7.2 seconds for one range data processing. This also proves that multi-threads improve the querying response.

In summary, for the design of **Partitioned by birth date,** the average execution time is approximately two third of the design of **Partitioned by administrative district**. However the size of its testing data is doubled of the design of **Partitioned by birth date**. So the better partition solution for population system is the design **Partitioned by birth date**. Through the testing we demonstrate that partitioned by last 5 digit of birth date of citizen is more feasible and acceptable.

We further validate the performance of this partition policy (by birth date), the production system is depicted in Fig. 1.(a) and we estimate data quantity of the future system. The testing environment is shown in Fig. 1.(b).



(a) architecture design of population information system



(b) test environment of population information system

Fig. 1. Architecture design and test environment

In the testing environment shown in Fig. 1., the SQL's responding time is checked for different data quantity level and this can be used to analyze the impact of data quantity on performance. Different SQL statements are used including Select, Insert, Delete and Update statement. The sampling testing data consist of 43,000,000 records with pictures, 0.1 billion composite records with pictures and 0.19 billion composite records with pictures and the details

are illustrated in Table 1..

Table 1. Test case details

Test Case	Description	Notes
1. Select	Querying based on ID number	1.Picture/Records=100% in 40,000,000 records.
2. Update	Updating based on ID number	2.Picture/Records=39% in 0.1 billion records (basic information 0.29k, picture 7k)
3. Delete	Deleting based on ID number	3. Picture/Records=22% in 0.19 records.
4. Insert	Inserting based on ID number	

To further validate the efficiency of this partition design solution, the testing is proceeded by 3 levels of quantity: 43,000,000, 1 billion and 1.9 billion. 40 million image data are included in each testing level. The performance and testing result is as Fig. 2..

Testing data(row)	43829992	109799992	190649992
Query(s)	0.0163	0.0171	0.0183
Insert(s)	0.0708	0.0712	0.0717
Delete(s)	0.015	0.0175	0.021
Modify(s)	0.0167	0.02	0.03

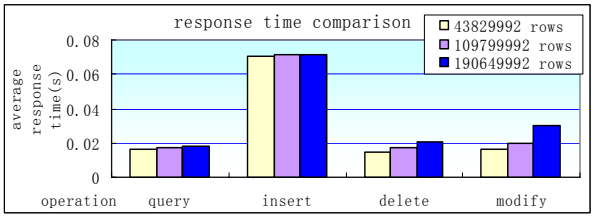


Fig. 2. Response time comparison

From the above results we can conclude that performance of query by SQL meets the requirement in the simulation environment. For those tables partitioned by month and data, the query whose criteria compound of partitioning column field and digits of ID respond in almost same duration no matter how many records the table contained. This indicates that the amount of data doesn't have significant effect on the performance. When testing with 0.19 billion records, the simple record querying respond time is under 0.05 second.

Concurrent	100	150	200	250	300
Response time(s)	1.321	1.462	1.48	1.688	2.042
Throughput	458141	425650	632111	628798	674189
Hit rate per second	107	140	148	148	158

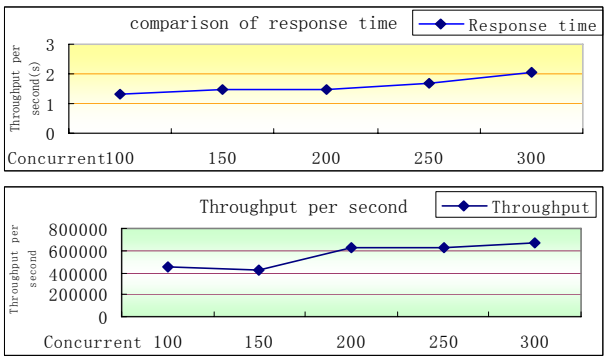


Fig. 3. Comparison of response time

The followed testing result is collected by LoadRunner tools in simulation environment, which processed the data include 43,000,000 images and result is as Fig. 2..

Locating for record in one range does not change by time at every testing point. About 0.1 second for locating one record in one range. From the Fig. 3. we can conclude that query response time does not grow by increasing the size of the testing data.

5. CONCLUSIONS

This paper proposes the partition solution for Country-Wide Population Information System. It provides two partitioning techniques and evaluates them in simulation environment. In actual China countrywide population information system the partitioning strategy by last 5 digits of birth date of citizen ID is adopted and proved its efficiency. Population information system includes information loading business and information maintenance business. The former benefit from the partition technology, while the latter is more relevant to the administrative constraints. This remains our further research work to determine which partition technique suits bests for the administrative constraints.

6. REFERENCES

[1] Oracle 9i DBA Handbook, Beijing, Mechanic Industry Press, 2004.

[2] J.H. Yin, "ORACLE Database Optimization and Tuning Design", Gas field Engineering, September 2005, 24(17): 47-48.

[3] F.Z. Jiang, "Performance Optimizing Technology of Oracle Database", Computer Engineering, Supplementary Issue, December 2004, Vol.30: 94-96.

[4] P. Yin, "The Functional Optimization of Sql Server Database", Computer Applications and Software, March 2005, 22(3): 58.

[5] Q.Y. Zheng, T. Ju, "Research and Application on Performance Optimization of Sybase SQL Server", Microcomputer Development, January 2003, 13(1): 85-86.

[6] G.H. Yue, "Some Technique Countermeasure to Improve on Oracle8i Database Answer Speed", Computer Applications and Software, May 2004, 21(5): 110-112.

[7] S.J. Huang, C.M. Li, "Processing massive data in ORACLE", Practices, December 2003, Vol.12: 71-72.

[8] T. Zhang, "Performance tuning analyze of Oracle, China Technology Information", China Science and Technology Information, May 2005, pp.17.

[9] F. Fang, "Logical design optimization of database", Communication and Broadcasting Television, March 2004, 64-69.

Nonlinear System Identification Using DAFNN

Ming Li
College of Power Engineering, Nanjing University of Science & Technology
Nanjing, 210094, China
E-mail: lm_2001@tom.com
Cheng-wu Yang
College of Power Engineering, Nanjing University of Science & Technology
Nanjing, 210094, China
E-mail: yangcw@njust.edu.cn

ABSTRACT

A new artificial neuron model was presented according to the mechanism of biological neurons. Because of a dynamic activation function, it was regarded as a dynamic nonlinear mapping. A new network called DAFNN, the abbreviation of Dynamic Activation Function Neural Network, of such neurons was also proposed. It could be used as dynamic plant identifiers or controllers. The performance of DAFNN was verified by comparing it with PIDNN, a neural network is composed of P, I and D neurons, and NARXNN, also called Nonlinear AutoReguressive eXgenous input filter. Simulation results show the performance of DAFNN was better than that of PIDNN. Without adding computation complexion, DAFNN could replace NARX recurrent neural network in nonlinear plant modeling and identification.

Keywords: Neuron Model, Activation Function, Nonlinear System Identification.

1. INTRODUCTION

The current emphasis in neural network research is on learning algorithms and architectures neglecting the importance of neuron models. In fact, choice of neuron models is considered by some experts to be as important as the network architecture and learning algorithm [1,2]. The function of neurons is mainly determined by its transfer function, a combination of the activation and output function. Because only are dynamic neurons and networks discussed, some static neurons (e.g., see [3,4,5]) have been omitted here. Dynamic neurons are usually regarded as dynamic nonlinear mappings from input space to output space. PID neuron models interest us for its simple transfer function [6]. PID neurons have three types: proportion neuron, integral neuron and differential neuron. Practically, PID network has a poor performance when used for modeling and identification for nonlinear dynamic systems. It is partially due to lack diversity of neuron, i.e. lack enough neuron types. According to the neurology bases, we will present a more flexible and multiform new neuron model, whose activation function has dynamic nature. Networks of such neurons, is called Dynamic Activation Function Neural Network (DAFNN), have the same architecture as PID networks, but they have better dynamic performance of identification for nonlinear dynamic systems in a more effective way.

The remainder of this paper is organized as follows. Section 2 introduces a new neuron model and briefly reviews its biological bases. Section 3 proposes a new dynamic neural network and its learning algorithm. Some simulation results of identification for dynamic nonlinear systems are shown in Section 4 to verify the performance of DAFNN. Finally, a conclusion is given in Section 5.

2. A NEW NEURON MODEL

The brain is a collection of about 10 billion interconnected neurons. Each neuron is a cell that uses biochemical reactions to receive process and transmit information. A single neuron receives information form a thousand neighboring neurons by its dendritic tree. After integrating, the aggregate input is then passed to the soma (cell body). If it is greater than the axon hillock's threshold value then the neuron fires and an output signal is generated and transmitted by its synapse.

To establish an artificial neuron model, it is most important to knowing the integration procedure of input information of biological neurons. There are three possible methods to deal with the input signals in biological neurons [6]:

(1) Spatial summation converts several weak signals into a single large one.

(2) Temporal summation converts a rapid series of weak pulses into one large signal.

(3) Shortening the span of two input signals may heighten the response velocity of output.

The above three biological hypotheses also are the bases of PID neurons, which is consist of proportion neuron (P neuron), integral neuron (I neuron) and differential neuron (D neuron). Each type has one function corresponding to one hypothesis respectively. Cascading a fan-in function, used as an activation function, and a multi-step function, used as an output function produces the transfer function of the PID neurons. There is a feedback from output of output function to its input in the I neuron. The input of output function of D neuron consists of two terms, the output of activation and its one order delay. However, a biological neuron may use above three methods simultaneously, i.e. one neuron may have three methods for processing the input information simultaneously. Because each neuron has only one method, PID neurons are not suitable representation of real ones. A new neuron model is proposed to overcome the drawbacks of PID neurons here.

The activation function can be regarded as the processing to integrate input signals and then generate a net input to soma. The output function generates an output signal of a neuron. According the above

hypotheses, the net input of a neuron is a combination of current and former net inputs. So the activation function ought to be a dynamic map. A sigmoid function usually is chosen as the output function for its good squashing functions for unbounded activation. Transfer function of the new neuron consists of a dynamic activation function and a sigmoid output function.

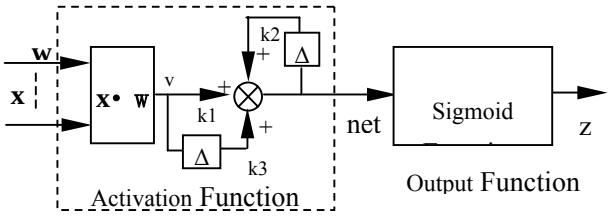


Fig. 1. Transfer function of a new neuron

Fig. 1 shows a block-diagram of the transfer function of a new neuron. x is the input vector and w is the weight vector. v is the inner conduct between x and w vectors. net is the current output of the activation function, also the net input to the neuron. z is the output of the neuron. The whole mapping from input to output is describing as follows:

$$\begin{cases} v(k) = \mathbf{x}(k) \cdot \mathbf{w}(k) \\ net(k) = k1 \cdot v(k) + k2 \cdot net(k-1) + k3 \cdot v(k-1) \end{cases}$$

(1)

$$z(k) = f(net(k))$$

(2)

The first equation is the activation function, whose output is the combination of former output, current and former weighted input sum three terms, K2, k1 and k3 are connection weights respectively. Having a feedback form output to input, the activation function is a dynamic mapping. Eq. (2) is the sigmoid output function, for instance hyperbolic tangent function.

3. DAFNN AND ITS LEANING ALGORITHM

DAFNN has similar feedforward architecture of MPLs. However, DAFNN is a dynamic network for its dynamic neurons in hidden layer. Another such dynamic network is PID neural network, whose neurons are PID ones in hidden layer. Architecture of a DAFNN is shown in Fig. 2, which is a single output network. The multi output one is the simple expansion of it. **U** is the input vector and $\mathbf{w}^1$ is connection weight vector from input to hidden layer. **V** and **z** are the input and output vectors of hidden layer respectively. $\mathbf{w}^2$ is the connection weight vector from hidden to output layer. Linear functions are used for input and output layer.

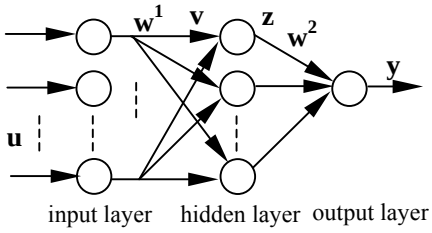


Fig. 2. The new neural network-DAFNN

Being a dynamic neural network, DAFNNs can be used for identification and control for dynamic nonlinear systems on-line or off-line. The classical gradient-based BP learning algorithm [7] is applied to train the network.

The neural network output is calculated as

$$\begin{cases} \mathbf{v(k)} = \mathbf{u(k)} \cdot \mathbf{w^1(k)} \\ \mathbf{net(k)} = \mathbf{k1} \cdot \mathbf{v(k)} + \mathbf{k2} \cdot \mathbf{net(k-1)} + \mathbf{k3} \cdot \mathbf{v(k-1)} \\ \mathbf{z(k)} = (e^{\mathbf{net(k)}} - e^{-\mathbf{net(k)}}) / (e^{\mathbf{net(k)}} + e^{-\mathbf{net(k)}}) \\ y(k) = \mathbf{z(k)} \cdot \mathbf{w^2}(k) \end{cases}$$

(3)

Where **k1**, **k2** and **k3** are three adjustable parameter vectors of activation functions of hidden neurons.

3.1 Off-line Learning Algorithm of DAFNN

In this paper, hyperbolic tangent functions are selected as the output functions of neurons in hidden layer. There are five adjustable parameter vectors, weight vector $\mathbf{w}^1$ and $\mathbf{w}^2$, the vector **k1**, **K2** and **k3**, which updated by BP algorithm.

The cost function using the familiar "sum of squared error" method is

$$J(k) = \frac{1}{2} \sum_{p=1}^{L} e_p^2(k)$$

$$e_p(k) = d_p(k) - y_p(k)$$

(4)

Where $e_p(k)$ is the network error computed as the desired response $d_p(k)$ minus the actual neural network output $y_p(k)$. L is the number of training samples and p is the index of them.

The weight $\mathbf{w}^2$ update is calculated as

$$e_p(k) = d_p(k) - y_p(k)$$

(5)

$$\frac{dJ(k)}{d\mathbf{w}_i^2(k)} = -\sum_p e_p(k) \cdot \frac{dy_p(k)}{d\mathbf{w}_i^2(k)} = -\sum_p e_p(k) \cdot \mathbf{z}_{p,i}(k) \qquad (6)$$

$$\mathbf{w}^2{}_i(k+1) = \mathbf{w}^2{}_i(k) - \eta_1 \cdot \frac{dJ(k)}{d\mathbf{w}^2{}_i(k)} \tag{7}$$

where η_1 is the small positive constant called the learning rate.

The **k1**, **k2** and **k3** vectors update is calculated as

$$\begin{aligned}\frac{dJ(k)}{d\mathbf{k1}_i(k)} &= -\sum_p e_p(k) \cdot \frac{dy_p(k)}{d\mathbf{k1}_i(k)} \\ &= -\sum_p e_p(k) \cdot \frac{dy_p(k)}{d\mathbf{z}_{p,i}(k)} \cdot \frac{d\mathbf{z}_{p,i}(k)}{d\mathbf{net}_{p,i}(k)} \cdot \frac{d\mathbf{net}_{p,i}(k)}{d\mathbf{k1}_i(k)} \\ &= -\sum_p e_p(k) \cdot \mathbf{w}^2_{p,i}(k) \cdot \left(1 - \mathbf{z}^2{}_{p,i}(k)\right) \cdot \mathbf{v}_{p,i}(k)\end{aligned} \tag{8}$$

$$\begin{aligned}\frac{dJ(k)}{d\mathbf{k2}_i(k)} &= -\sum_p e_p(k) \cdot \frac{dy_p(k)}{d\mathbf{k2}_i(k)} \\ &= -\sum_p e_p(k) \cdot \frac{dy_p(k)}{d\mathbf{z}_{p,i}(k)} \cdot \frac{d\mathbf{z}_{p,i}(k)}{d\mathbf{net}_{p,i}(k)} \cdot \frac{d\mathbf{net}_{p,i}(k)}{d\mathbf{k2}_i(k)} \\ &= -\sum_p e_p(k) \cdot \mathbf{w}^2_{p,i}(k) \cdot \left(1 - \mathbf{z}^2{}_{p,i}(k)\right) \cdot \mathbf{net}_{p,i}(k-1)\end{aligned} \tag{9}$$

$$\begin{aligned}\frac{dJ(k)}{d\mathbf{k3}_i(k)} &= -\sum_p e_p(k) \cdot \frac{dy_p(k)}{d\mathbf{k3}_i(k)} \\ &= -\sum_p e_p(k) \cdot \frac{dy_p(k)}{d\mathbf{z}_{p,i}(k)} \cdot \frac{d\mathbf{z}_{p,i}(k)}{d\mathbf{net}_{p,i}(k)} \cdot \frac{d\mathbf{net}_{p,i}(k)}{d\mathbf{k3}_i(k)} \\ &= -\sum_p e_p(k) \cdot \mathbf{w}^2_{p,i}(k) \cdot \left(1 - \mathbf{z}^2{}_{p,i}(k)\right) \cdot \mathbf{v}_{p,i}(k-1)\end{aligned} \tag{10}$$

$$\begin{cases} \mathbf{k1}_i(k+1) = \mathbf{k1}_i(k) - \eta_2 \cdot \dfrac{dJ(k)}{d\mathbf{k1}_i(k)} \\ \mathbf{k2}_i(k+1) = \mathbf{k2}_i(k) - \eta_2 \cdot \dfrac{dJ(k)}{d\mathbf{k2}_i(k)} \\ \mathbf{k3}_i(k+1) = \mathbf{k3}_i(k) - \eta_2 \cdot \dfrac{dJ(k)}{d\mathbf{k3}_i(k)} \end{cases} \tag{11}$$

The weight $\mathbf{w}^1$ update is calculated as

$$\begin{aligned}\frac{dJ(k)}{d\mathbf{w}_{i,j}^1(k)} &= -\sum_p e_p(k) \cdot \frac{dy_p(k)}{d\mathbf{w}_{i,j}^1(k)} \\ &= -\sum_p e_p(k) \cdot \frac{dy_p(k)}{d\mathbf{z}_{p,j}(k)} \cdot \frac{d\mathbf{z}_{p,j}(k)}{d\mathbf{net}_{p,i}(k)} \cdot \frac{d\mathbf{net}_{p,i}(k)}{d\mathbf{v}_{p,i}(k)} \cdot \frac{d\mathbf{v}_{p,i}(k)}{d\mathbf{w}_{i,j}^1(k)}\end{aligned} \tag{12}$$

$$\frac{d\mathbf{net}_{p,j}(k)}{d\mathbf{v}_{p,j}(k)} = \mathbf{k1}_j(k) + \mathbf{k2}_j(k) \cdot \frac{d\mathbf{net}_{p,j}(k-1)}{d\mathbf{v}_{p,j}(k-1)} \tag{13}$$

$$\frac{dJ(k)}{d\mathbf{w}_{i,j}^1(k)} = -\sum_p e_p(k) \cdot \mathbf{w}_j^2(k) \cdot (1 - \mathbf{z}^2{}_{p,j}(k)) \cdot \frac{d\mathbf{net}_{p,j}(k)}{d\mathbf{v}_{p,j}(k)} \cdot \mathbf{u}_{p,i}(k) \tag{14}$$

$$\mathbf{w}^1{}_{i,j}(k+1) = \mathbf{w}^1{}_{i,j}(k) - \eta_1 \cdot \frac{dJ(k)}{d\mathbf{w}^1{}_{i,j}(k)}$$

(15)

where adopts a same learning rate as $\mathbf{w}^2$ update.

3.2 On-line Learning Algorithm of DAFNN

The on-line learning algorithm is similar to off-line learning algorithm but the cost function is replaced by temporal form.

The cost function and the adjustable vectors updates are calculate as follows.

$$\begin{cases} J(k) = \dfrac{1}{2} \cdot (d(k) - y(k))^2 \\ e(k) = d(k) - y(k) \end{cases} \tag{16}$$

$$\frac{dJ(k)}{d\mathbf{w}_i^2(k)} = -e(k) \cdot \mathbf{z}_i(k) \tag{17}$$

$$\mathbf{w}^2{}_i(k+1) = \mathbf{w}^2{}_i(k) - \eta_1 \cdot \frac{dJ(k)}{d\mathbf{w}^2{}_i(k)} \tag{18}$$

$$\frac{dJ(k)}{d\mathbf{k1}_i(k)} = -e(k) \cdot \mathbf{w}^2{}_i(k) \cdot (1 - \mathbf{z}_i^2(k)) \cdot \mathbf{v}_i(k) \tag{19}$$

$$\frac{dJ(k)}{d\mathbf{k2}_i(k)} = -e(k) \cdot \mathbf{w}^2{}_i(k) \cdot (1 - \mathbf{z}_i^2(k)) \cdot \mathbf{net}_i(k-1) \tag{20}$$

$$\frac{dJ(k)}{d\mathbf{k3}_i(k)} = -e(k) \cdot \mathbf{w}^2{}_i(k) \cdot (1 - \mathbf{z}_i^2(k)) \cdot \mathbf{v}_i(k-1) \tag{21}$$

$$\begin{cases} \mathbf{k1}_i(k+1) = \mathbf{k1}_i(k) - \eta_2 \cdot \dfrac{dJ(k)}{d\mathbf{k1}_i(k)} \\ \mathbf{k2}_i(k+1) = \mathbf{k2}_i(k) - \eta_2 \cdot \dfrac{dJ(k)}{d\mathbf{k2}_i(k)} \\ \mathbf{k3}_i(k+1) = \mathbf{k3}_i(k) - \eta_2 \cdot \dfrac{dJ(k)}{d\mathbf{k3}_i(k)} \end{cases}$$

(22)

$$\frac{d\mathbf{net}_j(k)}{d\mathbf{v}_j(k)} = \mathbf{k1}_j(k) + \mathbf{k2}_j(k) \cdot \frac{d\mathbf{net}_j(k-1)}{d\mathbf{v}_j(k-1)}$$

(23)

$$\frac{dJ(k)}{d\mathbf{w}^1_{i,j}(k)} = -e(k) \cdot \mathbf{w}^2_j(k) \cdot (1 - \mathbf{z}^2_j(k)) \cdot \frac{d\mathbf{net}_j(k)}{d\mathbf{v}_j(k)} \cdot \mathbf{u}_i(k)$$

(24)

$$\mathbf{w}^1_{i,j}(k+1) = \mathbf{w}^1_{i,j}(k) - \eta_1 \cdot \frac{dJ(k)}{d\mathbf{w}^1_{i,j}(k)}$$

(25)

4. SIMULATION RESULTS

Two simple nonlinear SISO systems were chosen to demonstrate the principles of this paper. The collection includes both with and without disturbance plants. Since both are not motivated by any particular "real" dynamic system, the command signal and disturbance sources are artificial as well. In order to verify the performance of DAFNN used for identification for dynamic nonlinear systems, DAFNN was compared with PIDNN and NARXNN [8], a typical recurrent neural network used for identification for dynamic plants, in both examples.

System I: the first plant was a nonlinear system without disturbance in the output. The difference equations defining its dynamics are

$$x = 0.6\sin(2\pi k/30) - 0.4\sin(2\pi k/60)$$

$$y(k) = 5 \cdot \frac{y(k-1)}{2.5 + y^2(k-1)} + x^3(k-1)$$

System II: the second plant was a nonlinear system with disturbance in the output. The difference equations defining its dynamics are

$$s(k) = \frac{s(k-1)}{1 + s^2(k-1)} + u^3(k-1)$$

$$y(k) = s(k) + dist(k)$$

The input signal u(k) being i.i.d. Uniformly distributed between [0,1]. Disturbance was a first-order Markov process, generated by filtering a primary random process of i.i.d. Random variables, which were uniformly distributed in the range [-0.5, 0.5]. The filter was a one-pole filter with the pole at z=0.99. The disturbance was added directly to the output of the system.

In both examples, NARXNN was a $N_{(3,1)\,:5:1}$ network. This means the input of network is comprised of a tapped delay line with three delayed copies of the exogenous input vector and one delayed copy of the output vector. Furthermore, there are five neurons in the hidden layer and a single neuron in the output layer.

In off-line method, 60 input-output pairs were selected as training samples and trained 1000 iterations one time in both examples. The three networks were trained 30 times with different random initial values. In on-line method, both systems were running 9000 iterations. All learning rates are 0.01.

Fig. 3 is the off-line identification curves of three neural networks used as identification for system I. Homologous graphs for system II are shown in fig. 4. Fig. 5 and fig. 6 are the average MSE curves of three neural networks in on-line case of system I and II receptively.

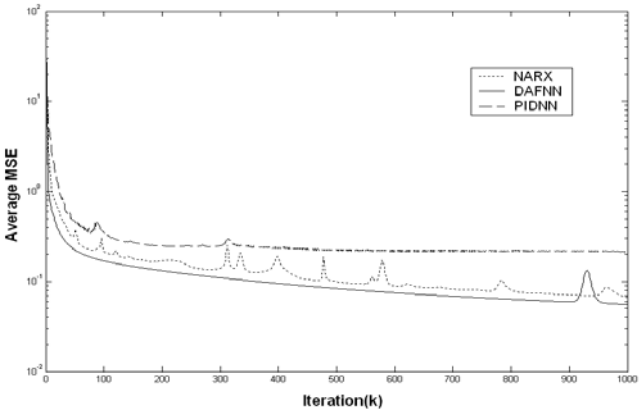


Fig. 3. Nonlinear plant off-line identification for system I

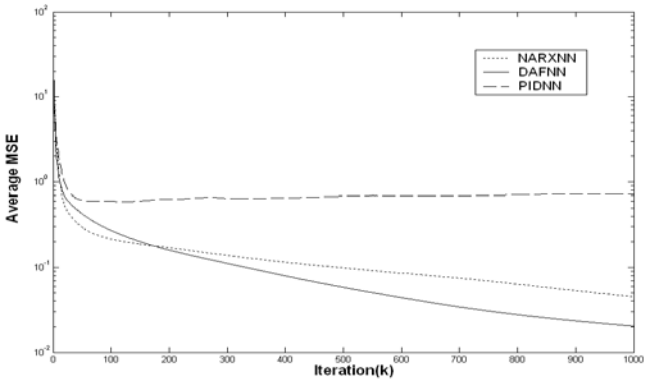


Fig. 4. Nonlinear plant off-line identification for system II

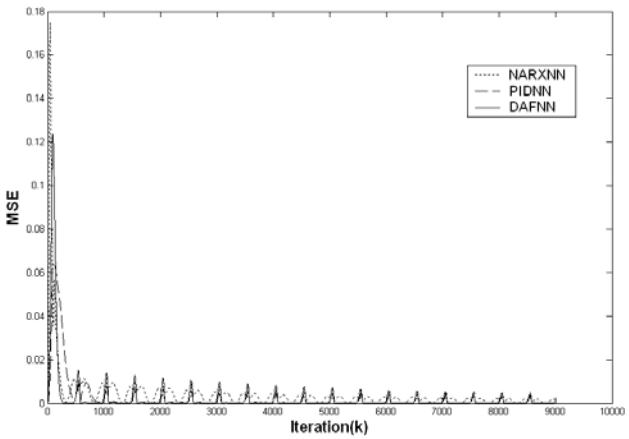


Fig. 5. Nonlinear plant on-line identification for system I

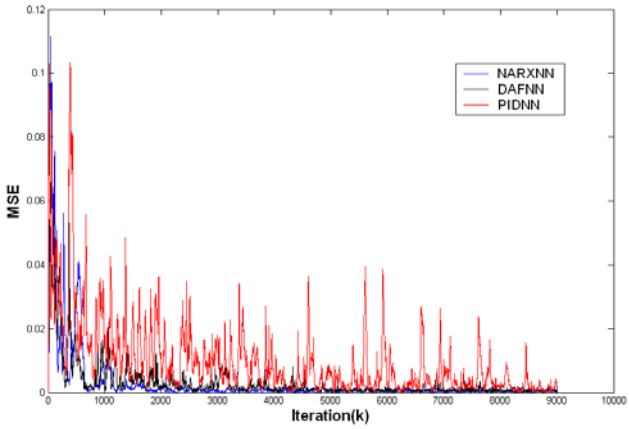


Fig. 6. Nonlinear plant on-line identification for system II

Simulation results show DAFNN has a better performance than that of PIDNN, especially when output of system was added a certain disturbance. In both systems, the number of adjustable parameters is same in DAFNN and NARXNN. So without increasing computation complexity, the performance of DAFNN is as good as that of NARXNN. However, DAFNN is more suitable for identification for nonlinear dynamic systems than NARXNN. Firstly, NARXNN is a recurrent network but DAFNN is non-recurrent one. So DAFNN has a simpler architecture. Secondly, three type variables need to determinate in NARXNN, which are the number of nodes in hidden layer and the delay order of input and output. Only the number of neurons in hidden layer needs to determinate in DAFNN.

5. CONCLUSION

A new neuron model proposed in this paper has a dynamic activation function, so it can be regarded as a dynamic nonlinear mapping from input to output space. Because the neurons in hidden layer are such neurons, DAFNN is a feedforward neural network with the dynamic natures. Such network is more suitable for identification for dynamic nonlinear systems than PIDNN. DAFNN has same good performance as NARXNN in practice. DAFNN is considered as a replacement for NARXNN in identification and control for nonlinear systems.

6. REFERENCES

[1] D.Wlodzislaw, J.Norvbert, “Survey of neural transfer functions,” Neural Computing Surveys, vol. 2, 1999,163-212, 1999.

[2] Y.S.Wu, M.S.Zhao, et al, “Beyond weights adaptation: a new neuron model with trainable activation function and its supervised learning,” in Proc. IEEE International Conference on Neural Networks, vol. 2, 1997, 1152-1157

[3] G. Cybenko, “Approximation by superpositions of a sigmoidal function”, Mathematics of Control, Signals and Systems, 2(4), 1989, 303–314.

[4] E. J. Hartman, J. D. Keeler, et al, “Layered neural networkswith gaussian hidden units as universal approximations”, Neural Computation, 1990,2(2): 210–215

[5] G. Dorffner, “A unified framework forMLPs and RBFNs: Introducing conic section function networks”, Cybernetics and Systems, 1994,25(4): 511–554.

[6] S.H.Lin, PID Neural Networks and its control systems, National Defense Industry Press, Beijing, 2006.

[7] H.R.Nielson, “Theory of the backpropagation neural network”, IEEE IJCNN, 1989,593-606.

[8] H.T.Siegelmann, B.G.Home, et al, “Computational capabilities of recurrent NARX neural networks,” IEEE Transactions on Systems, Man and Cybernetics-Part B, 1997,27(2): 208-215

Chaos Synchronization between Two Different Chaotic Dynamical Systems Using Nonlinear Control

Wei Wei [a, b], Lei Wang [a]
[a] Department of Mathematics, Huazhong University of Science and Technology
Wuhan, Hubei, 430074, P.R. China
[b] Department of Computer, Luoyang Technology College
Luoyang, Henan, 471003, P.R. China
Email: wwei@smail.hust.edu.cn

ABSTRACT

This work presents chaos synchronization between two different chaotic systems by nonlinear control laws. First, synchronization problem between a new chaotic system and Lorenz system has been investigated, and then the similar control is applied to the synchronization problem between two new chaotic systems proposed recently in the literature. Numerical simulations are shown to verify the results.

Keywords: Dynamical Systems, Chaos, Synchronization, Nonlinear Control, Runge-Kutta Method.

1. INTRODUCTION

Since Pecora and Carroll introduced a method [1] to synchronize two identical chaotic systems with different initial conditions, chaos synchronization, as a very important topic in the nonlinear science, has been developed extensively in the last few years. The idea of synchronization is to use the output of the master system to control the slave system so that the output of the response system follows the output of the master system asymptotically. A wide variety of approaches have been proposed for the synchronization of various chaotic systems which include PC method [1], OGY method [2,3], active control approach [4], adaptive control method [5,6,7,8], time-delay feedback approach [9], and backstopping design technique [10], etc. However, most of the methods mentioned above synchronize two identical chaotic systems. In fact, in many practical world such as laser array and biological systems, it is hardly the case that every component can be assumed to be identical. As a result, more and more applications of chaos synchronization in secure communications make it much more important to synchronize two different chaotic systems in recent years [11]. In this regard, some works on synchronization of two different chaotic systems have been performed [11, 12, 13].

This article considers the synchronization problems of two different chaotic systems. The new system proposed in [14] is taken as drive system. The Lorenz system and a new chaotic system devised by Chen and Lee [15] are selected as response system. In this work, we apply control theory to synchronize two different chaotic systems.

2. SYNCHRONIZATION BETWEEN A NEW CHAOTIC SYSTEM AND LORENZ SYSTEMS

In order to observe the chaos synchronization behavior in a new chaotic system and Lorenz systems, it is assumed that a new chaotic system drives the Lorenz system. Thus, the drive and response systems are as follows:

$$\begin{cases} \dot{x}_1 = ax_1 - y_1z_1 \\ \dot{y}_1 = -by_1 + x_1z_1 \\ \dot{z}_1 = -cz_1 + x_1y_1 \end{cases} \quad (1)$$

and

$$\begin{cases} \dot{x}_2 = \alpha(y_2 - x_2) + u_1 \\ \dot{y}_2 = \gamma x_2 - x_2z_2 - y_2 + u_2 \\ \dot{z}_2 = x_2y_2 - \beta z_2 + u_3 \end{cases} \quad (2)$$

where a, b, c, α, β and γ are positive scalars, and u_1, u_2 and u_3 are the control inputs to be designed. The aim of this section is to determine the control laws u_i for the global synchronization of two different chaotic systems. The new system has a chaotic attractor as shown in Fig. 1. when $a = 4.5$, $b = 12$, $c = 5$, and the Lorenz system has a chaotic attractor as shown in Fig. 2 when $\alpha = 10$, $\beta = 8/3$, $\gamma = 28$. Let us define the state errors between the response Eq. (2) that is to be controlled and the controlling Eq. (1) as

$$e_1 = x_2 - x_1, e_2 = y_2 - y_1, \quad e_3 = z_2 - z_1 \quad (3)$$

Subtracting (1) from (2) and using the notation (3) yields

$$\begin{cases} \dot{e}_1 = -\alpha x_2 + \alpha y_2 + y_1z_1 - ax_1 + u_1 \\ \dot{e}_2 = -e_2 - x_2z_2 - x_1z_1 + (b-1)y_2 + \gamma x_2 + u_2 \\ \dot{e}_3 = -\beta e_3 + x_2e_2 + e_1y_1 + (c-\beta)z_1 + u_3 \end{cases} \quad (4)$$

Now we define the control functions u_1, u_2 and u_3 as follows:

$$\begin{cases} u_1 = \alpha x_2 - \alpha y_2 - y_1z_1 + ax_1 - e_1 \\ u_2 = x_2z_2 + x_1z_1 - (b-1)y_1 - \gamma x_2 \\ u_3 = -x_2e_2 - e_1y_1 - (c-\beta)z_1 \end{cases} \quad (5)$$

Hence the error Eq. (4) becomes

$$\begin{cases} \dot{e}_1 = -e_1 \\ \dot{e}_2 = -e_2 \\ \dot{e}_3 = -\beta e_3 \end{cases} \quad (6)$$

The error Eq. (6) is a linear system of the form $\dot{w} = Aw$. Thus by linear control theory, if the system matrix A is Hurwitz, the system is asymptotically stable. Hence the error Eq. (6) with

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -\beta \end{pmatrix}$$

has all eigenvalues with negative real parts. This guarantees the asymptotic stability of the Eq. (6), which implies that the Lorenz Eq. (2) synchronize the new chaotic Eq. (1).

3. SIMULATION RESULTS

In this section, fourth order Runge-Kutta integration method with time step 0.001 is used to solve two systems of differential equations (1) and (2).We select the parameters of Lorenz system as $\alpha = 10$, $\beta = 8/3$, $\gamma = 28$ and the parameters of the new chaotic system as $a = 4.5$, $b = 12$, $c = 5$, so each of those systems exhibits a chaotic behavior. The initial values of the drive and response systems are $x_1(0) = 2$, $y_1(0) = 3$, $z_1(0) = 8$, and $x_2(0) = -2$, $y_2(0) = -3$, $z_2(0) = 3$, respectively, while the initial states of the error Eq. (6) are $e_1(0) = -4$, $e_2(0) = -6$, $e_3(0) = -5$. The diagram of Lorenz system controlled to be the new system is shown in Fig. 3 (a-c). The dynamics of synchronization errors for the drive system and response system is shown in Fig. 4..

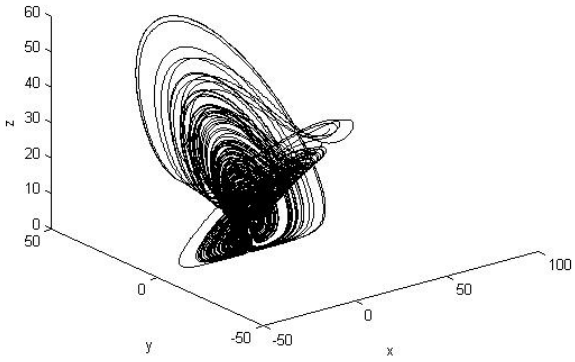


Fig. 1. The new chaotic attractor of Eq. (1).

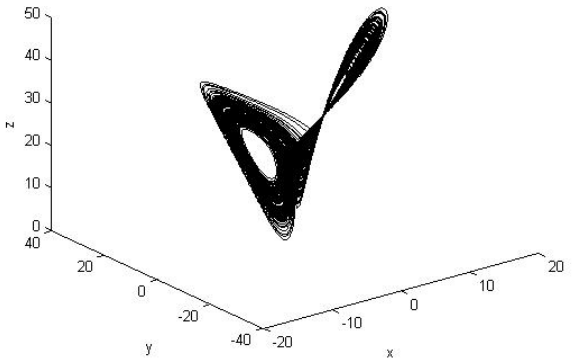
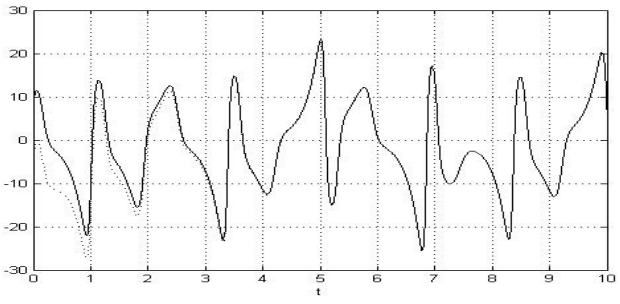
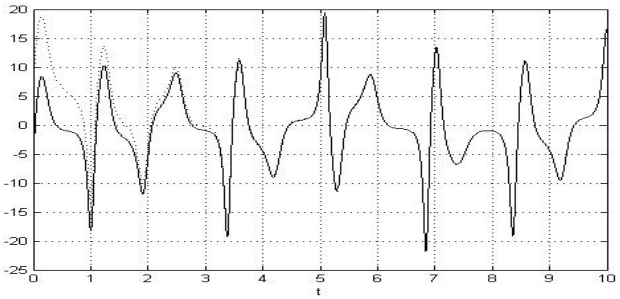


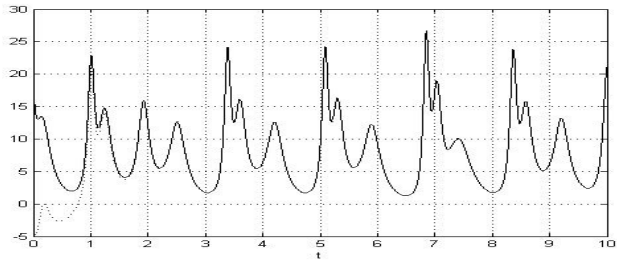
Fig. 2. Lorenz chaotic attractor.



(a)



(b)



(c)

Fig. 3. The diagram of the Lorenz system controlled to be the new system by using active control, (a) shows the time series of signals x_1 and x_2, (b) shows the signal y_1 and y_2, (c) shows the signal z_1 and z_2.

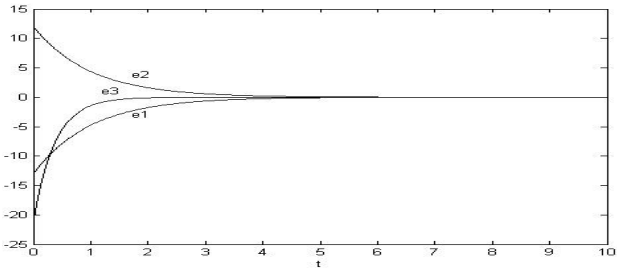


Fig. 4. Dynamics of synchronization errors states (e_1, e_2, e_3) for Eq.(6) with time t .

4. SYNCHRONIZATION BETWEEN TWO NEW CHAOTIC SYSTEMS

Recently, on studying anti-control of chaos, Chen and Lee [15] introduced a new chaotic system, which is described by the following nonlinear differential equation,

$$\begin{cases} \dot{x} = \alpha x - yz \\ \dot{y} = -\beta y + xz \\ \dot{z} = -\gamma z + (1/3)xy \end{cases} \quad (7)$$

where α, β and γ are positive control parameters. This system exhibits a chaotic attractor at the parameter values $\alpha = 5, \beta = 10$ and $\gamma = 3.8$.

In order to observe the chaos synchronization behavior in the two new Eq. (1, 7), it is assumed that Eq. (1) drives Eq. (7). Thus, the response system is as follows:

$$\begin{cases} \dot{x}_2 = \alpha x_2 - y_2 z_2 + u_1 \\ \dot{y}_2 = -\beta y_2 + x_2 z_2 + u_2 \\ \dot{z}_2 = -\gamma z_2 + (1/3)x_2 y_2 + u_3 \end{cases} \quad (8)$$

We subtract (1) from (8). We define the error system as the differences between the new Eq. (1) and the controlled

new Eq. (8). Let us define the state errors between the response Eq. (8) that is to be controlled and the controlling Eq. (1) as $e_1 = x_2 - x_1$, $e_2 = y_2 - y_1$, $e_3 = z_2 - z_1$.

Subtracting (1) from (8) and using the notation (3) yields

$$\begin{cases} \dot{e}_1 = \alpha e_1 - (a-\alpha)x_1 - y_2 z_2 + y_1 z_1 + u_1 \\ \dot{e}_2 = -\beta e_2 + x_2 z_2 - x_1 z_1 + (b-\beta) y_1 + u_2 \\ \dot{e}_3 = -\gamma e_3 + (1/3) x_2 y_2 - x_1 y_1 + (c-\gamma) z_1 + u_3 \end{cases} \quad (9)$$

Now we define the control functions u_1, u_2 and u_3 as follows:

$$\begin{cases} u_1 = (a-\alpha)x_1 + y_2 z_2 - y_1 z_1 - (\alpha+1)e_1 \\ u_2 = x_1 z_1 - x_2 z_2 - (b-\beta) y_1 \\ u_3 = x_1 y_1 - (1/3) x_2 y_2 - (c-\gamma) z_1 \end{cases} \quad (10)$$

Hence the error Eq. (3) becomes

$$\begin{cases} \dot{e}_1 = -e_1 \\ \dot{e}_2 = -\beta e_2 \\ \dot{e}_3 = -\gamma e_3 \end{cases} \quad (11)$$

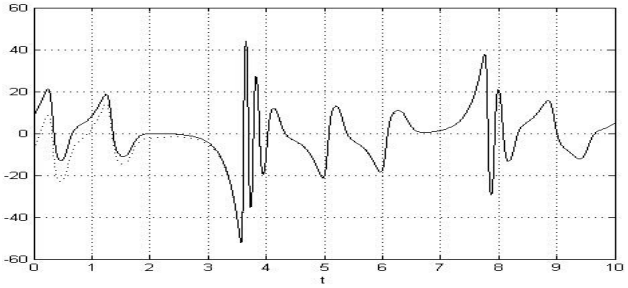
The error Eq. (11) is a linear system of the form $\dot{w} = Aw$. Thus by linear control theory, if the system matrix A is Hurwitz, the system is asymptotically stable. Hence the error Eq. (11) with

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\beta & 0 \\ 0 & 0 & -\gamma \end{pmatrix}$$

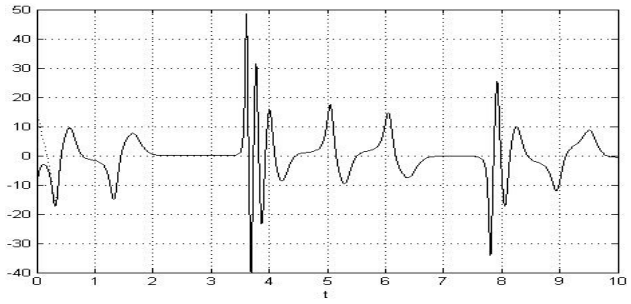
has all eigen values with negative real parts. This guarantees the asymptotic stability of the Eq. (6), which implies that Eq. (8) synchronize Eq. (1).

5. SIMULATION RESULTS

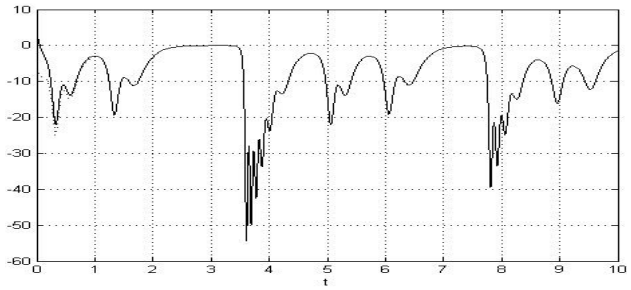
In this section, fourth order Runge-Kutta integration method with time step 0.001 is used to solve two systems of differential equations (1) and (8). We select the parameters of system as $\alpha = 5$, $\beta = 10$ and $\gamma = 3.8$, and the parameters of the new chaotic system as $a = 4.5$, $b = 12$, $c = 5$, so that each of the two new systems exhibits a chaotic behavior. The initial values of the drive and response systems are $x_1(0) = 3$, $y_1(0) = 0$, $z_1(0) = 4$, and $x_2(0) = -2$, $y_2(0) = -3$, $z_2(0) = -3$, respectively, while the initial states of the error Eq. (9) are $e_1(0) = -5$, $e_2(0) = -3$, $e_3(0) = -7$. The diagram of the new Eq. (7) controlled to be Eq. (1) is shown in Fig. 5 (a-c).



(a)



(b)



(c)

Fig. 5. The diagram of the Eq. (7) controlled to be Eq. (1) by using active control, (a) shows the time series of signals x_1 and x_2, (b) shows the signal y_1 and y_1, (c) shows the signal z_1 and z_2.

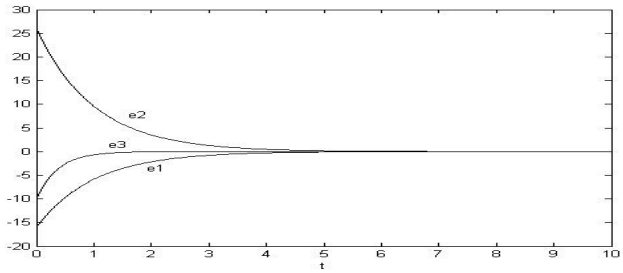


Fig. 6. Dynamics of synchronization errors states (e_1, e_2, e_3) for Eq. (7) with time t.

6. CONCLUSIONS

This article demonstrates that chaos in two different systems can be controlled using nonlinear control techniques. The stability of the resulting closed-loop error signals is easily proved by linear control theory. The Lorenz system is controlled to be a new system. Also, the new system is controlled to be the other new chaotic system. We can use active control theory to synchronize two identical or different chaotic systems. Two numerical simulations have been done to show the effectiveness of control scheme proposed.

7. REFERENCES

[1] L.M. Pecora, T.L. Carroll, "Synchronization in Chaotic Systems", Physical Review Letters, February 1990, 64(8): 821-824.

[2] E. Ott, C. Grebogi et al, "Controlling Chaos", Physical Review Letters, March 1990, 64(11): 1196-1199.

[3] Y.C. Hsiao, P.C. Tung, "Controlling Chaos for Nonautonomous Systems by Detecting Unstable Periodic

Orbits", Chaos, Solitons and Fractals, October 2002, 13(5): 1043-1051.
[4] H.N. Agiza, M.T. Yassen, "Synchronization of Rossler and Chen Dynamical Systems Using Active Control", Physics Letters A, January 2001, 278(4): 191-197.
[5] L.O. Chua, M. Itah et al, "Chaos Synchronization in Chua's Circuits", Journal of Circuits Systems and Computers, March 1993, 3(1): 93-108.
[6] M. Chen, Z. Han, "Controlling and Synchronizing Chaotic Genesio System Via Nonlinear Feedback Control", Chaos, Solitons and Fractals, August 2003, 17(4): 709-716.
[7] J. Lü, G. Chen et al, "Bridge the Gap between the Lorenz System and the Chen System", International Journal of Bifurcation and Chaos, December 2002, 12(12): 2917-2926.
[8] J. Lu, X. Wu et al, "Synchronization of a Unified Chaotic System and the Application in Secure Communication", Physics Letters A, December 2002, 305(6): 365-370.
[9] J.H. Park, O.M. Kwon, "A Novel Criterion for Delayed Feedback Control of Time-delay Chaotic Systems", Chaos, Solitons and Fractals, April 2005, 23(2): 495-501.
[10] X. Wu, J. Lu, "Parameter Identification and Backstepping Control of Uncertain Lü System", Chaos, Solitons and Fractals, August 2003, 18(4): 721-729.
[11] M.T. Yassen, "Chaos Synchronization between two Different Chaotic Systems Using Active Control", Chaos, Solitons and Fractals, February 2005, 23(1): 131-140.
[12] M.C. Ho, Y.C. Hung, "Synchronization of Two Different Systems by Using Generalized Active Control", Physics Letters A, September 2002, 301(5): 424-428.
[13] J.H. Park, "Chaos Synchronization between Two Different Chaotic Dynamical Systems", Chaos, Solitons and Fractals, January 2006, 27(2): 549-554.
[14] M.T. Yassen, "Controlling Chaos and Synchronization for New Chaotic Systems Using Linear Feedback Control", Chaos, Solitons and Fractals, November 2005, 23(6): 913-920.
[15] H.K. Chen, C.I. Lee, "Anti-control of Chaos in Rigid Body Motion", Chaos, Solitons and Fractals, September 2004, 21(4): 957-965.

Design and Simulation of Autonomous Fault Diagnosis and Reconfiguration Control System for Satellite Constellation Based on Distributed Control Architecture

JunHua Xiang
Institute of Aerospace and the Material Engineering, National University of Defense Technology
Changsha, Hunan 410073, China
Email: xiangjunhua@126.com

and

YuLin Zhang
Space Technology Research Center, Tsinghua University
Haidian, Beijing 100084, China
Email: y.l.zhang@mail.tsinghua.edu.cn

ABSTRACT

Based on the distributed control architecture， designed the system scheme of autonomous fault diagnosis and reconfiguration control system for satellite constellation. Studied and analyzed the constellation reconfiguration control schedule under existed the backup satellites or not. Finally carried on simulation and analysis of reconfiguration control for navigational satellite constellation applied even phase reconfiguration strategy, simulation results indicated the control scheme designed was feasible and effectively enhanced the autonomy of satellite constellation.

Keywords: Satellite Constellation, Autonomous Fault Diagnosis, Reconfiguration Control, Distributed Control Architecture.

1. INTRODUCTION

When some satellite is fault, no matter what kinds of station-keeping methods used by constellation station-keeping strategy, then it would affect the overall performance of satellite constellation. Specially, regarding the constellation using relative station-keeping method, control system collects the information of each spacecraft, and decides what satellite required carrying on controlling after centralism processing, and sends out the control command for this spacecraft, so when one or several satellites is fault, it enormously would influence the validity of entire station-keeping system, even appears the serious consequence. When appears fault satellite, it needs carry on control for fault satellite to leave the constellation autonomously, and when constellation exists the backup satellite, it must complete the control for the backup satellite independently to enter constellation and complete the reconfiguration of satellite constellation. If satellite constellation has not backup satellite, in order to satisfy the mission performance requirement, it need carry on overall analysis to satellite constellation and complete reconfiguration for satellite constellation [1, 2, 3].

This paper mainly aims at mission of constellation autonomous station-keeping, and references similar large-scale system for fault diagnosis and reconfiguration, carries on research and design of autonomous fault diagnosis and on-orbit reconfiguration control system for satellite constellation. Introduced some main descriptions of fault diagnosis system for satellite constellation, and carried analysis on three kinds of patterns of satellite constellation reconfiguration control: satellite increases, satellite number invariable, and satellite expiration, and proposed some reconfiguration strategies. Finally conducted the simulation research to navigational constellation, simulation results indicated that design plan of satellite constellation fault diagnosis and reconfiguration was feasble and reconfiguration schedule plan satisfied the requirements of autonomous reconfigure.

2. DISTRIBUTED CONTROL ARCHITECTURE

Constellation autonomous fault diagnosis and reconfiguration control system required the entire constellation has ability of information recognition, at the same time each satellite also has ability of information recognition and to complete constellation reconfiguration according to fault diagnosis results of satellite itself or received from constellation fault diagnosis system realized the performance improvement for satellite constellation. Autonomous fault diagnosis and reconfiguration control for satellite constellation based on distribute control architecture designed shows as figure 1.

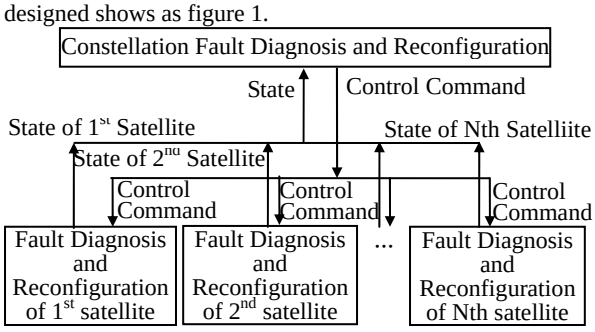


Fig. 1 Distributed control structure.

In figure 1, constellation level fault diagnosis and reconfiguration control subsystem does not have the on-orbit control executor, which mainly carries on on-orbit fault diagnosis and reconfiguration plan's production for satellite constellation according to satellite state information and constellation configuration information, then decomposes the reconfiguration control duties and transmits them to satellite level fault diagnosis and reconfiguration control subsystem. Satellite level fault diagnosis and reconfiguration control subsystem completes fault diagnosis of satellite itself, and sends its fault diagnosis information to constellation level fault diagnosis and reconfiguration control subsystem in order to obtain the control schedule of reconfiguration

control of overall constellation. But satellite level fault diagnosis and reconfiguration control subsystem is relates with the on-orbit control executor which finishes the reconfiguration control according the reconfiguration orders obtained from constellation level fault diagnosis and reconfiguration control subsystem or produced by satellite level fault diagnosis and reconfiguration control subsystem, so the overall system can complete autonomous fault diagnosis and reconfiguration control for satellite constellation.

Satellite level fault diagnosis and reconfiguration subsystem shows as figure 2, which includes residual error generator composed for dynamics model and autonomous navigation position system, diagnosis analysis and reconfiguration strategy generator[4, 5]. The residual error generator carries on operation to obtain residual error based on actual output (satellite orbit parameter and so on information obtained form autonomous navigation system) and the model's anticipated output. When analyzing the faults of actual movement, the residual error has the vital role, generally supposed that model is not failure, and the modeling error is smaller compared with system noise and survey noise, therefore if actual system failure, the residual error between model output and system actual output can exceed limit. According to the residual error signal, the fault analyzer use satellite orbit dynamics model, and takes some measures to strengthen the non-fault information such as the fault information, the suppression disturbance, the model error and so on, which contains in the residual error sequence. Using statistical analysis about the residual error sequence, it can examine the fault occurrence and carry on the diagnosis. Reconfiguration strategy module can produce the corresponding decision-making according to the fault diagnosis results, and sends out the control command.

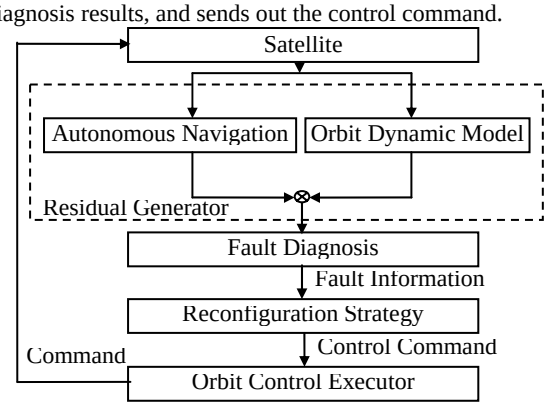


Fig. 2 Control architecture for one satellite.

Constellation level fault diagnosis and reconfiguration control subsystem like figure 3 shows, which includes constellation state identification module, constellation configuration fault diagnosis module, constellation reconfiguration strategy module, constellation configuration autonomous control module and so on. Constellation state identification module can diagnose the fault based on constellation state information obtained from each satellite, and constellation state information includes state, position, and orbital deviation of each satellite. Constellation reconfiguration strategy module can produce the corresponding decision-making according to the fault diagnosis results, and send out the reconfiguration control command. If fault only is the backup partially of satellite, it can take the switch measure, namely replaces the fault component using the backup component. If satellite itself is fault with no backup, but does not affect satellite performance, either the dropping is slow, then it can reduce the satellite weighting coefficient, either until supposes for zero. But if it is the fatal fault (for example satellite control system expires completely or thruster unable to work and so on), it need use the backup satellite or carry on the orbit adjustment in the orbit plane, and make overall constellation satisfy the performance index as far as possible. But the system motion equation of constellation and the control rule all should correspondingly change, and guarantee that the constellation configuration performance satisfies the requirement of stabilization and control. Constellation configuration autonomous control module mainly is transmit the control command to satellite controller based on the reconfiguration command send from constellation reconfiguration control strategy module, and finished the autonomous reconfiguration control of satellite constellation finally.

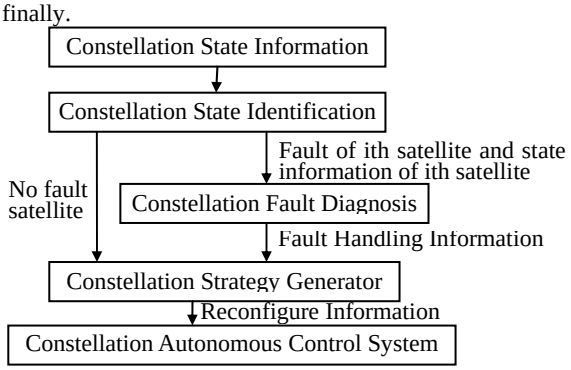


Fig. 3 Control architecture of satellite constellation.

The fault diagnosis algorithm is the core of fault diagnosis system. Whether the fault diagnosis system does succeed, mainly it is decided by the validity of fault diagnosis algorithm. To each performance of fault diagnosis algorithm, their order of priority degree is dissimilar to different tendency systematic and different monitoring goals, therefore it is need to aim at the different application domains and the concrete question and propose appropriately requirement. Currently fault diagnosis and constellation reconfiguration mainly is aim at the fault that can foresee and recover. Regarding complex fault diagnosis, it has the certain technical difficulty. If attitude, orbit, and load are abnormity, it can use the backup measure to solve to the fault known reason. It also can maintain this satellite in absolute station if the failure is unable to know the reason, and if the performance drops to unable to use, then it only can be abandoned. Regarding the constellation configuration reconfiguration control, it has control in the same orbit plane and shift between orbit planes. Because the cost of orbit plane control is expensive, it does not carry on the adjustment, and it needs to expire the working of this satellite from the constellation and wait for the backup satellite's launch. So constellation reconfiguration control mainly is phase control in the plane of satellite constellation.

3. RECONFIGURATION CONTROL PLAN

Orbit control of constellation reconfiguration mainly aims at some typical faults of the typical constellation such as satellite expiration or failure. Using constellation reconfiguration control plan design module to design the constellation configuration reconfigure plan, then passes the reconfiguration plan to constellation reconfiguration plan evaluation module to complete the evaluation to the improvement of the constellation operating performance after reconfiguration and the reconfiguration cost, then carries on the further revision in view of the performance

improvement and the cost caused by the reconfiguration plan, including proposed to use the backup satellite or launch the new satellite and so on, or carries on the adjustment to the satellite in constellation to ensure the most effective work of satellite constellation. The design process of optimized control for constellation orbit reconfiguration shows below as figure 4.

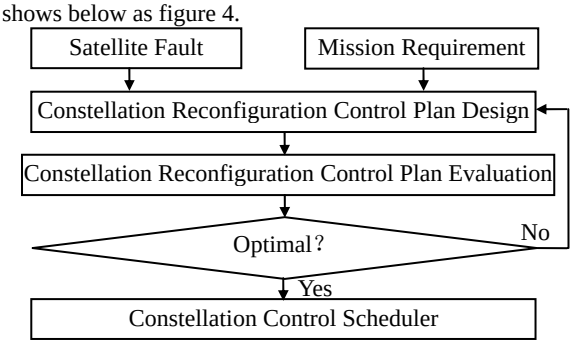


Fig. 4 Process of optimization design.

Because the orbit reconfiguration of satellite constellation need consume the propellant carried by satellite, it can reduce the working life of satellite. Therefore, the main optimized control goals for constellation orbit reconfiguration is that can complete the reconfiguration by the minimum price and does not affect the whole operating performance of satellite constellation, and at the same time require the control number to be least.

Regarding actual constellation mission, if constellation is not exposed by the initiative attack, then at the same time the situation which many satellites expires or failure is not least to see. Regarding the majority constellation, at least one backup satellite exists in each orbital plane. At the same time, along with the maturity of astronautics launch technology that the satellite supplement launch cycle also is more and more short. In the view of fuel consumption, reconfiguration control mainly is the phase adjustment, but regarding the phase control, it only need control semi-major axis deviation of satellite orbit, and consumption fuel for the control of semi-major axis deviation is very small.

Constellation reconfiguration is to evolve constellation A to constellation B, if satellite number in constellation A is N (A), satellite number in constellation B is N (B). According to the change situation of satellite number, constellation reconfiguration can be divided into three kinds of mainly working patterns[6]: satellite expired, the satellite increases, and the satellite number invariable.

(1)Satellite expired

Because of the satellite expires (function lose), the influence to constellation performance big because of unable promptly to increase the capacity. Then satellite number of constellation B is smaller than constellation A, namely N(B)<N(A). This kind situation does not need launch the new satellite, and existed one kind of type of reconfiguration, namely N(A) on-orbit satellites in constellation transfer to constellation B;

(2) Satellite increased

If satellite constellation is divided into two stages to deploy, it must consider stage deployment to increase the capacity of constellation, so satellite number of constellation B is higher than constellation A, namely N(B) >N(A). Constellation reconfiguration control will consider two kind of types transfer control: N(A) on-orbit satellites transfer to constellation B and N(B) - N(A) satellites launched to supplement constellation B.

(3) Satellite number invariable

When change of constellation configuration is caused by the change of constellation function, it need transfer the satellite in constellation A to satellite's corresponding position in constellation B, then N(B) = N(A).

Constructed three kinds of working patterns in view of satellite constellation, constellation on-orbit reconfiguration control mainly includes phase control in same orbital plane and ascending node longitude control for orbital plane, especially phase control.

4. RECONFIGURATION CONTROL SCHEDULE

The essential target of constellation reconfiguration control schedule is that obtains new nominal station of normal work satellite under maintaining the constellation operating performance. Because the fuel consumption is too big orbital plane transfer, it usually only has phase reconfiguration control in the same orbital plane.

The orbital control of constellation reconfiguration can use two parameters to attribute: transfer time Δt and energy required to provide. Usually defers to the change value of speed ΔV to compare the transfer energy required. Transfer velocity increment ΔV is independent in the flight vehicle quality. But Δt and ΔV all directly will affect the transfer cost. If the transfer time Δt will be too long, the satellite will not be able to provide the service in very long period of time, if transfer velocity increment ΔV will be too big, the satellite will need consume more fuels. Therefore, the question of constellation reconfiguration control is a optimization problems, the constraint condition and the control goal rely on Δt and ΔV .

In the supposition every orbital plane altogether has P satellite, and j^{th} satellite failure. Then it has two kinds of situations, one kind exist the backup satellite in the orbital plane, and other kind of situation does not exist the backup satellite. Now we carry on the satellite reconfiguration schedule in the orbital plane in view of these two kinds of situations separately.

Backup Satellite Existence

In order to look for the most superior control plan satisfied both the fuel consumption and the transfer time spends to complete reconfiguration control schedule of satellite in the orbital plane, to define satellite orbital reconfiguration control schedule's objective function is

$$J = \min \sum |\Delta v_i| + \lambda_i \min \sum |\Delta t_i| , \quad i = 0,1,\cdots,P-1$$

For λ_i is time consumption weight factor.

Backup satellite's parking orbit usually is higher or lower than normal orbit, and satellite reconfiguration control schedule mainly designs the transfer orbit which insures backup satellite to arrive the new goal orbital position in the shortest time using the smallest energy consumption that shows as figure 5.

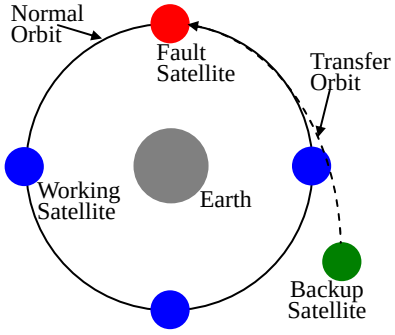


Fig. 5 Reconfiguration for parking orbit backup satellite.

If the backup satellite also works in the normal orbit according to the normal working pattern, then all normal satellites in orbital plane need reconfigure according to the strategy of phase uniform distribution. Rejects the expiration satellite, and carries on the phase control to all normal satellite that shows as figure 6.

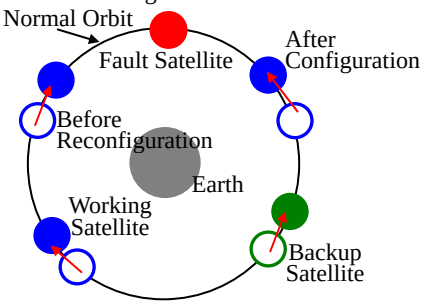


Fig. 6 Reconfiguration for working backup satellite.

Supposed that the number of all normal satellites in the orbital plane is P, the backup satellite is work in normal orbit as well as normal satellite, and the phase angular between neighboring satellites is

$$\Delta u_{j,j+1} = 360^{\circ} / (P+1)$$

After a normal satellite expiration or failure, the phase angular between neighboring satellites becomes

$$\Delta u'_{j,j+1} = 360^{\circ} / P$$

Then according to the constellation operating performance determined the nominal position of new satellite, and assigned the control duty of every satellite according to the results of constellation reconfiguration schedule.

No backup satellite

Because does not have the backup satellite, it must carry on the orbital plane adjustment of satellite and be enable the operating performance in the short time to achieve bigger. The essential target of reconfiguration control plan is that all satellites in the same orbital plane carry on the adjustment. The objective function becomes

$$J = \min \sum_{k} |\Delta V_{jk}| + \lambda_{jk} \min \sum_{k} |\Delta t_{jk}|$$

$$k = 0,1,\cdots,P-2$$

Then it need the nominal position in the same orbital plane with fault satellite according to the goal to maximum improve the whole operating performance of satellite, usually it only considers the orbital transfer in the orbital plane same as the expired satellite or fault satellite, and not consider the transfer between different orbital planes, which mostly can save propellant. But when several satellites in same orbital plane expires or failure, it has the serious influence to the whole operating performance of satellite constellation, for example appeared the big coverage gap, and can not launch satellite in the long-term time, it may consider the satellite transfer between different orbital planes, although it need consume massive propellants, in order to i the short-term improvement for constellation operating performance, its cost also can be accepted.

5. SYSTEM SIMULATION

Now using a navigational constellation to simulate and experiment for constellation autonomous fault diagnosis and reconfiguration control system based on distributed control architecture. Navigational constellation is the circular orbital Walker constellation, constellation configuration as 24/3/1, the orbital altitude as 21500km, and the orbital inclination as 55°. The navigational constellation may realize 6 assets global coverage, and the GDOP value is smaller than 1.9, which can see from figure 6 and 7. In the supposition that two satellites in the first orbital plane of navigational constellation are expired, namely Satellite 11, and Satellite 12, then the coverage performance and GDOP value are worse. The global coverage asset has decreased to 4, and the maximum value of GDOP has increased to 2.1.

Then using constellation autonomous fault diagnosis and reconfiguration control system to reconfigure the first orbital plane based on the strategy of evening phase. Obviously, after constellation configuration reconfiguration control, the global coverage performance of navigational constellation has repaired to 5, moreover the GDOP value also has the improvement such as figure 6 and figure 7.

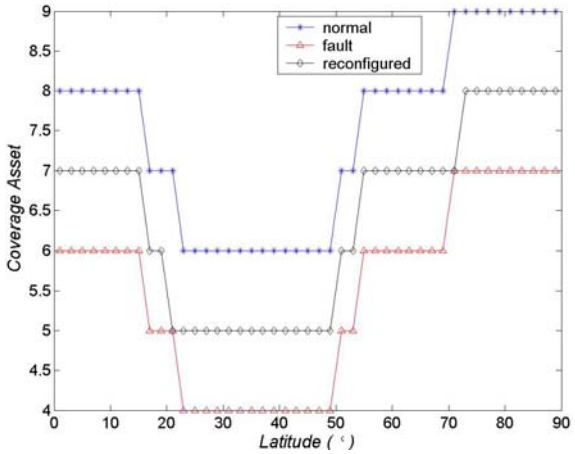


Fig. 6 Improvement for coverage asset.

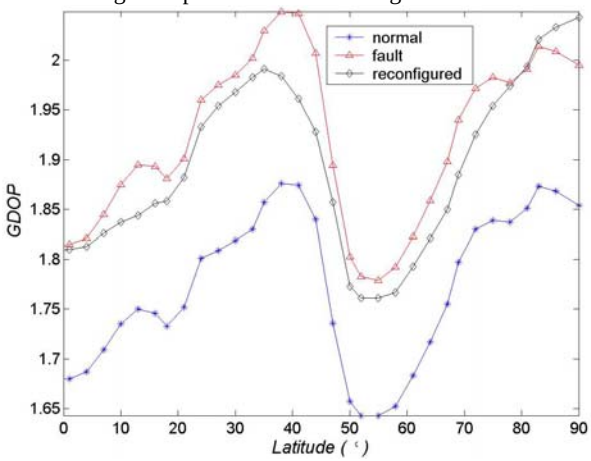


Fig. 7 Improvement for GDOP value.

6. CONCLUSION

Based on the fault-tolerant technical method, develop and apply autonomous fault diagnosis and reconfiguration control system, it can strengthen the autonomy of the satellite constellation and enable satellite constellation to own the fault-tolerant ability of on-orbit fault diagnosis and reorganization, reconfiguration control and so on, and effectively enhance the reliability of satellite constellation.

In order to realize on autonomous reconfiguration of satellite constellation, each satellite must own the ability independent fault diagnosis and component level reconfiguration. But, satellite receive the function of many kinds of disturbing forces in space environment, and computer resources is

limited, manual intervention ability is limited and so on, these factors all have restricted the realization of autonomous fault diagnosis[7]. At the same time, satellite constellation reconfiguration need synthesizes many factors as far as possible and optimize the reconfiguration control for satellite constellation.

7. ACKNOWLEDGMENT

Thanks Dr. Fan Li, Dr. Zhang Yasheng for their helps in constellation configuration design and reconfiguration strategy design.

8. REFERENCES

[1] Yu Xiaohong, Feng Shuxing, “Study on reconfiguring method of small satellite constellations for regional observation”, Journal of Astronautics, Vol.24, No.2, Mar. 2003, pp. 168-172.

[2] Zhang Yasheng, Zhang Yulin, “Research on the Fast Reconfiguration Method for the Performance Restore Type Constellation”, Journal of the Academy of Equipment Command & Technology, Vol.16, No.1, Aug. 2005, pp. 66-72.

[3] PARK, Chan Wang, “Coverage Analysis in Case of Faults of Some Satellites in Low Earth Orbit Satellite Constellations”, SSC98-XII-5, Proc. 12th Annual AIAA/USU Conference on Small Satellites, Logan, Utah, Aug.31-Sep.3, 1998.

[4] Wu Hongxin, Li Zhibin, “Space Station Intelligent Autonomous Control”, Journal of Aerospace Control, No.3, Jun. 1997, pp. 28-34.

[5] Dai Shuwu, Sun Huixian, “Technical overview of autonomous control and on-board data processing for spacecrafts”, Journal of Astronautics, Vol.24, No.1, Jan. 2003, pp. 17-22.

[6] Mathieu Chaize, Enhancing the Economics of Satellite Constellations via Staged Deployment and Orbital Reconfiguration, Master’s thesis, May 2003.

[7] Xing Yan, Wu Hongxin, Wang Xiaolei, LI Zhibi, “Survey of fault diagnosis and fault- tolerance control technology for spacecraft”, Journal of Astronautics, Vol.24, No.3, May 2003, pp. 221-226.

Simulation Model of Brushless Excitation System: Houshi Power Plant (China) *

Ahmed N. Abd-Alla, Ling Luo, J. Y. Wen , S. J. Cheng , Kui Li
College of Electric and Electronics Engineering, Huazhong University of Science and Technology, Wuhan , 430074, Hubei, China.
Email: najem20042003@yhaoo.com, jinyu.wen@163.com, sjcheng@hust.edu.cn, li_kui@126.com

ABSTRACT

This paper presents the performance of Excitation system for Houshi power plant (China). A standard simulation model of the proposed system is used by commercial software "Power System Stability" (PSS/E) and the performance under steady-state and dynamic conditions is studied. Two different identification methods, i.e., genetic algorithm and Prediction Error Method (PEM) are used and compared. In the genetic algorithm, first a block-diagram for the EXS is suggested, therefore in the investigated Excitation system measurement was performed at no-load operation condition to obtain step response to be able to identify the parameters of the model. The simulation results and their comparison, show the good accuracy of both methods. Although the model based on parameter obtained through genetic algorithm shows a better fit when its output response is compared with output response based on prediction Error method.

Keywords: Excitation system, parameter estimation, genetic Algorithm method, prediction error method (PEM).

1. INTRODUCTION

The Power system small signal, transient, and dynamic stability studies are only as accurate as the underlying models used in the computer analysis. The validity of the results of these studies depends heavily on the accuracy of the model parameters of the system components. In practice, the parameters commonly used in stability studies are manufacturer specified values, or "typical" values. These typical values may be grossly inaccurate, as various parameters may drift over time or with operating condition. Thus, to avoid such problems and to obtain more realistic simulation results, the identification of the system parameters based on field test is recommended. This paper reports the results of an on-going project to develop parameter estimation techniques for HouShi power plant, in Fijian province, China.

Several attempts have been made to obtain EXS models from field tests. A second order static excitation system has been discussed in [1]. In [2] generalized least square approach is used to model an excitation system. Parameter estimation of a pumped storage power plant using stochastic approaches is discussed in [3]. Identification of exciter constants using Prediction error Method (PEM) is addressed in [4]. In [5], the necessity to represent the EXS in full and close to the practical implementations for accurate and reliable results has been addressed. The feasibility and necessity of a nonlinear structure for EXS is discussed in [6].

* Project supported by National Natural Science Foundation of China (50595410).

This paper presents two identification methods to identify EXS model from some field tests. The methods are GA and PEM.

First a block diagram for the EXS parameter is proposed based on the documents and maps in the power station. To identify the parameters of this model, a test procedure, which is safe to perform, is presented. It should be emphasized that this method is not model-dependent and, therefore, it is readily applicable to a variety of model types and different test procedures.

When the EXS is identified using the genetic algorithms and Prediction error Method, the derived models are simulated. Their results are compared with each other and with the measured variables.

In this is paper is organized as follows: In section2, the EXS system is described and the test procedure is outlined. Section III, the genetic algorithm & Predication Error Method are introduced into parameter identification of Excitation system parameter. In section IV, the simulation results are compared with the measured actual variables to validate the derived model. Section V, concludes the paper.

2. SYSTEM DESCRIPTION AND THE TEST PROCEDURE

The unit under study is a 600[MW], 23[KV] Steam turbine generator set at the HouShi power plant in China. In HouShi a Brushless EXS (IEEE ESAC2A type exciter mode) is used.

2.1 System description

The proposed block diagram of the EXS is shown in Fig. 1. This block diagram has been proposed based on the documents and circuit diagrams in the power station. Details of such studies are not described here for the sake of brevity. This models a high initial response field controlled alternator-rectifier excitation system in which the alternator main exciter is used with noncontrolled rectifier. This model is applicable for simulating the performance of Westinghouse high initial response brushless excitation systems [7].

A direct negative feedback, V_H, around the exciter field time constant reduces its effective value and thereby increases the bandwidth of the excitation system small signal response. The time constant is reduced by the gain ($1 + K_B K_H$) of the compensation loop and is normally more than an order of magnitude lower than the time constant without compensation. To obtain high initial response with this system, a very high forcing voltage, V_{RMAX}, is applied to the exciter field.

A limiter sensing exciter field current allows high forcing, but limits the current. By limiting the exciter field current, exciter output voltage, V_E, is limited to a selected value, V_{LR},

which is usually determined by the specified excitation system response ratio. The output signals from the voltage regulator, V_A, and time constant compensation, V_H, elements are compared with the output signal, V_L, from the limiter in control logic circuitry, which functions to provide a sharp transition from regulator control to limiter control of excitation at the limit point. Excitation is controlled by the more negative of the two control signals.

Although the current limit is realized physically, the time constants associated with the loop can be extremely small. Therefore, the limit can be modeled as a positive limit on exciter voltage back of commutating reactance.

In this type of system, the $T_AT_CT_B$ represent Automatic Voltage Regulator's (AVR) time constants, K_A represents AVR gain .T_EK_E and S_E represents the exciter.

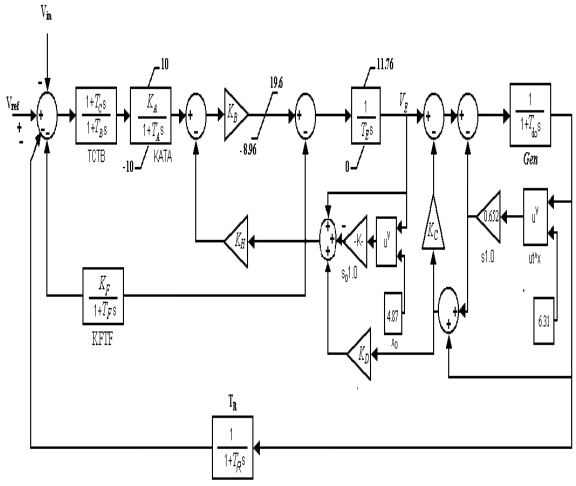


Fig. 1. Block diagram of Houshi power plant model

In this type of system, the TCTB & KFTF represents excitation system controller, KATA represents the exciter, and TR models the terminal voltage measurement device. L limits the output of excitation system.

2.2 Test procedure

The first step in the testing process is to prepare a test procedure. This requires a review of the information on the controls from the instruction manuals and block diagrams supplied by the manufacturer. A review of any plant specific concerns should also be made, for example, any operating restrictions imposed on the plant. This allows the test procedure to be adapted to the specific requirements exhibited by the plant.

In the defined test procedure of the EXS was treated as a single input single output. In this subsystems, *u(kT)* and *v(kT)* are the samples of the system input and output with constant sampling period T. The overall input signal *Vin* (input voltage for identification) was considered to be applied to the summing point of *Vref*.

The controller of excitation system is supplied by +10% step signal in sum input point of the excitation controller, this signals used to identify system parameter. Other Tests were performed by different step signals (5%,-5%, 2% and -2% step signal), this signals used to verify system parameter and Generator terminal Voltage.

3. EXCITATION SYSTEM PARAMETER IDENTIFICATION

In this section the theory of genetic algorithm and prediction error method used for EXS parameter identification are described.

3.1 Genetic Algorithm

The proposed identification procedure is a simulation based process that uses a genetic algorithm as optimization tool [8]. The simulation model of the system is excited by the same input. The output of the system, which is the set of available measurements, is compared to the simulated output of the model. The error between the two outputs is used as input to a genetic algorithm optimization module, which updates the model parameters in such a way that this error is minimized. The object function used to identify transfer function of excitation system can be calculated as:

$$Q=\sum\left(y-y_o\right)^2 \quad (1)$$

where, y is the output of identification result, is the output of actual process.

The transfer function of excitation system can be described as shown in figure 1:

The optimization process is to get the optimal parameters T_C, T_B, T_F, K_F, K_A, T_A, T_R, K_H ,K_E, K_D, T_{do} which can make Q minimum.

where, the searching area of the coefficients can be set according to experience, T_R,T_B,T_C=0~0.2, T_F and T_E =0.5~2.5, K_A=150~300, K_D and K_H=0.1~0.3, K_B=1~3and K_C and K_F =0.01~0.03, T_{do}=9~10.

A key feature of the approach is that the estimation process is not model-specific and it is therefore straight forward to switch between large varieties of models.

3.2 Prediction Error Method

In this method a time domain description of block is [4]:

$$y(t)=G(q).u(t)+H(q).e(t) \quad (2)$$

such that e(t) is white noise. Giving a description (2) and having observed input-output data u. y. the (prediction) errors e(t,) can be computed by:

$$e(t)=H^{-1}(q).[y(t)-G(q).u(t)] \quad (3)$$

For more information about PEM method one may refer to [9]. It should be mentioned that better estimations can be achieved by independent identification of nominator and denominator parameters of each block. For instance in ARX model we have:

$$A(q).y(t)=B(q).u(t)+w(t) \quad (4)$$

After employing identification algorithm, initial estimation for A(q) and B(q) is obtained, and we can rewrite equation(4) as:

$$y(t)=B(q).\frac{1}{A(q)}u(t)+\frac{1}{A(q)}w(t) \quad (5)$$

Defining

$$u_n=\frac{1}{A(q)}u(t) \quad , \quad w_n=\frac{1}{A(q)}w(t) \quad (6)$$

Then we have:

$$y(t)=B(q).u_n(t)+w_n(t) \quad (7)$$

From Eq.(6), and B(q) can be reestimated. This algorithm can be performed iteratively to improve the estimation.

4. THE IDENTIFICATION OF THE EXS FOR HUOSHI GAS POWER PLANT

In this section, we suppose to have the data obtained from the field test and estimation the parameters in Fig. 1 by using genetic algorithm and Prediction Error Method, from the obtained data is addressed.

4.1 Identification of the EXS Using genetic algorithm (GA)

In this section, the genetic algorithm identification method described in Section 3 is applied to the EXS at no-load operating conditions by 10% step signal. One signal is enough to got excitation system parameter [8].

Table 1.
The Excitation System Parameter Identification Result

T_R	0.006	K_H	0.29
T_B	0.0003	K_F	0.021
T_C	0.00001	T_F	1.12
K_A	280	K_C	0.05
T_A	0.07	K_D	0.12
K_B	2,51	K_E	1
T_E	2.01	T_{do}	9.2

4.2 Identification of the EXS Using the Prediction Error Method (PEM)

In this section, different sets of parameters were obtained by five step signal. Only the average model parameters of block diagram depicted in Fig. 1. are listed here for the sake of brevity.

Table 2.
The Average Model Parameters

T_R	0.01	K_H	0.27
T_B	0.0002	K_F	0.022
T_C	0.00001	T_F	1.09
K_A	250	K_C	0.045
T_A	0.02	K_D	0.105
K_B	2.48	K_E	1
T_E	2.004	T_{do}	9.2

5. MODEL VALIDATION

In any identification procedure, model validation is the most important step. The easiest way to validate a model is to compare the simulated model response to the measured output to the same input. This strategy was selected here for model validation [10, 11].

Fig. 2. to Fig. 6. compare the output result of the terminal voltage response of power plant model obtained through genetic algorithm (10%step signal) and Prediction Error method (five different step signals).

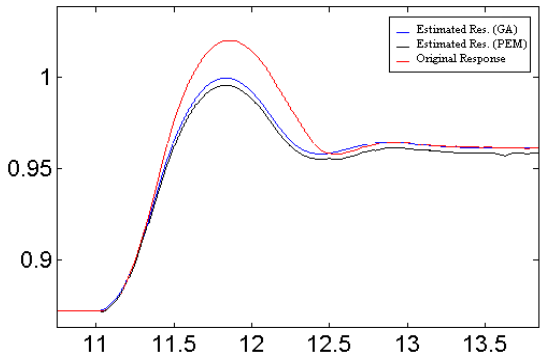


Fig. 2. Terminal voltage response Comparison under 10% step signal

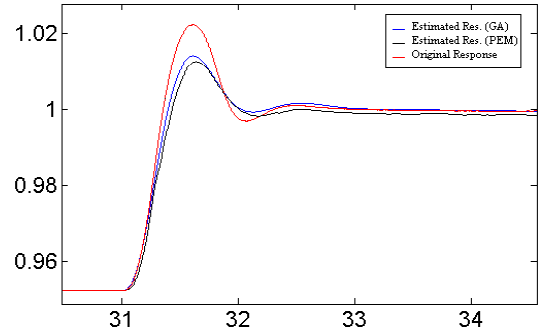


Fig. 3. Terminal voltage response Comparison under 5% step signal

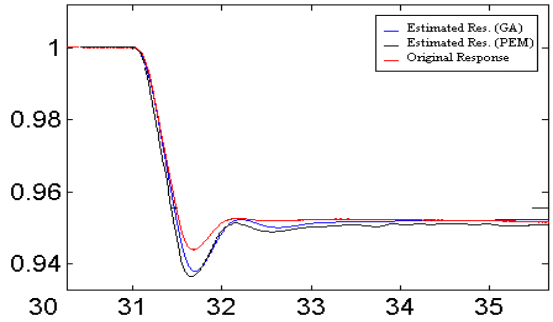


Fig. 4 .Terminal voltage response Comparison under -5% step signal

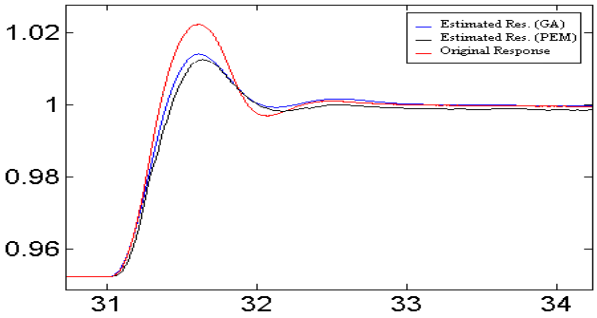


Fig. 5. Terminal voltage response Comparison under 2% step signal

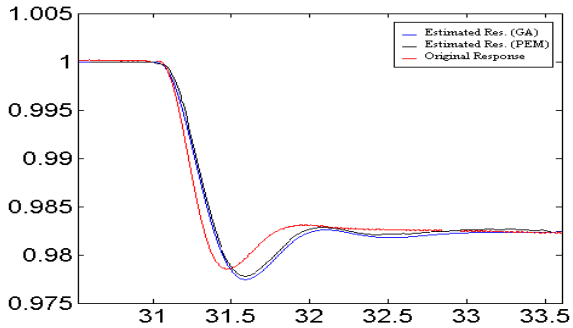


Fig. 6. Terminal voltage response Comparison under -2% step signal

Fig. 7. compares the output result of the terminal voltage response of power plant model obtained through genetic algorithm (5%step signal) and Prediction Error method (average model parameter).

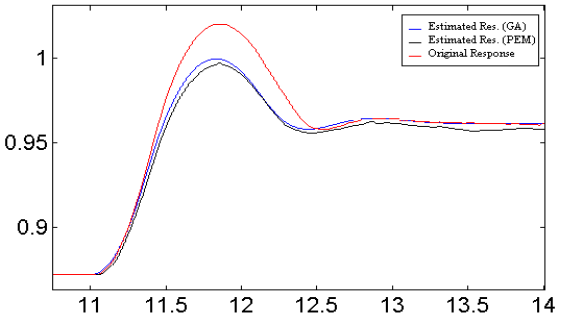


Fig. 7. Terminal voltage response Comparison under 10% step signal for average model parameter PEM and GA

Fig. 7. shows good accuracy of both methods. Although the power plant model parameter obtained through genetic algorithm shows a better fit when compared its terminal voltage response with response of model parameter obtained through prediction Error method. Since the superiority of this model parameter is not quite clear from the figure, Table 3 compares the error margin for the two models.

Table 3.
Comparison between the Two Methods Responding to the 10% Step Signal

MODEL'S	PREDICTION ERROR METHOD		GENETIC ALGORITHM		
Charact.	S.R.	R.M.	S.R.	R.M.	A. R.
t_d (s)	0.085	0.033	0.107	0.011	0.118
t_r (s)	0.457	0.033	0.449	0.025	0.424
t_p (s)	0.927	0.060	0.835	0.032	0.867
t_s (s)	1.933	0.210	2.080	0.063	2.143
M_p (%)	39.33	26.73	42.768	23.301	66.06
V.R (%)	0.669	—	0.470	—	0.132

Where: td=delay time, tr= rise time, tp=peak time, ts=stable time, Mp=max. peak, V.R.=Voltage Regulation, S.R= Simulation Result , R.M.= Error Margin , Actual Response = A. R., Charact.= Characteristic.

6. CONCLUSION

In this paper, two different identification methods were used for parameter identification of HuoShi Power Plant, Fijian province (China), genetic algorithm and Prediction Error Method. In the Genetic algorithm, model parameter of the power plant was obtained at no-load operating condition by 10% step signal. The advantages of the this method is that -
1. The few input data required, and the simplicity of its mechanism;
2. The identification process is not model-specific.
3. It is shown by simulation research that accurate identification results can be got no matter what kind of input signal is used;

In the Prediction Error Method, model parameter of the power plant was obtained at no-load operating condition by six different step signals. A single acceptable model was also obtained and validated by averaging the parameters.

The simulation results of both model parameters were close to the responses obtained from the field measurements. Although the model parameter obtained by genetic algorithm showed superior fitness to the measured variables.

7. REFERENCES

[1] G. Rogers, "Demystifying power system oscillations", IEEE Computer Applications in Power, vol. 9, July 1996, pp. 30~35.
[2] A. Zazo, J. L. Zamora et al., "Identification of excitation systems from time response tests", IEE Conference Publication, Vol.1, No.389, March 1994, pp. 839~843.
[3] T. Y. Guo, C. S. Liu et al., "Identification of excitation system model parameters via finalization field tests", International Conference on Advances in Power System Control, Operation and Management, Vo.12, December 1993, pp.833~838.
[4] M. Rasoli, M. karrari, "Identification of Brushless Excitation System Model Parameters at different operating conditions Via Field Tests", 14th PSCC, Seville, 24-28 June 2002.
[5] C. M. Liaw, T. S. Liu et al., "Parameter estimation of excitation systems from sampled data", IEEE Transactions on Automatic Control, vol. 37, May 1992, pp.663~666.
[6] R. Bhaskar, M. L. Crow et al., "Excitation system nonlinear parameter estimation for power system stability studies: Feasibility study", in Proceedings 30th North American Power Symposium., 1998, pp.453~458.
[7] T. L. Dillman, J. W. Skoog Lund et al., "High Initial Response Brushless Excitation System", IEEE Transactions on Power Apparatus and Systems, Vol. 90, September 1971, 2089-2094.
[8] C. L. Liu, J. Z. Liu et al., "The Application of Genetic Algorithm in Model Identification", in Proceedings IEEE TENCON'02, pp. 1261~1264.
[9] L. Ljung, "System Identification, Theory for the user", Prentice Hall, 1987.
[10] IEEE Guide for Identification, Testing, and Evaluation of the Dynamic Performance of Excitation Control Systems.
[11] IEEE Recommended Practice for Excitation System Models for Power System Stability Studies, August

1992.

Adaptive PID Control Based on Improved BP Neural Network

Ming-hui Qiang Ming-guang Zhang
School of Electrical and Information Engineering, Lanzhou University of Technology
Lanzhou,Gansu 730050, China
E-mail:zhangmingg8@163.com

ABSTRACT

An improved Back Propagation Neural Network (BPNN) based on self-leaning adaptive PID control strategy is presented in this paper. With the learning of a BPNN controller, it gradually compensates the deficiency of the feedback PID controller, until they gradually control the system together perfectly. The design of the controller is independent on the empirical knowledge of the system, and the parameters are tuned based on the testing information and error feedback-learning algorithm. The results of the simulation show that the proposed controller has the adaptability, strong robustness and satisfactory control performance in the nonlinear and time varying system.

Key words: Back Propagation (BP); Neural Network; Adaptive PID control; Simulation.

1. INTRODUCTION

The proportional-integral-derivative (PID) controllers are based on the most common control algorithm. Due to their simplicity and robustness, the PID controllers are widely used in the industrial process control [1]. But conventional PID controller with fixed parameters can hardly adapt to time varying of characteristics in wide range. To improve the control performance, several schemes of self-tuning PID controllers were proposed in the past. Junghui Chen, Tien-Chih Huang proposed using neural network to on-line updated PID controllers for nonlinear process control [2]. Ching-hung Lee, et, al designed a PID controller by fuzzy neural network (FNN)[3].

Back Propagation (BP) neural network is powerful computational tools that have been used extensively in the areas of pattern recognition, systems modeling and identification. The back propagation (BP) neural network is used to build the model of the relationship between process parameters and product properties. Now the application of BP model is most widely because of its simplicity and its power to extract useful information from samples. It allows specification of multiple input criterion, and generation of multiple output recommendations, and no assumption regarding the form of the functions relating input and output variables. BP model eliminates the limitations of the regression method, and accurately establishes the mapping between the input and output variables. BP neural network form a special architecture of neural network, which is characterized by several main advantages: the simplicity of its structure, faster learning algorithms, and better approximation capabilities. Due to the popularity of BP neural network, many researchers have been working during the last decade to develop more efficient training algorithms and applications [4,5,6,7,8].

According to advantages of Back Propagation (BP) Neural Network (NN) and PID controller, the adaptive PID control based on BPNN is proposed in this paper. We use self-learning ability of BPNN to automatically tune and modify the robust PID parameters on-line for the parametric interval system. Simulation results show that the PID controller obtained by using these improved BP neural network gives satisfactory results.

2. BACK PROPAGATION NEURAL NETWORK

2.1. Structure of BP Neural Network

BP neural network is basically a gradient decent algorithm designed to minimize the error function in the weights space. During training of the neural network, weights are adjusted to decrease the total error. In principle, it has been proved that any continuous function can be uniformly approximated by BP network model with only one hidden layer. So a three-layer BP model is employed in our study. It is easy to determine the number of neurons in the input layer and out-put layer during applications. A BP network is a three-layer feed-forward neural network [9,10,11,12]. A typical BP neural network structure is shown in Fig. 1..

In the structure of BP Neural Network, the first layer X is the input vector of the network, whose numbers are determined by complexity of controlled plant. Neurons at the second layer are hidden layer. The third layer is output layer including PID parameters k_p, ki, kd.

The digital incremental PID algorithm is usually adopted in computer control, the algorithm is as follows:

$$u(k) = u(k-1) + k_p(e(k) - e(k-1)) + k_i e(k) + k_d(e(k) - 2e(k-1) + e(k-2)) \quad (1)$$

Where input of the network is formulation (2).

$$O_j^{(1)} = x(j) \qquad j = 1,2,\cdots M \quad (2)$$

Input and output of hidden layer is formulation (3).

$$net_i^{(2)}(k) = \sum_{j=0}^{M} w_{ij}^{(2)} O_j^{(1)}$$
$$O_j^{(2)} = f(net_i^{(2)}(k)) \quad (3)$$

Where $w_{ij}^{(2)}$ is weight value of hidden layer, superscript (1), (2), (3) is input layer, hidden layer and output layer respectively.

The activation function of hidden layer adopts symmetrical sigmoid function.

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (4)$$

Inputs and outputs of output layer is as formulated.

$$net_l^{(3)}(k) = \sum_{j=0}^{Q} w_{li}^{(3)} O_i^{(2)}(k)$$
$$O_j^{(3)} = g(net_l^{(3)}(k)) \qquad l = 1,2,3$$
$$O_1^{(3)} = k_p$$
$$O_2^{(3)} = k_i \quad (5)$$
$$O_3^{(3)} = k_d$$

The activation function of output layer adopts non-negative sigmoid function because value of k_p, ki, kd can't be negative.

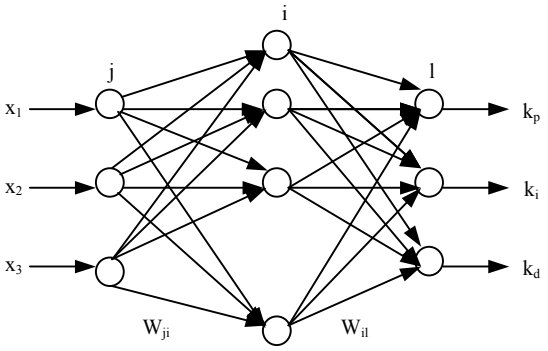


Fig. 1. BP neural network structure

$$g(x)=\frac{e^{x}}{e^{x}+e^{-x}} \tag{6}$$

2.2. Algorithm of BP neural network

The objective function to train the network is as follows:

$$E(k)=\frac{1}{2}(r(k)-y(k))^{2} \tag{7}$$

In order to minimize the error between the input *r(k)* and output *y(k)*, gradient descent method is adopted here to modify weights of the output layer. And inertial item is appended to speed up convergence and minimize global minima. The corresponding modifier formulas are as follows.

$$\Delta w_{li}^{(3)}(k)=-\eta\frac{\partial E(k)}{\partial w_{li}^{(3)}}+\alpha\Delta w_{li}^{(3)}(k-1) \tag{8}$$

where η is learning rate, α is inertial coefficient.

$$\frac{\partial E(k)}{\partial w_{li}^{(3)}}=\frac{\partial E(k)}{\partial y(k)}\cdot\frac{\partial y(k)}{\partial u(k)}\cdot\frac{\partial u(k)}{\partial O_{l}^{(3)}(k)}\cdot\frac{\partial O_{l}^{(3)}(k)}{\partial net_{l}^{(3)}(k)}\cdot\frac{\partial net_{l}^{(3)}(k)}{\partial w_{li}^{(3)}(k)} \tag{9}$$

$$\frac{\partial net_{l}^{(3)}(k)}{\partial w_{li}^{(3)}(k)}=O_{i}^{(3)}(k) \tag{10}$$

Because $\frac{\partial y(k)}{\partial u(k)}$ is unknown, it can approximately be substituted by signal function $\mathrm{sgn}(\frac{\partial y(k)}{\partial u(k)})$,the error brought by calculating can be compensated by tuning learning speed.

We obtain

$$\frac{\partial u(k)}{\partial O_{1}^{(3)}(k)}=e(k)-e(k-1) \tag{11}$$

$$\frac{\partial u(k)}{\partial O_{2}^{(3)}(k)}=e(k) \tag{12}$$

$$\frac{\partial u(k)}{\partial O_{3}^{(3)}(k)}=e(k)-2e(k-1)+e(k-2) \tag{13}$$

In the control and learning process, learning algorithm of weight of output layer is as formulated.

$$\Delta w_{li}^{(3)}(k)=\alpha\Delta w_{li}^{(3)}(k-1)+\eta\delta_{l}^{(3)}O_{i}^{(2)}(k) \tag{14}$$

$$\delta_{i}^{(3)}=e(k)\,\mathrm{sgn}(\frac{\partial y(k)}{\partial u(k)})\frac{\partial u(k)}{\partial O_{l}^{(3)}(k)}g'(net_{l}^{(3)}(k)) \quad l=1,2,3 \tag{15}$$

Learning algorithm of weight of hidden layer is as formulated.

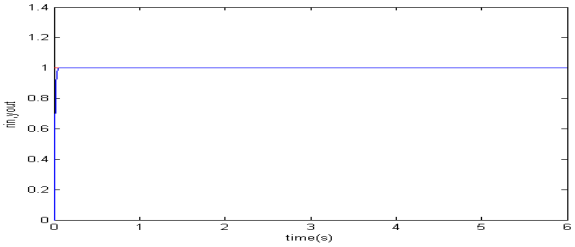


Fig. 3. The step response

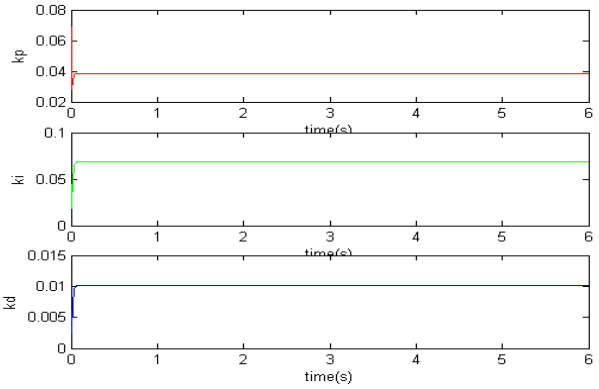


Fig. 4. Parameters adaptive tuning curve of the step

$$\Delta w_{ij}^{(2)}(k)=\alpha\Delta w_{ij}^{(2)}(k-1)+\eta\delta_{i}^{(2)}O_{j}^{(1)}(k) \tag{16}$$

$$\delta_{i}^{(2)}=f'(net_{i}^{(2)}(k))\sum\delta_{l}^{(3)}w_{li}^{(3)} \quad i=1,2,3 \tag{17}$$

Where $g'(\bullet)=g(x)(1-g(x)),\ f'(\bullet)=(1-f^{2}(x))/2$.

When the initial error is a bit larger, this kind of correcting can increase the inertial of the tuned weights, and restrain the influence of some big error in disturbing point, avoiding the oscillating of the system and strengthening the system robustness. In addition, it can overcome system's steady-state error and prevent the System's learning of the controller from getting into local minimum.

3. DESIGN OF ADAPTIVE PID CONTROLLER BASED ON RBFNN

The structure of adaptive PID control based on BP neural network is shown in Fig. 2. So the adaptive PID control algorithm and step is summarized as.

Step1 Choose the structure of the BPNN, namely, choose the node number of each layer and give the initial value of choose the learning speed rateη and inertial coefficient α;

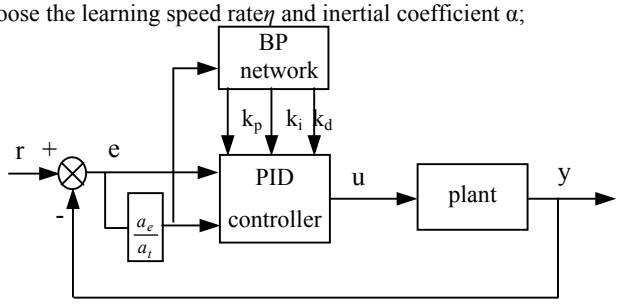


Fig. 2. Structure of adaptive PID control based on BPNN

Step2 sample *r(k)* and *y(k)*, calculate error e(k);

Step3 input *e(k),e(k-1),e(k-2)* into BPNN controller and get PID parameters k_p,k_i,k_d;

Step4 calculate output u(k) of PID controller;

Step5 according to BPNN learning, on-line adjust weight and achieve adaptive PID control;

Step6 set *k=k+1*,return to step 1.

4. SIMULATION EXPERIMENT AND RESULTS

The simulation adopts a nonlinear discrete system in this paper, whose transfer function is described by

$$y(k)=\frac{a(k)(k-1)}{1+y^2(k-1)}+u(k-1) \qquad (18)$$

Where *a(k)* is slowly time-vary, namely

$a(k)=1.2(1-0.8e^{-0.1k})$.

We adopt proposed algorithm and controller in the simulation. According to Fig. 1., BP network adopts a 4-5-3 structure. Learning rate η=0.28, inertial coefficient α=0.04 and weight∈ [-0.5,0.5]. The PID parameters *kp, ki, kd* is adjusted by self-learning of BP neural network until error approach zero.

Suppose input signal is step signal $r(t)=1.0$ and thus the output results of PID controller based on BPNN are shown in Fig. 3. and Fig. 4..When input signal is sine signal $r(t)=\sin(2\pi t)$, the simulation results are shown in Fig. 5. and Fig. 6. The simulation results show the outputs of the controller can match the output of the closed-loop controlled plant excellently. We can see that the system output tract the reference input satisfactorily and the performance of the proposed controller is better than that of the conventional PID controller. So the improved BPNN PID control has better dynamic characteristic and the self-learning ability.

5. CONCLUSIONS

Conventional PID controller tuned at typical operating point can hardly work well at different operating condition. A novel adaptive PID control strategy based on improved back propagation (BP) Neural Network (NN) is presented in this paper. The proposed controller has advantages of both self-learning capability of neural network and simplicity of PID controller. During the practice, adaptive controller has the superiority such as strong robustness, simple theory. Simulation results show that the proposed controller has the adaptability, strong robustness and satisfactory control performance in the nonlinear and time varying system.

6. ACKNOWLEDGEMENTS

The National 863 Scientific Project Development Foundation P.R. China under grant No.2002BA90128A supported this work.

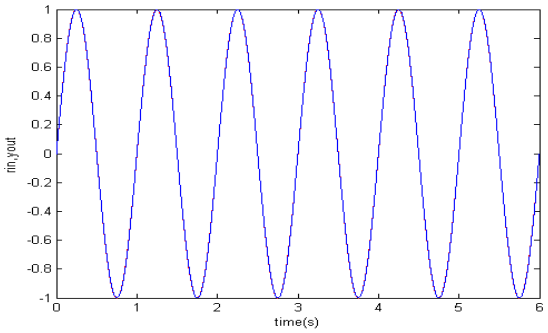


Fig. 5. The sine response curve

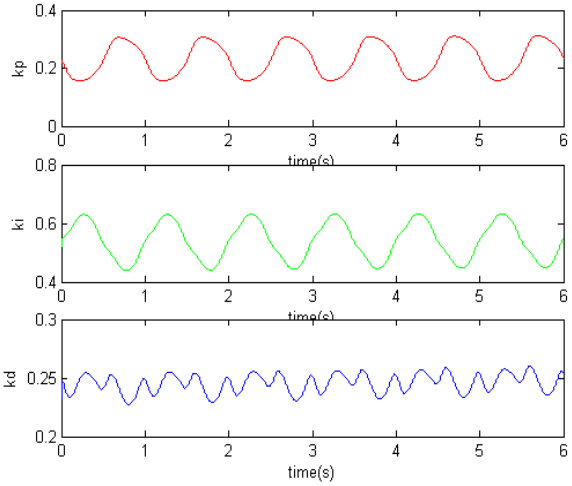


Fig. 6. Parameters adaptive tuning curve of the sine

7. REFERENCES

[1] K. J. Astrom, T. Hagglund, et al. “Automatic tuning and adaptation for PID controllers A survey”, IFAC J. Contr. Eng. Practice, 1993, 1(4): 699-714.

[2] J.H. Chen, T.C. Huang. “Applying neural network to on-line updated PID controllers for nonlinear process control”, Journal of Process Control, 2004, No.14: 211-230.

[3] C.H. Lee, Y.H. Lee,”A Novel robust PID controller design by fuzzy neural network”, Proceedings of the American Control conference, May 8-10, 2002, 1561-1566.

[4] L.G. Rudolph, T. R. Martinez, “A transformation strategy for implementing distributed, multilayer feed-forward neural networks: Back propagation transformation”, Future Generation Computer Systems, June 1997, 12(6): 547-564.

[5] L.S. Tan, N.X. Xiong et al. “A consolidation algorithm for multicast service using proportional control and neural network predictive techniques”, Computer Communications, 1 December 2005, 29(1): 114-122.

[6] K. Cai, J.T. Xia et al. “Analysis of the electrical properties of PZT by a BP artificial neural network”, Computational Materials Science, September 2005, 34(2): 166-172.

[7] J.T. Liu, H.B. Chang. “Prediction of the flow stress of high-speed steel during hot deformation using a BP artificial neural network”, Journal of Materials Processing Technology, 15 June 2000, 103(2): 200-205.

[8] D. Sha, V.B. Bajii’. On-line Adaptive Learning Rate BPAlgorithm For MLP and Application to An Identification Problem, available at http://www.cer.co.za/papers/ JACS_ 1999.pdf.

[9] B.L. Wang, L. Zhao Tan et al. “Neural network based online self-learning adaptive PID control”, Proceedings of the 3d World Congress on Intelligent Control and Automation, Vol.2, 28 June - 2 July 2000, 908-910.

[10] M.G. Zhang, X.G. Wang et al. “Adaptive PID Control Based on RBF Neural Network Identification”, Tools with Artificial Intelligence, 2005. ICTAI 05. 17th IEEE International Conference on, 14-16 Nov. 2005, 681-683.

[11] Y.J. Lee, J.H. Suh et al.“Adaptive PID control of an AGV system using humoral immune algorithm and neural network identifier technique”, Control Applications, 2004. Proceedings of the 2004 IEEE International Conference on, Vol.2, 2-4 Sept. 2004, 1576-1581.

[12] Y.L. Chen, W.L. Chen et al.“Development of the FES system with neural network + PID controller for the stroke”, Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on, 23-26 May 2005, 5119-5121.

Stabilization of Quantum Control in Spin-1/2 Systems

Bin Ye[1], WenBo Xu[1]
[1]School of Information Technology, Southern Yangtze University
WuXi, 214122, China
Email: yebinxie@yahoo.com.cn

ABSTRCT

We study effects of static imperfections on the stability of the quantum control of a two-level system. Our approach is based on linear response theory applied to quantum unitary evolution in presence of imperfections. Extensive simulation of steering a spin-1/2 particle via external electromagnetic fields demonstrates that the stability of quantum control increases notably with respect to appropriately tuned control parameters.

Keywords: Quantum Control, Spin 1/2 System, Fidelity, Linear Response Theorem

1. INTRODUCTION

Since the discovery that a quantum computer can in principle solve the problems intractable on classical computers [1], quantum computation and quantum information have become a major theoretical and experimental research topic. At present, however, one is able to perform laboratory computations only on a few qubits quantum computer. The main obstacles are errors in the evolution, either due to unwanted coupling with the environment or due to internal errors. Therefore, in order to precisely implement quantum gates, understanding the stability of quantum control with respect to perturbation is an important problem.

Quantum fidelity has been widely used as criteria for stability in the field of quantum dynamics [2, 3]. In this paper, we show that this approach can be efficiently used to study quantum control systems. Using linear response formalism in terms of correlation function of the perturbation, we analysis the fidelity decay of controlling two-level quantum systems. We propose an improved control function, which is more robust with respect to static perturbations. Numerical simulations with various control inputs are presented to confirm our predictions.

2. QUANTUM CONTROL OF A SPIN-1/2 PARTICLE

The problem of controlling quantum systems has long been an important scientific and technological challenge [4, 5]. The important applications of steering atomic and molecular quantum states includes control of chemical reactions, control of rotational coherence in linear molecules, quantum computation and laser cooling of internal molecular degrees of freedom [6]. In the last two decades, a number of theories were proposed to investigate the issue of controlling systems in the quantum regime [6-10].

The spin-1/2 system is a two-state quantum system. The two basis states spanning the Hilbert space are denoted by $|0\rangle$ and $|1\rangle$. Any other state $|\psi\rangle$ in the space can be written as a superposition $\alpha|0\rangle + \beta|1\rangle$ for some α and β such that $|\alpha|^2+|\beta|^2=1$. The Hamiltonian of a spin 1/2 particle driven by an external electromagnetic field is given by

$$H(t) = H_0 + \sum_{k=x,y,z} H_k u_k(t) \tag{1}$$

Where the internal Hamiltonian H_0 and the external Hamiltonian H_k are Hermitian linear operators, and $u_k(t), k = x, y, z$ are the components of the external electromagnetic field acted as the control functions. Given a initial state $|\psi(0)\rangle$, the Schrödinger equation is formulated in a bilinear equation

$$i\hbar|\dot{\psi}(t)\rangle = (H_0 + \sum_{k=x,y,z} H_k u_k(t))\cdot|\psi(t)\rangle \tag{2}$$

After rescaling of the time and control variables, the time evolution operator U (t) can be written in the form:

$$\dot{U} = -i(\sigma_x u_x + \sigma_y u_y + \sigma_z u_z)\cdot U$$

Where σ_x, σ_y, σ_z are Pauli matrices defined by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

By modeling the quantum system on a dynamical Lie group, finding the optimal solution of control functions are translated into controlling the evolution operator U(t). The optimal control problem for Eq.(2) is defined as: for a specified final time T and a target final matrix U_f, find a set of control inputs u_x, u_y, u_z such that $U(T, u_x, u_y, u_z)=U_f$ and the energy function

$$J(u) = \int_0^T \sum_{k,j=x,y,z} q_{kj} u_k(t) u_j(t) dt \tag{3}$$

with q_{kj} being a positive-definite constant number is minimized. Using algebraic dynamical method, the exact analytical solution can be obtained. In the case of two components of the electromagnetic field, say u_x and u_y, being varied, the optimal solutions for the problem of flipping the spin have the form

$$u_x^e(t) = a\cos(\omega t + \phi)$$
$$u_y^e(t) = a\sin(\omega t + \phi) \tag{4}$$

where a, ω and ϕ are constant parameters [10]. These parameters need to be chosen to match the target state. As a result, the possible minimum cost of the energy function Eq.(3) has the form:

$$J(u) = (2n+1)^2 \frac{\pi^2}{4T} \qquad n = 0,1,2,\cdots \tag{5}$$

and the corresponding evolution operator is

$$U(t) = \begin{pmatrix} e^{-i(\frac{\omega t+\phi}{2})} & 0 \\ 0 & e^{i(\frac{\omega t+\phi}{2})} \end{pmatrix} \begin{pmatrix} \cos at & -i\sin at \\ -i\sin at & \cos at \end{pmatrix} \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix} \tag{6}$$

3. FIDELITY DECAY IN THE QUANTUM

CONTROL SYSTEM

Due to unitarity of quantum dynamics, quantum stability cannot be defined in the same way as classical stability, namely through the exponential sensitivity on the variation of initial conditions. However, one can study the quantum fidelity defined as the stability of quantum motion with respect to a small variation in the Hamiltonian, or more generally, a variation of the quantum unitary evolution operator. When this concept applies to individual trajectories or phase space distribution functions, it is equivalent to the sensitivity to initial conditions. Whereas for chaotic orbits varying the Hamiltonian has a similar effect as varying the initial condition: exponential divergence of two nearby chaotic Hamiltonians [2].

Quantum fidelity is defined as an overlap between the state $|\psi(t)\rangle$ obtained by the unperturbed evolution $U_0(t)$ and the state $|\psi_\delta(t)\rangle$ obtained by the perturbed evolution $U_\delta(t)$:

$$F(t) = |\langle \psi_\delta(t) | \psi(t) |^2 \tag{7}$$

For a perturbation of strength δ generated by Hermitian operators V(t), the fidelity can be written to second order in δ as [2]:

$$F(t) = 1 - \frac{\delta^2}{2} \sum_{t,t'}^{T} C(t,t') + o(\delta^3) \tag{8}$$

Where the two-point time correlation function of the perturbation C(t,t') is

$$\begin{aligned} C(t,t') &= \langle V(t) V(t') \rangle \\ &= \mathrm{tr}(V(t)V(t'))/N \end{aligned} \tag{9}$$

With N being the Hilbert space size. The brackets $\langle \cdot \rangle$ denote the expectation value in the initial state. We note here that the time t is a discrete integer variable, denoting some basic single-step propagator.

In a physical situation with perturbation $V(t)=V_{static}+V_{random}$, the fidelity decay due to a static component is expected to dominate over the noise component [3]. Therefore, we will only consider the fidelity decay in terms of static perturbations, which is modeled by hermitian matrices from a gaussian unitary ensemble (GUE). For small perturbation strength δ the correlator in Eq.(8) is

$$\langle C(t,t') \rangle = \tfrac{1}{N} \langle tr(V(t,t')V) \rangle = |\tfrac{1}{N} trU(t,t')|^2 \tag{10}$$

Since the diagonal correlator C (t,t') is a static quantity for a fixed V, the fidelity is mainly determined by the off-diagonal correlator. From Eq. (10), it can be seen that the evolution operator U (t) that is traceless reduces the correlation function effectively. This is one of the main results previously obtained in [3].

Recalling the quantum optimal control system above, the sum of the off-diagonal correlator corresponding to evolution operator U (t) in Eq. (6) is

$$\begin{aligned} \sum_{t \neq t'} C(t,t') \quad &= 2(1 + 2\cos at)\cos(\tfrac{t}{2})\cos(\tfrac{t+\phi}{2}) \\ &+ \cos^2(\tfrac{t}{2})\cos(at) \end{aligned} \tag{11}$$

Hence enhancing the stability of the two-level control system is equivalent to minimizing Eq.(11). By minimizing Eq. (11) for a given target time T, we will get cos(aT)=0 or $\cos(\frac{t+\phi}{2}) = 0$ Because the amplitude a is more easy to control in physical situation, it is natural to set $a = \pi/(2T)$.

Numerical simulations with various amplitude a are performed to test our predictions. The initial density matrix ρ(0) is chosen as I/N where I is the identity matrix to calculate the correlation function over the whole Hilbert space. The amplitudes of control functions with $a = \frac{\pi}{2}, \frac{\pi}{4}$ and $a = \frac{\pi}{8}$ are compared with $a = \frac{\pi}{T}$. The simulation results are shown in Fig. 1. It can be seen clearly that the increasing rate of the correlation sum with $a = \frac{\pi}{T}$ is smallest compared with others control amplitudes. Thus the control system is more robust in the case of $a = \frac{\pi}{T}$.

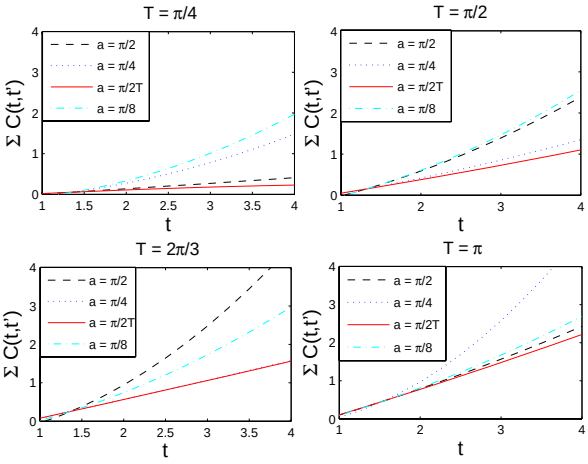


Fig. 1. The finite time correlation sum $\sum_{t \neq t'} C(t,t')$ for GUE perturbation in the two-level control system. The target control times are show on the top of each figure with $T = \frac{\pi}{4}, \frac{\pi}{2}, \frac{2\pi}{3}$ and π respectively.

4. CONCLUSIONS

Static perturbations have a dominating effect on the quantum dynamic systems compared with random noises. We have presented a novel approach based on linear response theory to study the stability of the two-level control systems with respect to static perturbations. We have shown that the control system is more stable after proper optimization of the control amplitudes. Simulation with GUE matrix demonstrates good stability with these control inputs.

5. REFERENCES

[1] R.P. Feynman, "Quantum Mechanical Computers", Foundations of Physics, Vol.16, 1986, 507-531.

[2] T. Prosen, M. Znidaric, "Stability of Quantum Motion and Correlation Decay", Journal of Physics A: Mathematical and General, Vol.35, 2002, 1455-1481.

[3] T. Prosen, M. Znidaric, "Can Quantum Chaos Enhance the Stability of Quantum Computation?", Journal of Physics A: Mathematical and General, Vol.34, 2001, 681-687.

[4] S. Habib, K. Jacobs, et al, "Quantum Feedback Control", Los Alamos Science Vol.27, 2002, 126-135.

[5] H. Mabuchi, N. Khaneja, "Principles and Applications of Control in Quantum Systems", International Journal of Robust and Nonlinear Control, Vol.15, 2005, 647-667.

[6] S.G. Schirmer, A.D. Greentree, et al, "Constructive Control of Quantum Systems Using Factorization of Unitary Operators", Journal of Physics A: Mathematical and General, Vol.35, 2002,8315-8339.

[7] A.C. Doherty, K. Jacobs, "Feedback Control of Quantum Systems Using Continuous State

Estimation", Physical Review A, Vol.60, 1999, 2700-2711.

[8] A.C. Doherty, S. Habib, et al, "Quantum Feedback Control and Classical Control Theory", Physical Review A, Vol.62, 2000, 12105-12107.

[9] G. Harel, V.M. Akulin, "Complete Control of Hamiltonian Quantum Systems: Engineering of Floquet Evolution", Physical Review Letters, Vol.82, 1999, 1-5.

[10] V. Ramakrishna, K.L. Flores, "Quantum Control by Decompositions of SU (2)", Physical Review A, Vol.62, 2000, 53409-53419.

An Adaptive Fuzzy Neural Network Nonlinear Control Algorithm for STT Anti-ship Missile

Changpeng Pan
Department of Command, Naval Aeronautical Engineering Institute
Yantai, Shandong, 264001, China
Email: navy@tom.com
And
Jie Chen, Wenjin Gu
Department of Control Engineering, Naval Aeronautical Engineering Institute
Yantai, Shandong, 264001, China
Email: hjhychenjie@tom.com

ABSTRACT

A new nonlinear control algorithm of combining dynamic inversion, adaptive fuzzy neural network and sliding mode control has been developed. The nonlinear dynamic theory is used to approximate linearization of the nonlinear system; the system dynamic inversion error is compensated by adaptive fuzzy neural network. The adaptive weights adjustment rule is derived by using lyapunov stability theory, and the system robustness is guaranteed by using sliding mode control and robustness control. The presented nonlinear control algorithm is applied to one STT missile control system, the simulation results show the proposed algorithm can eliminate the influence brought by disturbance availably, raise the precision of missile overload control system response.

Keywords: Nonlinear Control, Dynamic Inversion, Fuzzy Neural Network, Adaptive Adjustment Law, Sliding Mode Control.

1. INTRODUCTION

The research of anti-ship missile nonlinear control has received considerable attention during the last years, a great quantity of novel and creative nonlinear control algorithms were used in the missile control, and effects were satisfied [1,3,7,8]. Most nonlinear control techniques are based on linearizing the equations of motion by the application of nonlinear feedback. Known variously as feedback linearization or dynamic inversion, this method relies heavily on knowledge of the plant dynamics. More recently, adaptive fuzzy control has emerged as a means of explicitly accounting for uncertainties in the plant dynamics [2, 5, 6].

In this paper, we propose a adaptive fuzzy neural network sliding mode control (AFSMC) method, the proposed method combines a inverse dynamic feedback linearize control component and a sliding mode robustness control component, utilizes the adaptive fuzzy neural network control to compensate the error brought by inverse dynamic, builds up a control framework according to adaptive law, derives adaptive adjustment rules based on lyapunov function, and uses sliding mode control term and robustness control term to guarantee system stable. At last, an example of one STT supersonic anti-ship missile control system is given to illustrate the proposed method. Finally, a conclusion is provided that the proposed method can eliminate the influence brought by disturbance availably; raise the precision of missile overload control system response.

2. NONLINEAR CONTROL ALGORITHM

Consider a nonlinear system with the following structure:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})\boldsymbol{u} \tag{1}$$

Where $\boldsymbol{x} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots \boldsymbol{x}_n]^T$, $\boldsymbol{x}_i \in R^n, i = 1, 2, \cdots n$ is the state variables, $\boldsymbol{u} \in R$ denotes the control input, function $\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{g}(\boldsymbol{x})$ are sufficiently. Nonlinear dynamic inversion is used to proximate linearization of the nonlinear control system.

2.1. Dynamic inversion control law

If $\boldsymbol{f}(\boldsymbol{x})$ and $\boldsymbol{g}(\boldsymbol{x})$ are accurately known, One example, when it exists, is the dynamic inversion controller described by

$$\boldsymbol{u} = \boldsymbol{g}^{-}(\boldsymbol{x})[\boldsymbol{v} - \boldsymbol{f}(\boldsymbol{x})] \tag{2}$$

Substituting Eq. (2) into Eq. (1), yields the following equation:

$$\dot{\boldsymbol{x}} = \boldsymbol{v} \tag{3}$$

Regard $\boldsymbol{v} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots \boldsymbol{v}_n]$ as a new input, and then the original nonlinear system can be linearized into linear system by offsetting the nonlinear property.
In the tracking control, $\boldsymbol{v}$ is described by

$$\boldsymbol{v} = \dot{\boldsymbol{x}}_d - K_i \boldsymbol{z} \tag{4}$$

$\boldsymbol{x}_d = [\boldsymbol{x}_{d1}, \boldsymbol{x}_{d2}, \cdots \boldsymbol{x}_{dn}]$ Is the ideal final state, $\boldsymbol{z} = [\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots \boldsymbol{z}_n]$ is the system error vector, defined as $\boldsymbol{z}_i = \boldsymbol{x}_i - \boldsymbol{x}_{di}, i = 1, 2, \cdots n$, then design of the control law can be realized by the selecting of K_i.

But in some complexity nonlinear systems, such as missile control systems, the accurately $\boldsymbol{f}(\boldsymbol{x})$ and $\boldsymbol{g}(\boldsymbol{x})$ mathem-atical model could not be given, so the control of systems can't be realized only using dynamic inversion control law. Then a synthetically control fashion can be proposed, which adopts adaptive fuzzy neural network control algorithm to approach the unmodeled term of system, adopts additional sliding control component to guarantee adequately the system robustness, and adopts robust control component to compensate the system

uncertainty.

2.2. Synthetic control law

Assumption 1: the nominal model of system nonlinear terms $\boldsymbol{f}(\boldsymbol{x})$, $\boldsymbol{g}(\boldsymbol{x})$ are known.

Then the system becomes

$$\begin{aligned}\dot{\boldsymbol{x}} &= \bar{\boldsymbol{f}}(\boldsymbol{x}) + \Delta\tilde{\boldsymbol{f}}(\boldsymbol{x}) + (\bar{\boldsymbol{g}}(\boldsymbol{x}) + \Delta\tilde{\boldsymbol{g}}(\boldsymbol{x}))\boldsymbol{u} \\ &= \bar{\boldsymbol{f}}(\boldsymbol{x}) + \bar{\boldsymbol{g}}(\boldsymbol{x})\boldsymbol{u} + \Delta(\boldsymbol{x},\boldsymbol{u})\end{aligned} \tag{5}$$

Where $\bar{\boldsymbol{f}}(\boldsymbol{x})$ and $\bar{\boldsymbol{g}}(\boldsymbol{x})$ are the nominal model of $\boldsymbol{f}(\boldsymbol{x})$ and $\boldsymbol{g}(\boldsymbol{x})$. $\Delta(\boldsymbol{x},\boldsymbol{u}) = \Delta\tilde{\boldsymbol{f}}(\boldsymbol{x}) + \Delta\tilde{\boldsymbol{g}}(\boldsymbol{x})\boldsymbol{u}$ is the system unmodeled term.

Substituting Eq. (2) into Eq. (5), yields the following equation:

$$\dot{\boldsymbol{z}} = -k\boldsymbol{z} + \Delta(\boldsymbol{x},\boldsymbol{u}) \tag{6}$$

Assumption 2: The system is full state feedback, and $\bar{\boldsymbol{g}}(\boldsymbol{x})$ is nonsingler for all $\boldsymbol{x}$.

Because of existing $\Delta(\boldsymbol{x},\boldsymbol{u})$, the ideal control law of nonlinear system (5) can't be seeked directly, so compensation term is used to eliminate the uncertainty of system to assurance the suitable system dynamic response.

Supposing the synthetically control law as follows

$$\boldsymbol{u} = \boldsymbol{u}_c + \boldsymbol{u}_r \tag{7}$$

$$\boldsymbol{u}_r = -\delta\left\|\boldsymbol{u}_c\right\| sgn(\boldsymbol{s}) \tag{8}$$

$\boldsymbol{u}_r$ is robust control component, δ is exceeding zero constant, the define of $\boldsymbol{s}$ can be found in Eq. (11)

Supposing

$$\boldsymbol{u}_c = \bar{\boldsymbol{g}}(\boldsymbol{x})^{-}[-\bar{\boldsymbol{f}}(\boldsymbol{x}) + \boldsymbol{v}] \tag{9}$$

From Eq. (9) the establish of pseudo control term $\boldsymbol{v}$ is the key to the design of $\boldsymbol{u}_c$.

The following is assumed [2].

$$\boldsymbol{v} = \boldsymbol{v}_d + \boldsymbol{v}_s - \boldsymbol{v}_f \tag{10}$$

The select of $\boldsymbol{v}_d$ can be shown as Eq. (4), where $K_i = diag\{k_{v_i}\} \in R^{m\times m}, k_{v_i}$ is exceeding zero constant.

$\boldsymbol{v}_f$ is adaptive control term to compensate the uncertainty of system.

$\boldsymbol{v}_s$ is additional sliding model control term to guarantee adequately the system robustness.

To simplify symbols, $\tilde{\boldsymbol{f}}(\boldsymbol{x})$ denotes $\Delta\tilde{\boldsymbol{f}}(\boldsymbol{x})$, $\Delta\tilde{\boldsymbol{f}}$ denotes $\Delta\tilde{\boldsymbol{f}}(\boldsymbol{x})$. And the other variables are similar as these.

2.3. Sliding mode control component

Now establish $\boldsymbol{v}_s$ component. Define sliding mode surface

$$\boldsymbol{s} = c\boldsymbol{z} = c(\boldsymbol{x} - \boldsymbol{x}_d) \tag{11}$$

In Eq. (11) $\boldsymbol{x}_d$ is expected value,

$$c = diag[c_1, c_2, \cdots c_{n-1}, 1]$$

Choose the parameter c_i, $i = 1, \cdots, n-1$, get the polynomial

$$p(s) = s^{n-1} + c_{n-1}s^{n-2} + \cdots + c_2 s + c_1$$

The above polynomial is Hurwitz stabilized, where s is laplace operator[4].

Supposing

$$\boldsymbol{v}_s = -\mu\boldsymbol{s} - \rho sgn(\boldsymbol{s}) \tag{12}$$

Where $sgn(\boldsymbol{s})$ is symbol function.

Select variable $0 < \eta < \varepsilon$. $sgn(\,)$ function is replaced by $sat(\,)$ to eliminate chatter.

$$sat(w) = \begin{cases} 1 & x > \sigma \\ w/\sigma & |x| \le \sigma \\ -1 & x < -\sigma \end{cases} \qquad (\sigma > 0) \tag{13}$$

2.4. Adaptive fuzzy neural network control component

Adaptive fuzzy neural network control component is generally a fuzzy inference system constructed from the structure of adaptive. Learning algorithms are used to adjust the weights of the fuzzy inference system [6]. Its output $\boldsymbol{v}_f$ can be expressed as following [7]:

$$\boldsymbol{v}_f = \frac{\Sigma_{i=1}^{h}\bar{y}^i\left(\Pi_{j=1}^{n}\mu_{A_j^i}(x_j)\right)}{\Sigma_{i=1}^{h}\left(\Pi_{j=1}^{n}\mu_{A_j^i}(x_j)\right)} = \boldsymbol{\theta}^T\boldsymbol{\psi}(\boldsymbol{x}) \tag{14}$$

where $\mu_{A_j^i}(x_j)$ is the membership function of fuzzy variable x_j, h is the total number of the fuzzy rules, $\bar{y}^i$ is the point at which $\mu_{B^i}(\bar{y}^i) = 1$, A_j^i and B^i denote fuzzy sets of input and output variable, respectively, $\boldsymbol{\theta} = [\bar{y}^1\ \bar{y}^2\ \cdots\ \bar{y}^h]^T$ is a weight vector of the output layer, $\boldsymbol{x} \in \Re^n$ is input vector, and $\boldsymbol{\psi}(\boldsymbol{x}) = [\psi^1\ \psi^2\ \cdots\ \psi^h]^T$ is a fuzzy basis vector, where ψ^i is defined as

$$\psi^i(\boldsymbol{x}) = \frac{\Pi_{j=1}^{n}\mu_{A_j^i}(x_j)}{\Sigma_{i=1}^{h}\left(\Pi_{j=1}^{n}\mu_{A_j^i}(x_j)\right)} \tag{15}$$

In this section, the fuzzy logic is used to approach the uncertainty term $\Delta(\boldsymbol{x},\boldsymbol{u})$.

Supposing $\tilde{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*$ Where $\tilde{\boldsymbol{\theta}}$ is the estimated error of $\hat{\boldsymbol{\theta}}$, $\boldsymbol{\theta}^*$ is the optimal value of $\hat{\boldsymbol{\theta}}$.

Now we use the adaptive fuzzy neural network to approximate $\Delta(\boldsymbol{x},\boldsymbol{u})$.

$$\sup|\,\boldsymbol{\theta}^{*T}\boldsymbol{\psi}(\boldsymbol{x}) - \Delta(\boldsymbol{x},\boldsymbol{u})\,| \le \varepsilon \tag{16}$$

Where ε is given positive vector, $\tilde{\varepsilon}_e = \varepsilon - \hat{\varepsilon}$, $\hat{\varepsilon}_i$ is estimated value of ε_i.

Therefore $\Delta(\boldsymbol{x},\boldsymbol{u})$ can be represented as

$$\Delta(\boldsymbol{x},\boldsymbol{u}) \le \boldsymbol{\theta}^{*T}\boldsymbol{\psi}(\boldsymbol{x}) + \varepsilon \tag{17}$$

However, $\boldsymbol{\theta}^*$ can't be used directly, so we introduce $\hat{\boldsymbol{\theta}}$ as an estimate of $\boldsymbol{\theta}^*$.

Next we will choose the learning algorithm of the weight vector; we will use Lyapunov function to construct it.

Submitting Eq. (7) ~ (10), (12), (14) into Eq. (6) yields the following equation

$$\dot{\boldsymbol{z}} = -K_i^T\boldsymbol{z} - \mu\boldsymbol{s} - \rho sat(\boldsymbol{s}) - \hat{\boldsymbol{\theta}}\boldsymbol{\Psi} + \bar{\boldsymbol{g}}\boldsymbol{u}_r$$

$$\therefore\ \dot{\boldsymbol{s}} = -K_i^T\boldsymbol{s} - c(\mu\boldsymbol{s} - \rho sat(\boldsymbol{s}) - \hat{\boldsymbol{\theta}}\boldsymbol{\Psi} + \bar{\boldsymbol{g}}\boldsymbol{u}_r) \tag{18}$$

3. STABILITY ANALYSIS

Supposing Lyapunov function:

$$V = \frac{1}{2}\boldsymbol{s}^T\boldsymbol{s} + \frac{1}{2}\Gamma^{-1}\cdot tr\{\tilde{\theta}^{\mathrm{T}}\tilde{\theta}\} + \frac{1}{2}\tau\boldsymbol{\varepsilon}_e^{\ T}\boldsymbol{\varepsilon}_e \tag{19}$$

$\Gamma = diag(\Gamma_1,\cdots\Gamma_n)$ is learning law, taking the derivative of V yields

$$\dot{V} = \boldsymbol{s}\dot{\boldsymbol{s}} + \frac{\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\dot{\tilde{\boldsymbol{\theta}}}}{\Gamma} + \tau\boldsymbol{\varepsilon}_e\dot{\boldsymbol{\varepsilon}}_e$$

Submitting Eq. (8), (10) into Eq. (14), yields the following equation

$$\dot{V} = \boldsymbol{s}[-K_i^T\boldsymbol{s} - c(\mu\boldsymbol{s} + \rho sat(\boldsymbol{s}) + \bar{\boldsymbol{g}}\delta\|\boldsymbol{u}_c\|sgn(\boldsymbol{s}))]$$

$$-c\boldsymbol{z}^{\mathrm{T}}(\boldsymbol{\theta}^{*\mathrm{T}}\boldsymbol{\Psi} + \varepsilon_i) + \frac{tr\{\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\dot{\tilde{\boldsymbol{\theta}}}\}}{\Gamma} + \tau\boldsymbol{\varepsilon}_e\dot{\boldsymbol{\varepsilon}}_e$$

$\therefore$ The learning algorithm of the weight vector can be chosen as

$$\hat{\boldsymbol{\theta}} = \dot{\tilde{\boldsymbol{\theta}}} = -\Gamma\boldsymbol{s}^{\mathrm{T}}\boldsymbol{\Psi} \tag{20}$$

$$\dot{\tilde{\varepsilon}}_e = -\frac{1}{\tau}|s_i| \tag{21}$$

$s_i = c_i z_i, i = 1,2,\cdots n$

Then

$$\dot{V} \le -\boldsymbol{s}k_{vi}\boldsymbol{z} - c_i\boldsymbol{z}^{\mathrm{T}}(\boldsymbol{\theta}^{*\mathrm{T}}\boldsymbol{\Psi} + \varepsilon_i) - \Gamma\|\boldsymbol{s}\|\boldsymbol{\Psi} - \tau\varepsilon_e|s_i|$$

Selecting k_{vi}, c_i, Γ_i, τ are exceeding zero constants, then $\dot{V} \le 0$ established.

So we can draw such conclusion that system is global asymptotical stability at balance point $x = 0$ when selected control law is $\boldsymbol{u} = \boldsymbol{u}_c + \boldsymbol{u}_r$, $\boldsymbol{u}_c$ can be established by Eq. (9), (10), (12), (14).

4. STT ANTI-SHIP MISSILE OVERLOAD CONTROL SYSTEM

Anti-ship missile is a channel-coupled time-variability nonlinear multivariate system. In this paper a three channels separated design method is proposed by using above nonlinear control algorithm. Transform missile overload output tracking problem into system status tracking control. Maintain the coupling terms among channels. That are considered as disturbance, the original system control can be simplified as one order proximate system control.

4.1 Three channels separated design

For axis symmetrical STT missile, considering small attack angle and small sideslip angle, missile dynamics can be expressed as[1]:

$$\dot{\gamma}_C = \omega_x - tg\vartheta\left(\omega_y\cos\gamma_C - \omega_z\sin\gamma_C\right)$$

$$\dot{\alpha} = \omega_z + Z_\alpha\alpha + Z_{\delta_z}\delta_z - \beta\omega_x + \frac{g\cos\vartheta\cos\gamma_C}{v_m}$$

$$\dot{\beta} = \omega_y + \alpha\omega_x + Y_\beta\beta + Y_{\delta_y}\delta_y + \frac{g\cos\vartheta\sin\gamma_C}{v_m}$$

$$\dot{\omega}_y = N_\beta\beta + N_{\omega_y}\omega_y + N_{\delta_z}\delta_y - \left(\frac{J_x - J_z}{J_y}\right)\omega_x\omega_z$$

$$\dot{\omega}_x = M_x^{\ \beta}\beta + M_x^{\ \delta_x}\delta_x + M_x^{\ \omega_x}\omega_x$$

$$\dot{\omega}_z = M_\alpha\alpha + M_{\omega_z}\omega_z + M_{\delta_z}\delta_z - \left(\frac{J_y - J_x}{J_z}\right)\omega_x\omega_y$$

$$\dot{\delta}_{z,y,x} = -w_\alpha\delta_{z,y,x} + w_\alpha u_{z,y,x}$$

$$n_y = -\frac{v_m}{g}(\dot{\alpha} - \omega_z), n_{\mathrm{Z}} = \frac{v_m}{g}(\dot{\beta} - \omega_y) \tag{22}$$

Where α is attack angle, β is sideslip angle, and γ_c is roll angle, $(\omega_z,\ \omega_y, \omega_x)$ is the missile angular velocity vector, $(\delta_z,\ \delta_y, \delta_x)$ is the rudder roll angle rate, $(u_X,\ u_Y, u_Z)$ is the rudder command, $(n_z,\ n_{\mathrm{y}})$ is the corresponding overload. w_α is rudder time constant, g is gravity overload, v_m is missile velocity, the aim to of design STT autopilot is to yield overload in (Y, Z) direction, keeping small roll rate.

One supersonic speed anti-ship missile is a STT missile, its roll channel has still stability, and pneumatic contour is axis symmetry, so missile pitch channel and yaw channel are elementary conformably. Here take the example of pitch channel.

4.2 Pitch channel autopilot design

The aim to design STT missile pitch channel autopilot is that the missile longitudinal overload n_y can track guidance command $n_y^{\ c}$.

Non-minimum phase characteristic exists in the transition function from rudder angle to overload output of tail-controlled anti-ship missile [1].

In the view of autopilot design, non-minimum phase characteristic is unacceptable, that can bring to the instability zero dynamic. But non-minimum phase characteristic is only a property of relationship between input and output, not of whole system. Therefore it is

feasible to resolve the stability of non-minimum phase system by virtue of output redefinition. So for the purpose that system statement can denote missile performance directly, coordinate transformation is necessary.

Define new system status $\left[\alpha-\kappa n_y{}^c, \omega_z, n_y-n_y{}^c\right]^T$, under new status conduction, invert linear part in system (22) dynamic to canonical form, then the system zero dynamic zero dynamic is stable.

$$\begin{bmatrix}\dot{\eta}\\ \dot{\zeta}_1\end{bmatrix}=\begin{bmatrix}\boldsymbol{Q}_1 & \boldsymbol{P}_1\\ \boldsymbol{Q}_2 & P_2\end{bmatrix}\begin{bmatrix}\eta\\ \zeta_1\end{bmatrix}+Bu+\tilde{f}+\tilde{\xi} \qquad (23)$$

Where

$$\boldsymbol{\eta}=\left[\alpha-\alpha^c \quad \omega_z\right]^T, \zeta_1=n_y-n_y{}^c$$

$$\mathbf{Q}_1=\begin{bmatrix}0 & 1\\ Q_3 & M_{\omega_z}\end{bmatrix}, \boldsymbol{Q}_2=-v\left[Z_\alpha w_\alpha \quad Z_\alpha\right]/g$$

$$Q_3=\left(M_\alpha-M_{\delta_z}Z_\alpha/Z_{\delta_z}\right), \mathbf{P}_1=\begin{bmatrix}-g/v\\ -gM_{\delta_z}/vZ_{\delta_z}\end{bmatrix}$$

$$P_2=\left(Z_\alpha-w_\alpha\right), \boldsymbol{b}=\left[0 \quad 0 \quad b_m\right]^T, b_m=vZ_{\delta_z}w_\alpha/g$$

$$\tilde{f}=\left[\frac{g\cos\vartheta\cos\gamma}{v}-\beta\omega_x;-\left(\frac{J_y-J_x}{J_z}\right)\omega_x\omega_y;0\right]^T$$

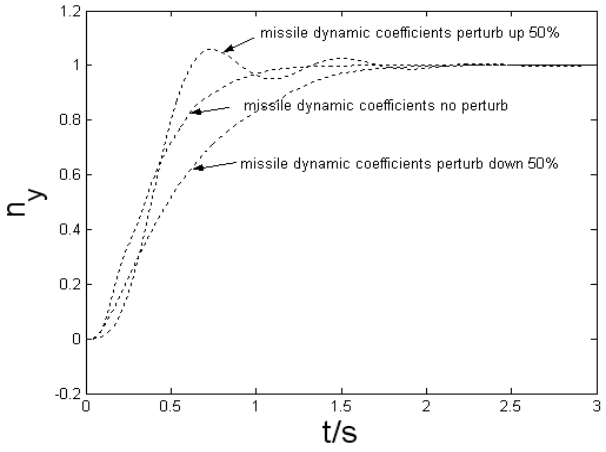
$$\tilde{\xi}=\begin{bmatrix}-\frac{g}{v}n_y{}^c-k_2\dot{n}_y{}^c\\ \left(Q_3k_2-\frac{gM_{\delta_z}}{vZ_{\delta_z}}\right)n_y{}^c\\ \left(P_2-\frac{v}{g}Z_\alpha w_\alpha k_2\right)n_y{}^c-\dot{n}_y{}^c\end{bmatrix}^T$$

where $M_\alpha, M_{\delta_z}, Z_\alpha, Z_{\delta_z}$ are missile dynamic coefficients, $\tilde{f}$ and $\tilde{\xi}$ are stir vector, which include all couple chatter influence of pitch channel.

Substituting above algorithm into Eq.(9), suppose expectation of original output n_y is $n_y^c=1$, when anti-ship missile achieve stable, angle acceleration ω_z is zero, so its expectation is considered as zero.

5. SIMULATION RESEARCH

An example of one axis symmetry supersonic anti-ship missile overload output tracking control is used to simulation. Missile motor has worked for 4 seconds, initials conditions are: $\alpha=0^\circ, \beta=0^\circ, \gamma=0^\circ, \omega_x=0^\circ, \omega_y=0^\circ, \omega_z=0^\circ$, altitude is 1000m, initial velocity is 649.26m/s, controller parameter: $\eta=3.25$, $c=\left[2.0, 1.46, 4.7\right]$. Considering missile dynamic coefficient perturbs 50%, then a group of missile overload n_y, attack angle α and sliding model variable $\boldsymbol{S}$ response curves are shown in Fig.1 and Fig. 2.



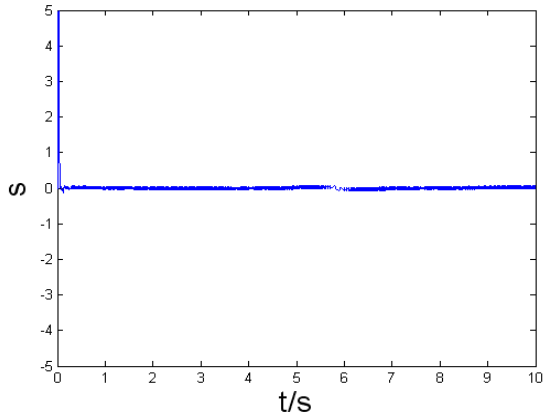


Fig. 1. Response curves of overload n_y

Fig. 2. Response curves of sliding mode variable S

6. CONCLUSIONS

This paper proposes a novel nonlinear control algorithm combing dynamic inverse, adaptive fuzzy neural network and sliding model control, adaptive adjustment law with the capacity of compensatory overcomes error brought by dynamic inverse feedback linearization. Sliding mode control component and robustness control component guarantee the system robustness. And above algorithm is adapted to one supersonic speed anti-ship missile overload output tracking control, the responding coordinate transfer realizes system zero dynamic stabilization, while the impact of system stability aroused by couple among channels disturbance is considered. Simulation results show the proposed algorithm possesses satisfactory tracking performance and guaranteed robust stability.

7. REFERENCES

[1] J.H. Ryu, M.J. Tahk, “Plant Inversion Control of Tail-Controlled Missiles”. Proc of the AIAA Guid.

Contr. Dyn. Conference. New Orleans, Louisiana, 1997. 1691-1696.

[2] M.Q. Hu, J.P. Wang, "Application of adaptive fuzzy-slide model control in reconfigurable flight controller". Acta Aeronautica et Astronautica Sinica, china.2002, 23(6): 11

[3] Y.A. Zhang, Y.A. Hu, "Nonlinear Design Approaches for Missile Control and Guidance". Beijing National defense backlist. China. 2003,54-55.

[4] J. Li, D.M. Xu, Adaptive "Sliding Mode Controller for Nonlinear Systems with Mismatched Uncertain ties Based on Adaptive Backstepping Scheme". Control and Decision, china. 1999, 14(1), 46-50.

[5] W. Y. Wang, C. Y. Cheng, et al, "An online GA-based output-feedback direct adaptive fuzzy-neural controller for uncertain nonlinear systems". IEEE Transactions on Systems, Man, and Cybernetics, 2004, 34(1): 334-345.

[6] W. Y. Wang, M. L. Chan, et al, "H∞ tracking-based sliding mode control for uncertain nonlinear systems via an adaptive fuzzy-neural approach". IEEE Transactions on Systems, Man, and Cybernetics, 2002, 32(4): 483-492.

[7] J. H. Ryu, C. S. Park, et al, "Plant inversion control of tail-controlled missiles". AIAA-97-3766, 1997.

[8] J. I. Lee, I. J. Ha, "Autopilot design for highly maneuvering STT missiles via singular perturbation-like technique". IEEE Transactions on Control System Technology, 7(5): 527-541.

Time Scale Risk-Sensitive Hierarchical Structure Control Problem*

Changming[1, 2] Yin, Huanwen[1] Chen, Lijuan[1] Xie
[1]College of Computer and Communicational Engineering, Changsha University of Science and Technology
Changsha, Hunan Province 410076, P.R.China
[2]Department of Automatic Control, LTH, Lund University
Lund, 22100, Sweden
Email: yinchm@csust.edu.cn

ABSTRACT

Hierarchical structure control problem solving is one of the keys to study a controlled Agent who learns its environment efficiently with large state space. It is more and more important to some practical applications in large state space control problems because of requiring an Agent to fit more complex environment, specially, in the area of studying for machine learning. We always regard the Markov Decision Processes (MDP) as the environment model of a reinforcement learning Agent. Bellman's optimal control equation is the base to solving this problem in simulating experiments and practical applications. We have introduced a kind of more complex and more practical environment for learning Agent in our previous work. Combining two concepts of risk-sensitive and multi-time scale, we have proposed a new conception which we refer it to multi-time scale risk-sensitive Markov Decision Processes. Under this new conception, we have modeled the basic Bellman's optimal control equation. Our motivation in this paper is to investigate this problem continually and gives a set of basic results. These results are all cores for framework of solving multi-time scale risk-sensitive control problems.

Keywords: Hierarchical Structure Control, Markov Decision Processes, Multi-time Scale, Risk Sensitive, Bellman Equation.

1. INTRODUCTION

How to solve a hierarchical structure control problem becomes one of the keys to study the Agent who learns its environment efficiently with large-scale state space [4]. It is more and more important to some practical applications in large state space control problems because of requiring the Agent to fit more complex environment, specially, in the area of studying for reinforcement learning. We always regard the Markov Decision Processes (MDP) as the environment model of reinforcement learning Agent [4]. Bellman's optimal control equation is the base to solving this problem in simulating experiments and practical applications.

H.S. Chang, P. Fard, S.I. Marcus and M. Shayman proposed multi-time scale Markov decision processes in their paper [1]. D. Hernandez-Hernandez and S.I. Marcus have concerned with risk sensitive Markov decision processes [3] for a long time and obtained a series of novel results. Their works are to be good for researching reinforcement learning problem [4]. We have introduced a kind of more complex and more practical environment for learning Agent in [5]. Combining two concepts of risk-sensitive from [3] and multi-time scale from [1], we have proposed a new concept which we refer to multi-time scale risk-sensitive Markov Decision Processes [5]. Under this new problem, we have modeled the basic Bellman's optimal control equation. Our motivation in this paper is to investigate this problem continually and gives some basic results. These results are all cores for framework of solving multi-time scale risk-sensitive control problem.

2. UTILITY FUNCTIONS AND PERFORMANCE

Define 1(Utility Functions [2, 3]) A fixed real number $\lambda \in \mathbf{R}$ (real set) stands for a constant risk-sensitive coefficient. Its corresponding utility function U_λ: $\mathbf{R} \to \mathbf{R}$ is defined as following: for $x \in \mathbf{R}$,

if $\lambda \neq 0$, then $U_\lambda(x) = \mathrm{sign}(\lambda) e^{\lambda x}$; else,

if $\lambda = 0$, then $U_\lambda(x) = x$.

It is not difficult to verify that this function is increasing strictly, and also satisfies that

$$U_\lambda(c + x) = e^{\lambda c} U_\lambda(x), \lambda \neq 0, x, c \in \mathbf{R}.$$

We assume that the Agent obtains a random reward R via the expectation of function $U_\lambda(R)$ in this paper. Given a bounded random reward R, the corresponding certainty equivalent $E[\lambda, R]$ with respect to $U_\lambda(.)$ is defined implicitly by $U_\lambda(E(\lambda, R)) = E[U_\lambda(R)]$. Combining the definition of utility function with it together we can yield that

$$E(\lambda, R) = \begin{cases} \frac{1}{\lambda} \log(E[e^{\lambda R}]), & \lambda \neq 0; \\ E[R], & \lambda = 0. \end{cases}$$

Using Jensen's inequality it follows that, when R is not a constant, then $E(\lambda, R) > E[R]$ (respectively $E(\lambda, R) < E[R]$), if $\lambda > 0$ (respectively $\lambda < 0$). A controller rating a random reward R according to the expectation of $U_\lambda(R)$ is referred to as risk-averse when $\lambda > 0$, and risk-seeking if $\lambda < 0$; if $\lambda = 0$, the decision maker is risk-neutral [3]. It notices that we will abuse the exponent notation e and exp.

3. MULTI-TIME SCALE RISK-SENSITIVE MDP

Here we consider two-time scale MDP model only being described in this paper [1] and then give some conceptions under risk-sensitive performance [2, 3].

High level (slow time scale) MDP with finite state space Z and finite action space D is proposed. At each decision epoch $n \in \{0, 1, 2\ldots\}$, the system is in state $i_n \in Z$, and applies the action $d \in D$, and then transfers its state from i_n to i_{n+1} according to probability distribution $P^u(i_{n+1} | i_n, d_n)$.

In the low level (fast time scale), all MDPs possess the same finite state space S and finite action space A, and $Z \cap S = \theta$, $D \cap A = \theta$, where θ denotes the null set.

We refer $t \in \{t_0, t1, t2\ldots\}$ to be epoch in fast time scale and $t_{nT} = n$ stands for epoch in slow time scale where T is a fixed finite scale factor between high and low levels.

* This work is supported by the NSFC No.60075019.

We also suppose $t_{nT} = n^+$, that is to say, there is a very small difference between t_{nT} and n , such that the decision at t_{nT} in fast time scale drops behind a little at n in slow time scale to be made [1].

We suppose that the initial state of low and high level MDP is $x \in S$ and $i \in Z$ respectively, i.e., $x_{t0} = x$, $i_0 = i$ when $n = 0$. High level MDP applies action $d_0 \in D$ at state i_0, and the next state i_1 is confirmed by probability P^u. At epoch t_0, t_1, …, t_{T-1}, the system is evolved in terms of low level MDP. More accurately speaking, for state $x \in S$ at epoch t_0, the system applies action $a \in A$, and then transfers to the next state $y \in S$ at epoch t_1 according to probability $P^l(y|x,a,i,d)$, and receives a non-negative bounded reward $R^l(x,a,i,d)$. The system repeats this process for state $y \in S$ at time t_1 until it reaches to epoch t_{T-1}.

In this paper, we will refer to the decision rule as relating infinite problem, and policy as finite one. Here we define a decision rule in low level as $g^l = \{\pi(l, n)\}$ which stands for a non-stable policy serial with T epochs, where $n = 0, 1, 2, \ldots$, and $\pi(l, n) = \{\varphi(t_{nT}), \ldots, \varphi(t_{(n+1)T-1})\}$ is a function serial for all of n. For all of $k \geq 0$, we define a function as $\Phi(t_k)$: $S \times Z \times D \to A$

We say a decision rule in low level is stable about slow time scale n if $\pi(l, n) = \pi(l, n')$ for all of n and n' . All of possible stable decision rules about slow time scale like this are denoted by a set Q, and we omit the index letter n in $\pi(l, n)$ and denote it in π^l. So, we can use the epochs t_0, $t_1 \ldots t_{T-1}$, to denote the function serial of π^l, and Π^l stands for the set of all of T epoch non-stable polices π^l. We will omit the subscript in φ also if π^l is stable.

Given a low level decision rule $g^l \in Q^l$, for non-negative bounded immediate reword function I^u on $Z \times D$ in low level, and all $n \geq 0$; $x \in S$, $i_n \in Z$ and $d_n \in D$, we define a function R^u as

$$R^u(x, i_n, d_n, \pi^l) = E^x_{i_n, d_n}\left\{\sum_{t=t_{nT}}^{t_{(n+1)T-1}} R^l\left(x_t, \varphi_t\left(x_t, i_n, d_n\right), i_n, d_n\right)\right\} + I^u\left(i_n, d_n\right).$$

Where the superscript x in E stands for the initial state $t_{nT} = x$ of low level, and the subscripts i_n, d_n are fixed for this expectation, φ_t means $\varphi(t)$. Function R^u is the total reward for T epochs. That's that this bounded total reward includes two parts, one is a reward function with 0 terminate reward and obeying T epochs non-stable policy π^l under the condition of giving $i_n \in Z$, $d_n \in D$, and initial state $x \in S$, the other is a immediate reward for high level at state i_n taking an action d_n.

The total expectation reward obtained for T epochs non-stable policy in low level reacts on a single step reward of high level MDP. We define a high level stable decision rule g^u as a function g^u: $S \times Z \to D$, and denote all possible of these high level stable decision rule as a set Q^u. Given the initial states $x \in S$, $i \in Z$ of two levels respectively, the goal of Agent is to get a decision rule pairs $g^l \in Q^l$, $g^u \in Q^u$ that are determined by a value of function defined on $S \times Z$ as following

$$V_\lambda(g^l, g^u; x, i) = \frac{1}{\lambda}\log\left(E^{x,i}_\lambda\left[e^{\lambda\sum_{n=0}^{\infty} R^u\left(x_{t_{nT}}, i_n, g^u\left(x_{t_{nT}}, i_n\right), \pi^l\right)}\right]\right)$$

According to the definition of utility function, this formula can be recast by

$$U_\lambda\left(V_\lambda(g^l, g^u; x, i)\right) = E^{x,i}_{\pi^l}\left[U_\lambda\left(\sum_{n=0}^{\infty} R^u\left(x_{t_{nT}}, i_n, g^u\left(x_{t_{nT}}, i_n\right), \pi^l\right)\right)\right]$$

and optimal value function of the system can be defined by the following formula. Note that the notation V^*_λ stands for two levels optimal value function with infinite epochs.

$$V^*_\lambda(x, i) = \sup_{g^u \in Q^u}\sup_{g^l \in Q^l}\left\{V_\lambda(g^l, g^u; x, i)\right\}$$

$$= \sup_{g^u \in Q^u}\sup_{g^l \in Q^l}\frac{1}{\lambda}\log\left(E^{x_{t_{nT}}, i_n}_{\pi^l}\exp\left(\lambda\sum_{n=0}^{\infty} R^u\left(x_{t_{nT}}, i_n, g^u\left(x_{t_{nT}}, i_n\right), \pi^l\right)\right)\right)$$

$$= \sup_{g^u \in Q^u}\sup_{g^l \in Q^l}\frac{1}{\lambda}\log E^{x_{t_{nT}}, i_n}_{g^u}\left[\exp\left(\lambda\left\{\sum_{n=0}^{\infty} E^{x_{t_{nT}}}_{i_n, d_n}\right.\right.\right.$$

$$\left.\left.\left.\left[\sum_{t=t_{nT}}^{t_{(n+1)T-1}} R^l\left(x_t, \varphi_t\left(x_t, i_n, g^u\left(x_t, i_n\right)\right), i_n, g^u\left(x_t, i_n\right)\right)\right] + I^u\left(i_n, d_n\right)\right\}\right)\right].$$

4. SOME BASIC RESULTS FOR HIERARCHICAL OPTIMAL CONTROL PROBLEM

This paper discusses hierarchical optimal control problem for estimating the optimal function V^*_λ. Under fixed state and action in high level, we denote all of possible low level non-stable polices with T epochs as a set defined in the following: for a given pairs of state/action, $i \in Z$ and $d \in D$,

$$\Pi^l[i,d] \triangleq \left\{\pi^l[i,d] \middle| \pi^l[i,d] \triangleq \left\{\phi^{i,d}_{t_0}, \cdots, \phi^{i,d}_{t_{T-1}}\right\}, \phi^{i,d}_{t_k} : S \times \{i\} \times \{d\} \to A, k = 0, \cdots, T-1\right\}$$

We denote $P^T_{xy}(\pi^l[i,d])$ to be a probability which starts at initial state x and reaches to state $y \in S$ within T steps according to T epochs non-stable policy $\pi^l[i,d]$. It takes notice of that this probability can be gained by P^l. The following theorem about Bellman optimal equation with multi-time scale risk-sensitive performance has been proved in our paper [5].

Theorem 4.1 For all x, $y \in S$, $d \in D$, and i, $j \in Z$, the functions $R^u(x,i,d,\pi^l[i,d])$, $P^T_{xy}(\pi^l[i,d])$ and $P^u(j|i,d)$ are all measurable, and we also assume that the value function is bounded. Then the optimal value function follows λ-optimal equation as following:

$$U_\lambda\left(V^*_\lambda(x,i)\right) = \sup_{d \in D}$$

$$\left(\sup_{\pi^l[i,d] \in \Pi^l[i,d]}\left\{e^{\lambda R^u\left(x,i,d,\pi^l[i,d]\right)}\sum_{y \in S}\sum_{j \in Z} P^T_{xy}\left(\pi^l[i,d]\right)P^u\left(j \middle| i,d\right)U_\lambda\left(V^*(y,j)\right)\right\}\right)$$

Next, we will give some theorems in the following which are all basic for studying hierarchical structure control problem.

Theorem 4.2 For all x, $y \in S$, $d \in D$, and i, $j \in Z$, the functions $R^u(x,i,d,\pi^l[i,d])$, $P^T_{xy}(\pi^l[i,d])$ and $P^u(j|i,d)$ are all measurable. We set the function W: $S \times Z \to [0, \infty]$, such that

$$U_\lambda(W(x,i)) \geq \sup_{d \in D}$$

$$\left(\sup_{\pi^l[i,d] \in \Pi^l[i,d]}\left\{e^{\lambda R^u\left(x,i,d,\pi^l[i,d]\right)}\sum_{y \in S}\sum_{j \in Z} P^T_{xy}\left(\pi^l[i,d]\right)P^u\left(j \middle| i,d\right)U_\lambda\left(W(y,j)\right)\right\}\right)$$

Then we have $W(x,i) \geq V^*_\lambda(x,i)$.

In the remain part of this paper we assume that if n is a relative stop time [2] and the mapping W: $S \times Z \to \mathbf{R}$ is a given real value function of the states pair, then $W(x_{t_{nT}}, i_n) = 0$ at the event $[n = \infty]$.

Theorem 4.3 For each non-negative integer n, we set F_n be a σ-field generated by trajectory (or history, we denote it h_n, and use H_p to stand for the set of all histories [4]). We also set a positive stochastic variable m be the relative stopping time about $\{F_n\}$. That is to say that $\Pr^{x,i}_{\pi^l}[m \in \{1,2,3,\cdots\} \bigcup \{\infty\}] = 1$ for every state $x \in S$, $i \in Z$ and policy $\pi^l = \pi^l[i, d] \in \Pi^l[i, d]$. For $k \in \mathbf{N} \setminus \{0\}$, the event $[m = k]$ belongs to F_k. Under these conditions above, we can obtain the following strong λ-optimality equation:

$$U_\lambda(V_\lambda^*(x,i)) = \sup_{\pi^l \in \Pi^l[i,d]}$$

$$E_\pi^{x,i}\left[\exp\left(\lambda\sum_{k=0}^{T-1} R^u\left(x_{t_k}, i_k, g^u\left(x_{t_k}, i_k\right), \pi^l[i,d]\right)\right) U_\lambda(V_\lambda^*(x_{t_{mT}}, i_m))\right],$$

$$x \in S, i \in Z, d \in D, \pi^l = \pi^l[i,d] \in \Pi^l[i,d].$$

Proof: We first set the positive variable n to be a relative stopping time, and the real number $\varepsilon>0$ is arbitrary. For each $x \in S, i \in Z$, a policy $\pi^l(\varepsilon,x,i) \in \Pi^l[i,d]$, which relatives to two levels states are chosen such that

$$V_\lambda(\pi^l(\varepsilon,x,i),x) \geq V_\lambda^*(x,i) - \varepsilon.$$

We also define a new policy δ^l as following: given $p \in \mathbf{N}$ and $h_p \in H_p$ (the history set),

$$\delta_p^l(\cdot \mid h_p) = \begin{cases} \pi_p^l(\cdot \mid h_p), & \text{if} \quad n > p; \\ \pi_{p-k}^l\left(\varepsilon, x_{t_{kT}}, i_k\right)\left[\cdot \mid x_{t_{kT}}, i_k, a_{t_{kT}}, \cdots, x_{t_{pT}}\right], & \text{if} \quad n = k \leq p. \end{cases}$$

On the other word, the Agent is driven by policy δ^l, and it choices an action according to the policy π^l, until to the relative time n, and then if this process begin again, the Agent will transfer the policy to $\pi^l\left(\varepsilon, x_{t_{nT}}, i_n\right)$, as if the system is reset to the beginning. Using the definition of δ^l and the increasing property of utility function, and Markov property, we can obtain that, for each $x \in S$, $i \in Z$, $k \in \mathbf{N}/\{0\}$,

$$E_{\delta^l}^{x,i}\left[I[n=k] U_\lambda\left(\sum_{m=0}^{\infty} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \delta^l[i,d]\right)\right)\right.$$

$$\left.\left| x_{t_0}, a_{t_0}, i_0, d_0, \cdots, x_{t_{(k-1)T}}, a_{t_{(k-1)T}}, i_{k-1}, x_{t_{(k-1)T+1}}\right.\right] \geq I[n=k]$$

$$\exp\left(\lambda\sum_{m=0}^{k-1} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \delta^l[i,d]\right)\right) U_\lambda(V_\lambda^*(x_{t_{kT}}, i_k) - \varepsilon).$$

Where the letter I denotes the indictor function. Therefore, we imply that

$$E_{\delta^l}^{x,i}\left[I[n=k] U_\lambda\left(\sum_{m=0}^{\infty} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \delta^l[i,d]\right)\right)\right] \geq e^{-\lambda\varepsilon}$$

$$E_{\delta^l}^{x,i}[I[n=k] \exp\left(\lambda\sum_{m=0}^{n-1} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \delta^l[i,d]\right)\right) U_\lambda(V_\lambda^*(x_{t_{nT}}, i_n))].$$

Where the equation above comes from the fact that two policies δ^l and π^l are consentaneous before the stopping time n. Again, we can get that, for all $x \in S$,

$$U_\lambda(V_\lambda^*(x,i))$$

$$\geq e^{-\lambda\varepsilon} E_{\pi^l}^{x,i}\left[\exp\left(\lambda\sum_{m=0}^{n-1} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)\right) U_\lambda\left(V_\lambda^*(x_{t_{nT}}, i_n)\right)\right].$$

Because of $x \in S, i \in Z$, $\varepsilon > 0$ and the arbitrariness of the policy π, we can obtain that, for all $x \in S$,

$$U_\lambda(V_\lambda^*(x)) \geq \sup_{\pi^l \in \Pi^l[i,d]} E_{\pi^l}^{x,i}\left[e^{\lambda\sum_{m=0}^{n-1} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)} U_\lambda\left(V_\lambda^*(x_{t_{nT}}, i_n)\right)\right].$$

To obtain the inverse inequality, we set the positive stochastic variable n be relative stopping time mentioned foregoing and for each policy $\pi^l = \pi^l[i,d] \in \Pi^l[i,d]$, $k \in N$, and all states $x \in S$, $i \in Z$, using Markov property and formula above, we can imply that

$$E_{\pi^l}^{x,i}\left[U_\lambda\left(\sum_{m=0}^{\infty} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)\right)\right.$$

$$\left.I[n=k] \left| x_{t_0}, a_{t_0}, i_0, d_0, \cdots, x_{t_{(k-1)T}}, a_{t_{(k-1)T}}, i_{k-1}, x_{t_{(k-1)T+1}}\right.\right]$$

$$\leq \exp\left(\lambda\sum_{m=0}^{n-1} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)\right) I[n=k] U_\lambda(V_\lambda^*(x_{t_{kT}}, i_k)),$$

where the transferred policy $\tilde{\pi}^l$ is determined by following formula,

$$\tilde{\pi}_n^l(\cdot \mid h_n) = \pi_{n+k}(\cdot \mid x_{t_0}, a_{t_0}, i_0, d_0, \cdots, x_{t_{(k-1)T}}, a_{t_{(k-1)T}}, i_{k-1}, x_{t_{(k-1)T+1}}), n \in \mathrm{N}, h_n \in H_n.$$

So, we can imply that

$$E_{\pi^l}^{x,i}\left[U_\lambda\left(\sum_{m=0}^{\infty} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)\right) I[n=k]\right]$$

$$\leq E_{\pi^l}^{x,i}\left[\exp\left(\lambda\sum_{m=0}^{n-1} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)\right) I[n=k] U_\lambda(V_\lambda^*(x_{t_{nT}}, i_n))\right].$$

Using this inequality, we can obtain the following:

$$E_{\pi^l}^{x,i}\left[U_\lambda\left(\sum_{m=0}^{\infty} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)\right)\right]$$

$$\leq E_{\pi^l}^{x,i}\left[\exp\left(\lambda\sum_{m=0}^{n-1} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)\right) U_\lambda(V_\lambda^*(x_{t_{nT}}, i_n))\right].$$

Applying the operator *sup* on policy π^l and using the λ-optimal equation, we can imply that, for each $x \in S$, $i \in Z$,

$$U_\lambda(V_\lambda^*(x,i)) = \sup_{\pi^l \in \Pi^l[i,d]} E_{\pi^l}^{x,i}\left[U_\lambda\left(\sum_{m=0}^{\infty} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)\right)\right]$$

$$\leq \sup_{\pi^l \in \Pi^l[i,d]} E_{\pi^l}^{x,i}\left[\exp\left(\lambda\sum_{m=0}^{n-1} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)\right) U_\lambda(V_\lambda^*(x_{t_{nT}}, i_n))\right].$$

Combining this formula above with other side inequality, we can get the final result in the theorem and finish this proof.

According to the foregoing assumption, we can know that the expectation value of the right hand of λ-optimal equation in theorem 4.1 is

$$E_{\pi^l}^{x,i}\left[\exp\left(\lambda\sum_{m=0}^{n-1} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)\right) U_\lambda(V_\lambda^*(x_{t_{nT}}, i_n)) I[n<\infty]\right] +$$

$$E_{\pi^l}^{x,i}\left[\exp\left(\lambda\sum_{m=0}^{\infty} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)\right) U_\lambda(0) I[n=\infty]\right].$$

Because the theorem above is very important, so, we prove it in detail. It notes that, when the relative stopping time n equals one, the formula of λ-optimal equation in theorem 4.1 can derive the general λ-optimal equation in [5].Now, given a subset G of $S \times Z$, we refer T_G to be the first positive time reaching to state set G, that is to say,

$$T_G = \min\left\{n > 0 \mid \left(x_{t_{nT}}, i_n\right) \in G\right\}.$$

It notices that T_G is a relative time and this symbol will be used in next theorem. The next theorem shows that the value function $V_\lambda(\pi^l, x, i)$ can be approached to the reward which will be obtained that the system settles in a finite subset of state space. This property is very important for those problems with large state space.

Theorem 4.4 We set that the reward functions and the state transfer probability functions are all measurable, and the reward functions are non-negative. We also set $x_{t0} \in S$, $i_0 \in Z$, $\pi^l = \pi^l[i,d] \in \Pi^l[i,d]$, and give a fixed number $\varepsilon>0$. Then, there exists a finite subset F of $S \times Z$, such that

$$(x_{t_0}, i_0) \in F, \text{and} \quad U_\lambda^{-1}\left(E_{\pi^l}^{x_{t_0}, i_0}\left[U_\lambda\left(\sum_{m=0}^{T_{F^c}-1} R^u\left(x_{t_{mT}}, i_m, g^u\left(x_{t_{mT}}, i_m\right), \pi^l[i,d]\right)\right)\right]\right)$$

$$\geq V_\lambda(\pi^l[i,d], x_{t_0}, i_0) - \varepsilon.$$

Where $F^c = (S \times Z)/F$.

5. CONCLUSIONS

This paper uses the new concept about multi-time scale risk-sensitive Markov decision processes put forward in [5]. Putting this Markov decision model as environment model of an Agent, we can conclude that an Agent is more fit for its learning requires about hierarchy control problem. We have proved some theorems above, which are the core for investigating the hierarchical control problem with multi-time scale and risk-sensitive performance. Satisfying these conditions pointed in the theorems of this paper, we will obtain the theory base to solve a series of hierarchical learning problems; especially it is possible to investigate more deeply for multi-time scale risk sensitive Markov decision processes problems. Using this theory, we have a new idea for researching reinforcement learning problems of an Agent with large-scale state space. We will investigate this problem latter.

REFERENCES

[1] H.S.Chang, P.Fard et al, "Multi-time Scale

Markov Decision Processes". IEEE Tran. On Automatic Control, 48 (June 2003), 976-987.

[2] S.Coraluppi, S.I.Marcus, "Mixed Risk-Neutral Minimax Control of Markov Decision Processes". Proc. 1997 Conf. on Info. Sciences and Systems, March 1997, Johns Hopkins University, Baltimore, MD, pp. 435-440.

[3] D.H.Hernandez, S.I. Marcus, "Corrigendum to Risk Sensitive Control of Markov Processes in Countable State Space". System and Control Letters, vol. 34(1998): 105-106.

[4] R.S. Sutton, A. G. Barto. "Reinforcement Learning: An Introduction". MIT Press, Cambridge, MA, 1998.

[5] C.M. Yin, H.X. Wang et al, "Optimal Equality for Multi-Time Scale Risk-Sensitive Markov Decision Processes". Proceedings in the International Symposium on Computer Science and Technology 2005, Ningbo, China.

Design of Stress-Strain Virtual Measurement System for Bionic Plough

Jianqiao Li [1], Shucai Xu [1, *], Zhanrong Cui[1], Rui Zhang[1, 2]
1. Key Laboratory of Terrain-Machine Bionics Engineering (Jilin University), Ministry of Education
Changchun, Jilin, 130025, China
Email: xushc@jlu.edu.cn
2. Department of Modern Mechanics, University of Science and Technology of China
Hefei, Anhui, 230026, China

ABSTRACT

The experimental complexities and the interferential factors influencing on signal logging were analyzed for bionic plough worked in the field and laboratory. During design, the anti-jamming methods including hardware and software technology were adopted. Based on hardware of DAQP-12 and QTC-300 data logging sets, and notebook computer, a stress-strain virtual measuring system for bionic plough, which was designed by the graphical programming language DASYLab, was used to measure the horizontal force F acting on the surface of bionic plough by soil. Moreover, the theory of entropy and variance relation was used to analyze the distribution law of the experimental signals after the field experiment, here obtained $\log A$ is 1.238. In the end, the experimental results show that measured data conforms to the statistical rule distribution and the error of the whole test system is under 5 percent, so the tress-strain virtual measuring system was stable and credible.

Keywords: Virtual instrument; stress-strain; data acquisition; anti-jamming; Ratio of performance to price.

1. INTRODUCTION

Aimed at many interferential factors influencing on signal gathering in the field experiment, it is very important how to design the measurement system having high precision and automaticity. Even if digital instruments and intelligent instruments improve the measurement accuracy and function of the traditional measurement and analysis system, the traditional measurement system has some obvious disadvantages. First of all, the system is close, and it depends on some function-fixed hardware, which results in high cost to develop its functions. Secondly, the equipment is large and ponderous, and it needs network power, which brings inconvenience while doing experiments in the field. And furthermore, the singleness of function and channel leads to high exploitation cost. Thirdly, the mode of data display and memory format is single, which brings less information. Finally, the methods used in data post processing and analyzing can not dissatisfied the demand of the experiment. According to the above demerits, a kind of virtual instrument developed with PC and correlative software emerged. The arising of virtual instrument was a challenge against traditional instruments, and simultaneously was also a technology innovation and a brand new trend in the measuring field. The virtual instruments based on computer technology have improved many demerits of the traditional instruments and replaced them depended on hardware gradually. The virtual instrument is usually called "Software is Instrument", it can adopted the anti-jamming methods including hardware and software technology to decree the interfere factors to signal gathering, such as electrostatic, magnetic field and random error. The virtual instruments can also use the battery power supply [1, 2, 3, 4, 5, 6, 7]. In view of the merits of the virtual instruments, a stress-strain virtual measuring system was designed by graphical programming language DasyLab and it was used to analyze the horizontal force acting on the bionic ploughs in the field experiment [8, 9].

2. DESIGN OF VIRTUAL STRESS-STRAIN MEASURING SYSTEM

When the stress-strain virtual measuring system was used to test the horizontal force acting on the bionic ploughs, the data acquisition principle and the anti-jamming methods of different design stages were shown in Fig. 1., according to the flow of signal.

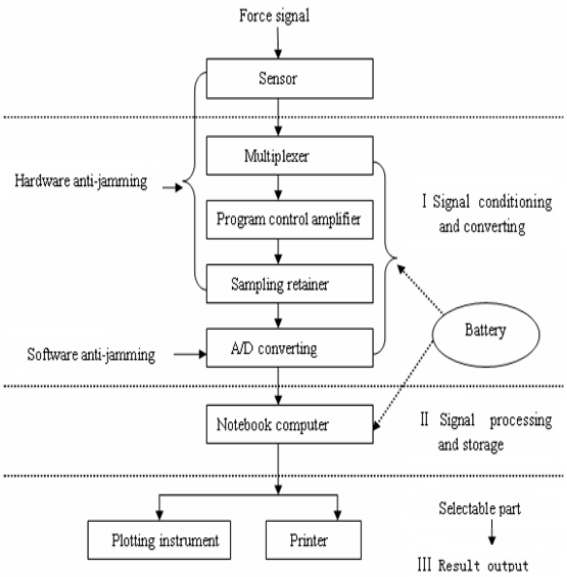


Fig. 1. Diagram of the data acquisition principle and the anti-jamming method

To improve the measuring accuracy and analyze the result of experiment with the theory of error and statistic, three demands were put forward as follows: firstly, the anti-jamming measures of hardware and software technology were used for suppressing the interferences through correctly analyzing the interference source; Secondly, size of sample must be same in every experiment, that is to say, the virtual measuring system had a fixed max degree of freedom; Thirdly, the effect of random noise could be minimized. Moreover, the sensor

* Corresponding author: Shucai Xu

could probably bring about errors, so the sensor must be validated for reducing them and realizing transformation between physical quantity and analog. The equation of validated-sensor was y=4.43x+0.015 after calibration.

2.1 Hardware Components and Configuration

From the aspect of hardware, the stress-strain virtual measuring system was constituted of computer, DAQP-12 and QTC-300 data logging sets as shown in Fig. 2..

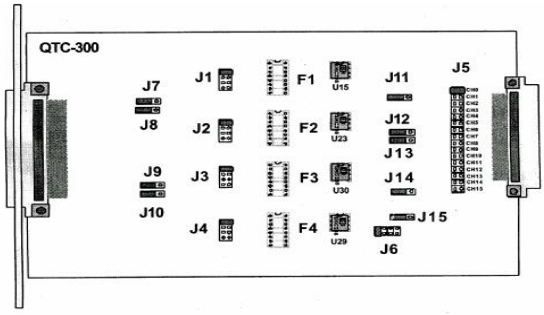


Fig. 2. Function location diagram of QTC-300

The functions of signal conditioning and converting were completed by the QTC-300 and daqp-12 commonly. The QTC-300 is a four-channel strain gage signal-conditioning adapter. Each channel is equipped with separate instrumentation amplifier, multiple excitation current source options and a fourth order low pass filter. Each channel has selectable gain of 1, 10,100 or 500. For the QTC-300 to accurately measure strain, it must be calibrated using the DaqCal Utility and configured correctly when a new sensor is added to the board, or a previously used sensor is reconnected to the board every time. Fig. 2. shows the location of the main A/D channel and board selection jumper blocks (J5/J6), SSH option jumper block (J11), A/D gain selection jumper blocks (J1 through J4 for input channels 0 through 4 respectively), filter selection jumper blocks (J12 through J15), filter blocks (F1 through F4) and the AC/DC coupling selection jumpers (J7 through J10). The settings for J5 and J6 are CH0 and BD0. The option for J11 is SSH disabled. Because the 0V value on data acquisition board corresponds to 0 mV at the pre-amplifier, the configuration for J1 is 1000. Using the A/D gain selection option provided on the QTC-300 interface board would result in less signal noise than using the A/D gain selection options available on the data acquisition adapter in the host computer. For filtering the high frequency disturbing signal, J12 and F1 are set LP and 80Hz. As shown in Fig. 1, DAQP-12 is a 12-bit analog input PCMCIA card. The parameters of daqp-12 were set using DAQDRIVE Config Utility. Base address and IRQ Level of DAQP-12 are set for 10A0 and 15. The channels, input mode and signal type of A/D Converter are configured as 15, single-ended and bipolar respectively. The select driver of DAQP-12 is "daqlab32.dll". From the aspect of interference, the alternate magnetic field coursed by the different electrical installations was a main interference source. Additionally, changes in the internal electric network, the surge aroused by the operation of switches, starting/stopping of the bigger electrical installations and harmonics brought about by AC/DC transmission system, all took bad effect on the measuring system. Aimed at the above interferences, the anti-interference technology of hardware could stop the source of them effetely, and block the transmission channels of them, so the two follow anti-jamming measures were adapted: ①Set up the wave filter to suppress the noise through J12 and F1 of QTC-300; ②Select the shielding body for using and isolating the electric field which has opened the space, the parts of magnetic field or electric magnetic field coupling, cut off the coupling passway of its space field. To make the shielding anti-interference result relatively excellent, need to accomplish the following two things: Firstly, guarantee the shielding layer with keeping in touch well; Secondly, the double twisted connecting wire was adopted to reduce the power space and interfere in second time [10,11,12,13]. In addition, for reflecting the stress-strain virtual measuring system had good expansibility, the signal conditioning adapter adopted was a series, therefore, the virtual measuring system can realize many other measurement functions, such as measuring temperature, velocity, and acceleration, via substituting corresponding signal conditioning adapters for QTC-300. Meanwhile, the channels of the virtual measuring system can be expended multiply through increasing the amount of the same signal-conditioning adapter. However, the traditional system for stress-strain changes the function or the channels through adding the amount of the NEC16 strain gauge made in Japan, as shown in Fig. 3.. The price of NEC16 strain gauge is 300 thousand RMB, so the total price of the traditional measuring system for stress-strain and the data logger is 380 thousand RMB. However, the total price of the virtual measuring system for stress-stain is 40 thousand RMB. Therefore, compared with the former, the latter indicates more superiority on the extensibility and price obviously.

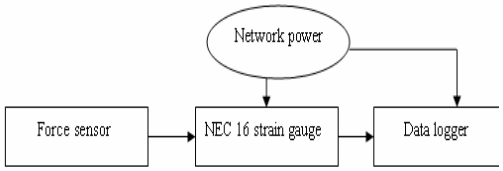


Fig.3. Principium diagram of traditional stress-strain measuring system

2.2 Design of Software

The virtual measuring system for stress-strain, which was used to test the horizontal force acting on the bionic ploughs, was designed by the graphical programming language DasyLab with functions including data acquisition, display, storage, analysis, and control. The software of virtual stress-strain measuring system was designed by choosing and configurating the function modules to realize the measurement function illustrated in Fig. 4..

A/D module was selected firstly, and then the settings for hardware and encouragement voltage of sensor were configured for expansion board QTC-300 and –10v to +10v in its property dialog box. To weaken some of interference in the input signal, two modules of Filter were selected for making up the adjustable digital filter in the design, the cutoff frequencies of filter was configured by the result that was analyzed by the spectrometer made up of two function modules, FFT and Y/t. Before testing, in order to eliminate the zero floats of sensors, two modules, Dig. Meter and scaling, were selected to deal

with them in the design. The parameters, a and b in the property dialog box of the second Scaling module, were set up for 4.43 and 0.015 to realize transformation between the digital signal and physics signal based on the validated parameters of sensor. The best method that dispels the random noise was fitting proposal in real time, so the Regression module was selected and set up for least square method to fit in the design. The arithmetic module, Arithmetic, was selected for use in the design, which could carry on operation including 1 /x, c/x, x/c, x +y, ∑xi etc to the test data according to the needs of the experiment results. To display data in many formats, the display modules, Y/t, Dig. Meter, and Bar Graph were selected, the output channel number and output range of them were all set for CH0 and –10v to +10v. Two write modules, whose file formats were configured for ASCII and *.dsb separately, were used to store data, the storage formats of them realized the data further analysis and replay conveniently. Two Switch modules were configured for Start/stop and Pause/continue to control the experiment status individually, as shown in Fig. 4.. In order to acquire the same sample capacity each time, the module, Stop, was adapted to control the sampling time, the property parameters of it were set up for follows: Action after: seconds, Action: stop, Parameter: 16. In order to design personal operation board and change the function parameters of above-mentioned modules expediently, the Action module was selected especially. Opening the property dialog box of Action module by double-clicking its icon, the parameters for channel 0 were specified (i.e. Event: Experiment Start, Receive module: <DASYLab> Channels (Empty=Module), Action: Layout Full Screen, Parameter Layout: 1), and the parameters for channel 1 were configured (i.e. Event: Experiment Stop, Receive module: <DASYLab> Channels (Empty=Module), Action: Worksheet Show) [14].

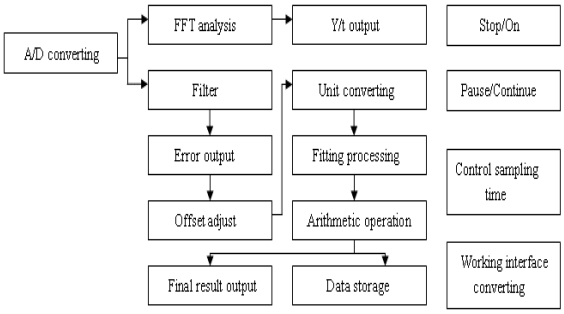


Fig. 4. Function diagram of software of virtual stress-strain measuring system

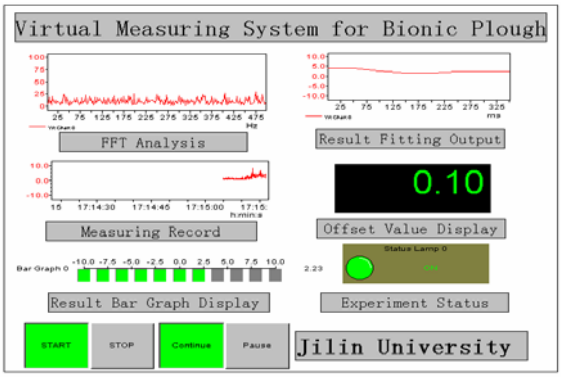


Fig. 5. Virtual operation panel of virtual stress-strain measuring system

Finally, it was important that all settings must be saved. From above all, the virtual operation panel was designed completely, as shown in Fig. 5..

3. RESULTS AND DISCUSSION

The virtual measuring system for stress-strain was used to measure the horizontal force acting on the bionic ploughs in the field. The experiments were done in the spring of 2004 in some countryside in Shulan, in northeast of China. The experimental conditions were that water content was from 25.65% to 27.49%, working speed was about 0.83 m/s, the depth of plough was from 150 mm to 200 mm, and width of plough was from 180mm to 200mm. Under the same experimental conditions, the acquired data in the same group was regarded to be measured at the same precision. One of six groups of acquired data was chosen to analyze under this experimental condition as listed in Table 1..

Table 1. Measure Data of Experiment Distribution

No	Measure value (R_i) (kN)	m_i
1	1.754	5
2	2.002	8
3	2.228	9
4	2.472	15
5	2.756	39
6	3.043	76
7	3.309	97
8	3.553	122
9	3.797	84
10	4.08	69
11	4.355	57
12	4.616	25
13	4.913	11
14	5.179	7
15	5.493	4

According to the gross error processing criterion, $\left|x_i - \bar{x}\right| > 3\sigma$ (where, xi was primitive data $\bar{x}$ was primitive mean value), the gross error of original data was rejected firstly this group of data was divided into q =15 groups at about 0. 27kN interval, m_i was the primitive acquire data integer included in each spacing 0.27kN, then mean value R_i of every group was filled in Tab.1. According to the relation between entropy and variance in Information Theory, the distribution of the measure result was calculated in the end [15]. The variable σ_R could be calculated from the Tab.1.
Mean value:

$$\bar{R} = (\sum_{i=1}^{15} R_i * m_i) / \sum_{i=1}^{15} m_i$$

$$= (1.754*5+2.002*8+\ldots+5.493*4)/628$$

$$=3.602 \qquad (1)$$

$$\sigma_R = \sqrt{\sum_{i=1}^{q} p_i^* (R_i - \overline{R})^2}$$

$$= \sqrt{\sum_{i=1}^{15} P_i^* (R_i - 3.602)^2} = 2.38 \qquad (2)$$

Where, P_i^* was the probability of every group data, which could be calculated from the equation ($m_i \Big/ \sum_{i=1}^{15} m_i$).

$$\log \sigma_R = \log 2.38 = 0.3751 \qquad (3)$$

$$H(R) = -\sum_{i=1}^{q} p_i^* \times \log p_i^*$$

$$= \frac{1}{628} \times$$

$$\left(5\log\frac{628}{5} + 8\log\frac{628}{8} + \ldots + 4\log\frac{628}{4} \right)$$

$$= 0.9946 \qquad (4)$$

$$A = 2 \times \left[H(R) - \sigma_R \right] = 2 \times \left[0.9946 - 0.3751 \right]$$

$$= 1.238 \qquad (5)$$

Based on reference [15], $\log A$ of normal distribution was 1.232. Because 1.238 was very close to 1.232, therefore, above set of measurements obeyed normal distribution approximately.

The stress-strain virtual measuring system was substituted for the traditional system successfully. The above analysis indicated that the experimental data accorded with the statistics law of the random error distribution basically. The methods of judging data distribution provide foundation for processing data further correctly.

4. CONCLUSIONS

The stress-strain measuring system for bionic plough based on virtual instrument technology indicated more merits than the traditional system on many aspects. Firstly, its functions, including instrument self-checking, data real-time acquire, data real-time storage, data processing and replay etc, were all completed by the computer, so it had the very high automaticity. Secondly, the data of experiment was analyzed with error theory, the result shown that the stress-strain virtual measuring system had good dependability and high measure precision. Thirdly, it was one tenth of the traditional stress-strain measuring system in volume, and it need not network power supply, so it is very suitable for the field experiments. In the end, it has much higher ratio of performance to price.

5. ACKNOWLEDGMENTS

The project was supported by National Natural Science Foundation of China (Grant No. 50175045), the Science and Technology Development Project of Jilin Province（Grant No. 20050539）, the Doctor Subject Special Foundation of College (Grant No. 20050183015), and the 2004 Innovation Foundation of Jilin University.

6. REFERENCES

[1] D.G. Li, J.G. Cao et al, “Measurement System of Quality Characteristic of Flight Object Based on Virtual Instrument Platform,” Journal of Scientific Instrument, 2003，24（1）：53-56.

[2] Q. Li, Z.Z. Chen, “Design of Hardware of Signal Processing Virtual Instrument System,” Journal of Scientific Instrument, 2004，25（3）： 413-417.

[3] A.P. Singh, T. S. Kamal et al, “Development of a Virtual Curve Tracer for Estimation of Transducer Characteristics Under the Influence of a Disturbing Variable,” Sensors and Actuators A, 2005，Vol. 120, 518-526.

[4] S. D. Bencomo, “Control Learning: Present and Future,” Annual Reviews in Control, 2004，Vol. 28, 115-136.

[5] M.P. Rigobello, “Virtual Instrumentation for pH Measurements in Biological Systems,” Computer Methods and Programs in Biomedicine, 1999 ，Vol. 60, 55-64.

[6] LayerE, GawedzkiW, “Theoretical Principles for Dynamic Errors Measurement,” Measurement, 1990，8（1）： 178-182.

[7] H. L. Han, X.D.Xu et al, “The Design of Software Platform for Virtual Instrument,” Journal of test and measurement technology, 2002， 16（1）： 1-3.

[8] J.Q. Li, L.Q. Ren et al, “A Study on the Bionic Plow Moldboard of Reducing Soil Adhesion and Plowing Resistance,” Transactions of the Chinese Society of Agricultural Machinery, 1996，27（2）： 1-4.

[9] L.Q. Ren, J.Q. Li et al, “Unsmoothed Surface on Reducing Resistance by Bionics,” Chinese Science Bulletin, 1995， 40（13）： 1077-1080.

[10] Z.J. Zhang, “Research on Error Analysis in Dynamic Measurement,” Journal of Scientific Instrument, 2004，Vol. 18（2）： 139-143.

[11] W.F. Ge, “The Hardware of Anti-interference Measures of PLC Control System,” Coal Mine machinery, 2004，No. 11, 57-59.

[12] StanislawLR, “Towards a Global Virtual Instrument Control and a Virtual Link between Experiment and Modeling,” ThermochimicaActa, 2000 ，No. 355, 107-111.

[13] Quatech. Inc, SignalPro Series Signal Conditioning Adapters Users Manual, 2004.

[14] Datalog and Dasytec, DasyLab Data Acquisition System Laboratory User’s Guide Version 5.6,2004.

[15] G.C. Li, Y.S. Yang et al, “Measuring Error and Data Processing Principle”, Edition II, Changchun: Jilin University Press, 1990, 107-109.

Structure Synthesis of 3r2t 5-Dof Symmetrical Parallel Robot Mechanisms

HaiRong Fang , YueFa Fang , ZhunQing Hu , Shuo Wang
School of Mechanical and Electronic Control Engineering, Beijing Jiaotong University
Beijing 100044, China
Email: huzhunqing@163.com

ABSTRACT

Using screw theory, motion properties and constraint conditions for 5-DOF parallel manipulators structural platform is analyzed, and the relationship between screw and reciprocal screw of structural platform are established. All possible limbs of force constraints are derived, and structure synthesis for low -DOF parallel robot mechanisms is advanced. Finally, a parallel manipulator with 3-rotation-DOF and 2-translation-DOF is used as an example to describe the synthesis procedure for symmetrical and non-symmetrical 5-DOF parallel manipulators.

Keywords: Screw theory, Robot, Parallel mechanisms Structure synthesis.

1. INTRODUCTION

As a low-DOF parallel manipulator is a complex mechanical constraint system, the type design of it is a complex design process. In fact, such design is considered as one of the open problems in parallel robot research by Merlet[1]. So far, the methods of structural synthesis of low-DOF parallel manipulators include screw theory method [2], Lie group theory method, put forward by Herve et al, and single open chain method presented by T. L. Yang et al[3]. The researches of low-DOF parallel manipulators are mainly focused on 3-DOF parallel manipulators.

In this paper, the constraint systems in low-DOF parallel manipulators are analyzed using the screw theory [4]. A systematic approach is presented for structural synthesis of low-DOF parallel manipulators. A parallel manipulator with 3-rotation-DOF and 2-translation-DOF is used as an example to demonstrate the feasibility of the approach.

2. APPLICATION OF SCREW THEORY IN STRUCTURE SYNTHESIS OF PARALLEL MECHANISMS

2.1 Screw and reciprocal screw

In the screw theory, screw $\$$ is denoted by a dual vector $(\boldsymbol{s};\boldsymbol{s}^0)$, or in the Plüker coordinate form $(l \quad m \quad n; p \quad q \quad r)$. If $\boldsymbol{s}\cdot\boldsymbol{s}=\mathbf{1}$, the screw is called the unit screw. The pitch of the screw h is defined by

$$h=\frac{\boldsymbol{s}\cdot\boldsymbol{s}^0}{\boldsymbol{s}\cdot\boldsymbol{s}} \qquad (1)$$

If $h=0$, namely, $\boldsymbol{s}\cdot\boldsymbol{s}^0=0$, the screw , $\$$ is called a line vector, denoted by $(\boldsymbol{s};\boldsymbol{s}_0)$, where $\boldsymbol{s}$ is a vector, $\boldsymbol{s}_0=\boldsymbol{r}\times\boldsymbol{s}$ defines the moment of the screw axis about the origin of a reference frame, $\boldsymbol{r}$ is the position vector of any point on the vector $\boldsymbol{s}$ with respect to the reference frame. A line vector represents the direction and location of a line in space, and represents a rotational velocity in kinematics or a force in static. If h is equal to infinite, namely, $(0;\boldsymbol{s})$. An infinite pitch screw can represents a linear velocity in kinematics or a couple in static.

If two screws, $\$$ and $\$_r$, satisfy the condition

$$\$_r \circ \$ = 0 \qquad (2)$$

They are said to be reciprocal, where the symbol "$\circ$" denotes the reciprocal product.

If the screw $\$$ represents an instantaneous motion of a rigid body, the first three components of it represent the angular velocity and the last three components represent the linear velocity of a point in the rigid body that is instantaneously coincident with the origin of a reference frame, and we call it a twist. In comparison, the reciprocal screw $\$_r$ denotes a system of forces and couples acting on a rigid body, then the first three components of it represent the resultant force and the last three components represent the resultant moment about the origin of a reference frame, and we call it a wrench. A wrench represents a single force if the pitch of it is equal to zero and a single force couple if the pitch of it is infinite.

A parallel manipulator typically consists of a moving platform that is connected to a fixed base by several serial limbs, all the joints in which can be denoted by screws, with the revolute joints represented by the screws of zero pitch, the prismatic joints by the screws of infinite pitch, and the screw joints by the screw of finite pitch. Other types of joint can be formed by combining these basic joints. For example, a spherical joint is equivalent to three intersecting but not coplanar revolute joints, a cylindrical joint to one prismatic joint in concentric with a revolute joint, and universal joint to two intersecting revolute joints. With the concept of joint substitution applied, all the joints of a limb can form a twist system. If there exists a reciprocal screw which is reciprocal to every joint screw in the twist system, the reciprocal screw represents the constraint for this limb acting on the moving platform, and the number of the reciprocal screws is equal to the number of the limb constraints.

2.2 Structure synthesis of low-DOF parallel robot mechanisms

The first step is to define constraint wrenches of a moving platform. To limit the DOF of the moving platform of a parallel robots mechanism, there must be at most six linearly independent constraint wrench, denoted by six unit basic screws as

* This project is supported by National Natural Science Foundation of China (No.50375001). Received June 30，2003; received in revised form May 28，2004; accepted June 5，2004

$$\begin{cases}\hat{\$}_r^1=(1\ \ 0\ \ 0;0\ \ 0\ \ 0),\hat{\$}_r^2=(0\ \ 1\ \ 0;0\ \ 0\ \ 0)\\ \hat{\$}_r^3=(0\ \ 0\ \ 1;0\ \ 0\ \ 0),\hat{\$}_r^4=(0\ \ 0\ \ 0;1\ \ 0\ \ 0)\\ \hat{\$}_r^5=(0\ \ 0\ \ 0;0\ \ 1\ \ 0),\hat{\$}_r^6=(0\ \ 0\ \ 0;0\ \ 0\ \ 1)\end{cases} \tag{3}$$

where $\hat{\$}_r^1,\hat{\$}_r^2,\hat{\$}_r^3$ —— Constraint forces acting on the moving platform along the moving directions of three axes

$\hat{\$}_r^4,\hat{\$}_r^5,\hat{\$}_r^6$ ——Constraint couples acting on the moving platform along three axes

When the degree of freedom of a mechanism is equal to m, there must be 6-m linearly independent constraint wrenches acting on the moving platform, which form the constraint wrench system of the moving platform.

Step two is to solve the twist system reciprocal to the constraint wrenches of the moving platform. Without loss of generality, for n-DOF serial robot limb, it can be assumed that all joints of a limb are single-DOF joints. If there is a multi-DOF joint, it can be substituted with several single-DOF joints. The axis of the ith joint is denoted by unit screw $\hat{\$}_i$. Then the joint screws of a limb form a limb twist system of order n, denoted by

$$T_L=\left[\hat{\$}_1,\hat{\$}_2,\cdots,\hat{\$}_n\right] \tag{4}$$

where n —— Number of 1-DOF joints in a limb

For spatial manipulators, if n = 6, there exists no screw that is reciprocal to the limb twist system. If n < 6, there exist 6-n linearly independent reciprocal screws that form a constraint wrench system, which represents constrained condition of the limb to the moving platform, shown as

$$\$_r^j\circ\hat{\$}_i=0\qquad i=1,2,\cdots,n;\ j=1,\cdots,6-n \tag{5}$$

where $\hat{\$}_i$ —— Unit twist of the ith joint

$\$_r^j$ —— jth reciprocal screw in the limb constraint wrench system

Step three is to determine the limb structures. This is a very important step in the process of structure synthesis of parallel robot mechanisms. In this step, we must not only find out all types of limbs that satisfy the motion characteristics, but also take into account the rationality of the limb configuration, complexity of the mechanism, continuity of the motion, singularity of the mechanism and so on.

3. STRUCTURE SYNTHESIS OF LIMBS OF FORCE CONSTRAINTS

3.1 Twist system of limbs of force constraints

To solve the twist of a limb, two special cases are investigated here.

Case 1. A limb is called the F-1 limb if there is only one constrained wrench reciprocal to the limb twist system. As this limb can only supply one constraint, the limb has five degrees of freedom. Case 2. A limb is called the F-2 limb if there is two constraint wrenches reciprocal to the limb twist system.Fig.1 shows the F-1 limb.To make the twist system more general, the constraint wrench is denoted by a unit wrench of zero pitch as

$$\hat{\$}_r=(\boldsymbol{s}_r;\boldsymbol{r}_r\times\boldsymbol{s}_r) \tag{6}$$

where $\boldsymbol{s}_r=\begin{bmatrix}l_r & m_r & 0\end{bmatrix}$—— Unit vector along the direction of the constraint force

$\boldsymbol{r}_r=\begin{bmatrix}x & y & z\end{bmatrix}$ —— Position vector pointing to any point on the constrained wrench axis

According to Eq. (5), five basic screws of the limb twist system which are reciprocal to the constrained wrenches denoted by Eq.(6) are solved as

$$\begin{cases}\$_1=(1\ \ 0\ \ 0;\ m_rz/l_r\ \ 0\ \ 0)\\ \$_2=(0\ \ 1\ \ 0;\ -z\ \ 0\ \ 0)\\ \$_3=(0\ \ 0\ \ 1;\ (l_ry-m_rx)/l_r\ \ 0\ \ 0)\\ \$_4=(0\ \ 0\ \ 0;\ -m_r/l_r\ \ 1\ \ 0)\\ \$_5=(0\ \ 0\ \ 0;\ 0\ \ 0\ \ 1)\end{cases} \tag{7}$$

Any twist in the limb twist system is obtained by a linear combination of the above five basis screws, namely,

$$\$=a\$_1+b\$_2+c\$_3+d\$_4+e\$_5=(a\quad b\quad c;\quad f/l_r\quad d\quad e) \tag{8}$$

where *a*, *b*, *c*, *d*, *e*——arbitrarily constants that can not be simultaneously equal to zero

$$f=azm_r-bzl_r+cyl_r-cxm_r-dm_r$$

From the above twist system, the limb configurations can be determined. According to the above assumption, a limb is formed by the basic joints such as revolute joints(R), prismatic joints (P) and so on. The twist expressions of these two joints are derived by Eq.(9).

Screw expression of revolute joint R: according to Eq.(1),when applying the conditions $\boldsymbol{s}\cdot\boldsymbol{s}_0=0$ and $\boldsymbol{s}\cdot\boldsymbol{s}=1$, the unit twist of the revolute joint is derived by Eq.(9) as

$$\hat{\$}=(\boldsymbol{s};\boldsymbol{s}_0)=(a\quad b\quad c;\frac{f}{l_r}\quad d\quad -\frac{af+bdl_r}{cl_r}) \tag{9}$$

where $a^2+b^2+c^2=1$.

According to Eq. (9), if the values of a,b,c are defined, the direction of the joint, $\boldsymbol{s}=\begin{bmatrix}a & b & c\end{bmatrix}$, is determined uniquely, while the axis position of the joint is determined by single parameter d. Without loss of generality, letting $\boldsymbol{r}_r=\begin{bmatrix}0 & 0 & 0\end{bmatrix}$, then $f=0$, and

$$\boldsymbol{s}_0=\boldsymbol{r}\times\boldsymbol{s}=-\frac{d}{cl_r}(\boldsymbol{s}_r\times\boldsymbol{s}) \tag{10}$$

From the above equation, if $\boldsymbol{s}\neq\boldsymbol{s}_r$,the twists of the corresponding revolute joints on the limb F-1 intersect the given wrench. On the other hand, if $\boldsymbol{s}=\boldsymbol{s}_r$, the twists of the corresponding revolute joints on the limb are parallel with the given constraint wenches, and the position can be defined arbitrarily.

Screw expression of prismatic joint P: when applying the conditions $a=b=c=0$, the unit twist of the prismatic joint of infinite pitch is derived by Eq. (9) as

$$\hat{\$}=(0;\boldsymbol{s})=(0\quad 0\quad 0;f/l_r\quad d\quad e) \tag{11}$$

Where $(dm_r/l_r)^2+d^2+e^2=1$. It can be observed that $\boldsymbol{s}^{\mathrm{T}}\boldsymbol{s}_r=0$, and so, all prismatic joints in the F-1 limb are perpendicular to the given constrained wrench.

In summary, to construct the F-1 limb, the revolute joints and the prismatic joints must satisfy the following conditions: ①the revolute joints in the F-1 limb must intersect or parallel

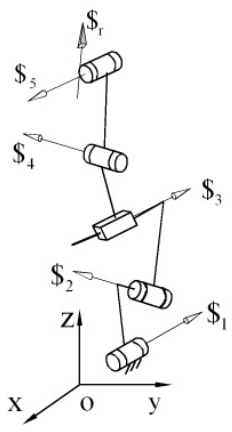


Fig. 1. Series limb

to the given constraint force. ② all prismatic joints in the Fig. 1. Limb must be perpendicular to the given constraint force.

3.2 Limbs of force constraints types

In view of condition 1, there are two types of the revolute joints in the Fig. 1. Limb: those revolute joints that are parallel to the direction of the constraint forces, called type R-1, and all other revolute joints that intersect the constraint force, called type R-2. Further, to uniquely define the location of the constraint force and to avoid redundancy, there must be at least two and at most three type R-1 joints. At the same time, the number of the corresponding linearly independent twists of zero pitch is at most five.

In view of conditions 2, there are at most two linearly independent prismatic joints in the Fig. 1. limb. Using revolute (R) and prismatic (P) joints as the basic joint types, we obtain 5R, 4R1P, 3R2P type Fig. 1. limbs. Here a numeral appearing in front of the joint signals denotes the number of the corresponding joints in the limb. With the

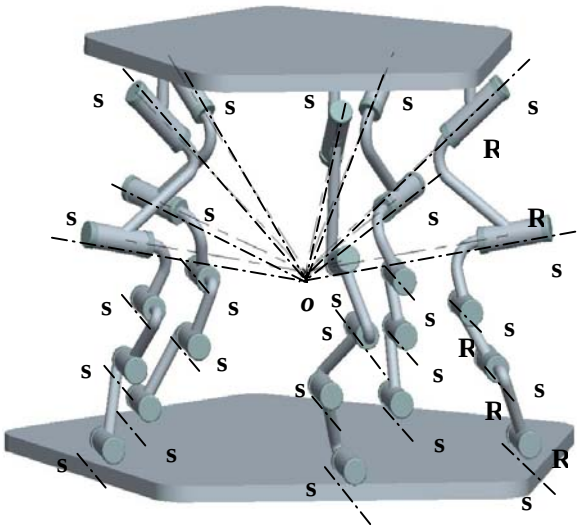


Fig. 2. 5−$R_2R_2R_2R_1R_1$ 3R-2Tsymmetrical parallel structure

concept of joint substitution, 3R1U, 2R1U1P, 3R1C, 2R1C1P and 1R1U2P as the feasible 3-link (4-jointed) limb configurations and 2R1S, 2U1R, 1R1S1P, 1R1U1C, 1S2P, 1R2C and 1U1C1P as the feasible 3-link (4-jointed) limb configurations can be derived.

Using the same method as described in case 1, 4R and 3R1P type Fig. 2. limbs are obtained. With the concept of joint substitution, 2R1U, 2R1C and 1R1U1P as the feasible 2-link (3-jointed) limb configurations can be derived. These Fig. 2 limbs can supply two pure forces instantaneously, but cannot satisfy constraint needs in the whole motion cycle.

4. Example of structure synthesis of 3R2T 5-DOF parallel manipulators

In a 3R2T 5-DOF parallel symmetrical manipulator, motion platform needs 3-rotation-DOF and 2-translation-DOF and single force constraints according to the request of structure symmetry, for a 5-DOF parallel manipulator, there must be five limbs, each limb supporting an force constraints. Therefore, each motion platform force constraints wrench system composes five constrains wrench

To make the twist system more general, if the force constraints wrench is a single force constraint, it will be expressed with a zero pitch wrench. Therefore, the motion platform constraints wrench $\mathbf{W}_P$ expresses as follow matrix

$$\mathbf{W}_P = \begin{bmatrix} \mathbf{s}_{r1} & \mathbf{r}_{r1} \times \mathbf{s}_{r1} \\ \mathbf{s}_{r2} & \mathbf{r}_{r2} \times \mathbf{s}_{r2} \\ \mathbf{s}_{r3} & \mathbf{r}_{r3} \times \mathbf{s}_{r3} \\ \mathbf{s}_{r4} & \mathbf{r}_{r4} \times \mathbf{s}_{r4} \\ \mathbf{s}_{r5} & \mathbf{r}_{r5} \times \mathbf{s}_{r5} \end{bmatrix} \tag{12}$$

In order to make the five force constrains of motion platform only support a real force constraints effect, the five limb might provide five constraints wrench the same direction.named $\mathbf{s}_{r1} = \mathbf{s}_{r2} = \mathbf{s}_{r3} = \mathbf{s}_{r4} = \mathbf{s}_{r5}$. In order to avoid the five force constraints wrench produce couples constraints. All of them need to make one point. That is $\mathbf{r}_{r1} = \mathbf{r}_{r2} = \mathbf{r}_{r3} = \mathbf{r}_{r4} = \mathbf{r}_{r5}$,where $r(\mathbf{W}_P) = 1$.

Thus it can be seen that when structure synthesis for 3R-2T parallel symmetrical manipulators is progressed, in order to ensure whole limbs in workspace satisfy above conditions, it is possible for F-type limb to have 4 poles and 5 pairs.

According to above and structure symmetrical request, when structure synthesis for five sameness limbs chosen from F-1 type limb is progressed, it will be obtained 5-R1R1R2R2R2 、 5-R1R1R2PR2 、 5-R1R1PR2R 、 5-R1R1PPR2 、5-R1R2R2R2R1、5-R1PR2PR1 parallel manipulator structure. Fig.2 gives 5-R2R2R2R1R1 3R-2T 5-DOF symmetrical parallel manipulator structure.

In order to satisfy motion platform structure 3R-2T, R2R2R2R1R1 in Fig. 2 arranges as follow

Rotate pair axes R_2 $\mathbf{s}_{i1} // \mathbf{s}_{i2} // \mathbf{s}_{i3} // \mathbf{s}_{r1} (i = 1 \sim 5)$, all of limb rotate pairs axes R_1 $\mathbf{s}_{i4}$ and $\mathbf{s}_{i5}$ must be intersect to point O, and in order to make structure all motion limbs axes R_2 the same direction, its axes must be sequence arranged from besed platform.

5. CONCLUSIONS

In this paper, the explicit screw expression of the degrees of freedom and the constraints of the moving platform in low-DOF parallel manipulators are presented. The relation of the constraint screw system and the twist system is brought forward, and the steps of structure synthesis of low-DOF parallel robots mechanisms are defined. Also put forward is a systematic method for structure synthesis of low-DOF parallel robot mechanisms. As an example, the 5-DOF parallel manipulators with three rotation and two translation moving platform are synthesized. This method has important guidance action to low-DOF parallel manipulator structure synthesis.

6. REFERENCES

[1] K.H. Hunt, “Kinematic Geometry of Mechanisms”, Oxford University Press, 1978.

[2] J.M. Herve, F. Sparacino, Structural Synthesis of parallel Robots Genrating Spatial Translation.
[3] Y.F. Fang, L.W. Tsai, “Structure Synthesis of a Class of 4-DoF and 5_DoF Parallel Manipulators with Identical Limb Structures”. The International Journal of Robotics Research, 2002,21(9):799-810.
[4] D. Zlatanov, C.M. Gosselin, “A Family of New Parallel Architectures with Four Degrees of Freedom, Journal of Compulational Kinematics”, F.C.Park and C.C.Iurascu:57-66.

Property Identification of the Singularity Loci of the 6/6-SPS Stewart Manipulator by Using the Method of *Z*-Plane

Yi Cao, Weixi Ji, Hui Zhou
School of Mechanical Engineering, Southern Yangtze University
Wuxi, Jiangsu 214122, China
Email: caoyi@ysu.edu.cn
And
Baokun Li, Xuanyao Wang
Department of Mechanical Engineering, Anhui University of Science and Technology
Huainan, Anhui 232001, China
Email: li_bao_kun@126.com

ABSTRACT

The problem of identifying the property of singularity loci of 6/6-SPS Stewart manipulators is addressed. After constructing the Jacobian matrix of the 6/6-SPS Stewart manipulator, a cubic polynomial expression in the mobile platform position parameters, which represents the constant -orientation singularity locus of the manipulator, is derived. Graphical representations of the singularity locus of the manipulator for different orientations are illustrated with examples. Based on this analytical expression, a quadratic expression that represents the singularity locus of the manipulator in *Z*-plane is derived, and further the property of singularity loci of the manipulator in parallel *Z*-planes is identified. It shows that singularity loci of 6/6-SPS Stewart manipulators in parallel *Z*-planes are all quadratic expressions including infinite hyperbolas, four pairs of intersecting lines and a parabola.

Keywords: Stewart Manipulator, Singularity Locus, Property Identification, Singularity Classification, *Z*-Plane.

1. INTRODUCTION

During the past two decades, parallel manipulator systems have become one of the research attentions in robotics. This popularity has been motivated by the fact that parallel manipulators possess some specific advantages over serial manipulators. Among them, the best-known parallel manipulator is the Gough-Stewart Platform that was introduced as an aircraft simulator by Stewart [1] in 1965.

One of the important problems in robot kinematics is the special configuration. As to parallel manipulators, in such configurations, the end-effector gains at least one unwanted degree of freedom while all actuators are locked. Meanwhile, infinite active forces must be applied to balance the loads exerted on the platform, otherwise causes the mechanism breakdown. This phenomenon has attracted the attention of many researchers. Hunt [2] first discovered a special configuration for the parallel manipulator that occurs when all the segments associated with the prismatic actuators intersect a common line. Fichter [3] pointed out that a special configuration is attained when the moving platform rotates around *Z*-axis by $\psi = \pm\pi/2$, whatever the position of the moving platform is. Merlet [4, 5, 6] studied the singularity of the 3/6-SPS Stewart manipulator more systematically using the Grassmann Line Geometry, including 3c, 4b, 4d, 5b and 5a. Gosselin and Angeles [7] showed that singularities of parallel manipulators could be classified into three different types. Recently, St-Onge and Gosselin [8] pointed out that the singularity locus equation of the general 6/6-SPS Stewart manipulator should be a polynomial expression of degree three, and so on [9,10,11,12,13,14,15,16]. Most of the reported work in this area has been performed in geometric conditions of singular configurations, deriving the analytical symbolic expression and obtaining graphical representations of the singularity locus. However, to the best of the authors' knowledge, much less work on singularity has been reported for the topic of property identifications of the singularity loci themselves, which is of great importance to the understanding of the singularity, the most relevant investigations have been maid by Huang et al. [13, 14] for the property identification of singularity loci of the 3/6-SPS Stewart manipulator.

This paper will further analyze the property of singularity loci of the 6/6-SPS Stewart manipulator. It is the most general form of 6/6-SPS Gough-Stewart manipulators and widely used in practice. Its mobile and base platforms are two similar semiregular hexagons in shape but opposite in direction in the initial pose. Our study is briefly as follows. After constructing the Jacobian matrix of the manipulator according to the theory of statical equilibrium, a cubic analytical expression that represents the singularity locus of the manipulator for a constant orientation is derived. Singularity loci of the manipulator for different orientations are illustrated with examples to demonstrate the result. Based on this polynomial expression, a quadratic expression that represents the singularity locus of the manipulator in *Z*-plane is derived, and further the property of singularity loci of the manipulator in parallel *Z*-planes is analyzed. It is shown that singular loci of the 6/6-SPS Stewart manipulator in parallel *Z*-planes include four pairs of intersecting lines, a parabola, and infinite hyperbolas.

2. SINGULARITY LOCI ANALYSIS IN THREE-DIMENSIONAL SPACE

A moving reference frame $P\text{-}X'Y'Z'$ and a fixed one *O-XYZ* are respectively attached to the mobile platform, B_1B_2, … , B_5B_6, and the base one, C_1C_2, …, C_5C_6, of the manipulator, as shown in Fig. 1., where origins *P* and *O* are corresponding geometric center of the mobile platform and the base one. The Cartesian coordinates of the mobile platform are given by the position of point *P* with respect to the fixed frame, designated by (*X*, *Y*, *Z*), and the orientation of the mobile platform is represented by three standard *Z-Y-Z* Euler angles (ϕ, θ, ψ).

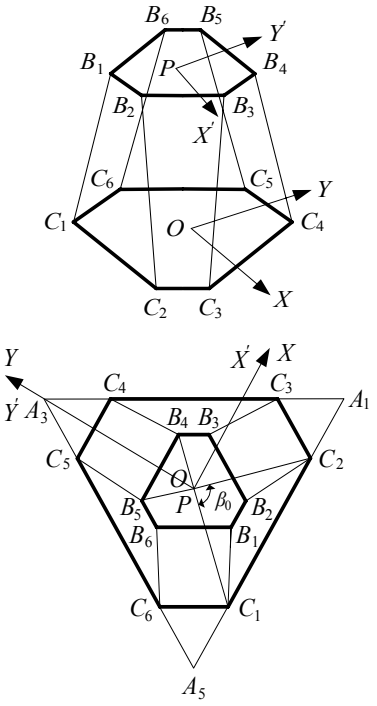


Fig. 1. A schematic of a 6/6-SPS Stewart manipulator

Furthermore, design parameters of the 6/6-SPS Stewart manipulator can be described as follows. The circumcircle radius of the base hexagon is R_a, and the one of the mobile hexagon is R_b. While β_0 denotes the central angle of circumcircles of the hexagons corresponding to sides C_1C_2 and B_4B_5, as shown in Fig. 1.. The coordinates of six vertices, B_i (i=1, 2,…, 6), of the mobile platform are denoted by $\boldsymbol{B}_i'$ with respect to the moving frame, and $\boldsymbol{B}_i$ with respect to the fixed frame. Similarly, $\boldsymbol{C}_i$ and $\boldsymbol{A}_j$ represent the coordinates of nine vertices, C_i (i=1, 2,…, 6) and A_j (j = 1, 3, 5), of the base platform with respect to the fixed frame.

According to the theory of statical equilibrium, the Jacobian matrix $\boldsymbol{J}$ of the 6/6-SPS Stewart manipulator can be constructed as follows [12]

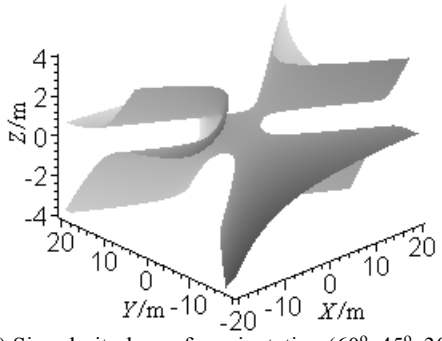
$$[J]^{T}=\begin{pmatrix} \frac{(\boldsymbol{B}_1-\boldsymbol{C}_1)}{|\boldsymbol{B}_1-\boldsymbol{C}_1|} & \frac{(\boldsymbol{B}_2-\boldsymbol{C}_2)}{|\boldsymbol{B}_2-\boldsymbol{C}_2|} & \frac{(\boldsymbol{B}_3-\boldsymbol{C}_3)}{|\boldsymbol{B}_3-\boldsymbol{C}_3|} & \frac{(\boldsymbol{B}_4-\boldsymbol{C}_4)}{|\boldsymbol{B}_4-\boldsymbol{C}_4|} & \frac{(\boldsymbol{B}_5-\boldsymbol{C}_5)}{|\boldsymbol{B}_5-\boldsymbol{C}_5|} & \frac{(\boldsymbol{B}_6-\boldsymbol{C}_6)}{|\boldsymbol{B}_6-\boldsymbol{C}_6|} \\ \frac{(\boldsymbol{C}_1\times\boldsymbol{B}_1)}{|\boldsymbol{B}_1-\boldsymbol{C}_1|} & \frac{(\boldsymbol{C}_2\times\boldsymbol{B}_2)}{|\boldsymbol{B}_2-\boldsymbol{C}_2|} & \frac{(\boldsymbol{C}_3\times\boldsymbol{B}_3)}{|\boldsymbol{B}_3-\boldsymbol{C}_3|} & \frac{(\boldsymbol{C}_4\times\boldsymbol{B}_4)}{|\boldsymbol{B}_4-\boldsymbol{C}_4|} & \frac{(\boldsymbol{C}_5\times\boldsymbol{B}_5)}{|\boldsymbol{B}_5-\boldsymbol{C}_5|} & \frac{(\boldsymbol{C}_6\times\boldsymbol{B}_6)}{|\boldsymbol{B}_6-\boldsymbol{C}_6|} \end{pmatrix} \quad (1)$$

where $\boldsymbol{B}_i, \boldsymbol{C}_i$ (i = 1, 2, … , 6) are vertex vectors defined above, and symbol × denotes the cross-product of two vectors and $|\cdot|$ indicates the norm of a vector, and $[.]^T$ denotes the transpose of [.].

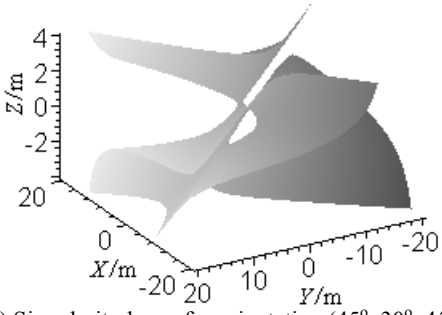
Recently, Gosselin and Angeles pointed out that singularities of parallel manipulators could be classified into three different types, i.e., inverse kinematic singularity, direct kinematic singularity and architecture singularity [7]. For the second type of singularity occurring when different branches of the direct kinematics problem converge, it is difficult to analyze and has received much attention from many researchers. So, this paper will only discuss the direct kinematic singularity of the 6/6-SPS Stewart manipulator considered in the present paper, which occurs when the determinant of the Jacobian matrix of the manipulator being equal to zero, i.e., det ($\boldsymbol{J}$)=det($\boldsymbol{J}^T$)=0. Expanding and factorizing the determinant of the Jacobian matrix $\boldsymbol{J}$, the singularity locus equation of the manipulator can be written as follows

$$\begin{aligned} &f_1Z^3 + f_2XZ^2 + f_3YZ^2 + f_4X^2Z + f_5Y^2Z + \\ &f_6XYZ + f_7Z^2 + f_8X^2 + f_9Y^2 + f_{10}XY + \\ &f_{11}XZ + f_{12}YZ + f_{13}Z + f_{14}X + f_{15}Y + f_{16} = 0 \end{aligned} \quad (2)$$

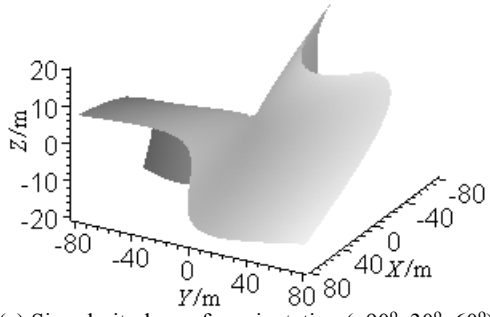
Eq. (2), a polynomial expression of degree three in the mobile platform position parameters XYZ and the maximum degree of variables X, Y is 2 and Z is 3, which represents the constant-orientation singularity locus of the manipulator in the Cartesian space for a constant orientation (ϕ, θ, ψ) of the mobile platform. Graphical representations of the constant-orientation singularity locus of the manipulator for different orientations are given to illustrate the theoretical result. Design parameters used here are given as R_a =2 m, R_b =3/2 m, $\beta_0 = \pi/2$.



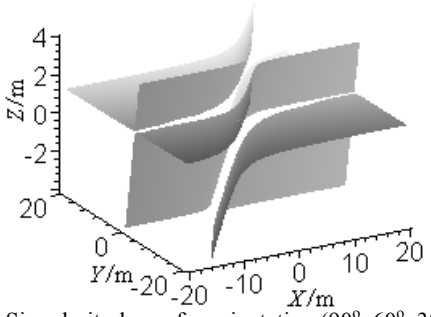
(a) Singularity locus for orientation (60°, 45°, 30°)



(b) Singularity locus for orientation (45°, 30°, 45°)



(c) Singularity locus for orientation (–90°, 30°, 60°)



(d) Singularity locus for orientation (90°, 60°, 30°)

Fig. 2. Singularity loci of the 6/6-SPS Stewart manipulator in three-dimensional space for different orientations

3. SINGULARITY LOCI ANALYSIS IN PARALLEL *Z*-PLANES

Huang et al. [14] pointed out that cross sections of the cubic singularity locus equation of the 3/6-Gough-Stewart manipulator in parallel θ-planes are all quadratic expressions that include a parabola, four pairs of intersecting lines and infinite hyperbolas. This conclusion is of great importance for the property identification of singularity loci of the 3/6-Gough-Stewart manipulator. One point to note is that the deriving process of the analytical expression of the singularity locus of 3/6-Gough-Stewart manipulators in θ-plane is very complex. However, in order to identify the property of singularity loci of the 6/6-SPS Stewart manipulator in three-dimensional space, a novel method called *Z*-plane is proposed by which singularity loci of the 6/6-SPS Stewart manipulator in parallel *Z*-planes will be discussed in this section. The advantage of *Z*-plane over θ-plane is that deriving process of the analytical expression of the singularity locus of 6/6-SPS Stewart manipulators in *Z*-plane is very simple.

Eq. (2) can be written as follows after some rearrangements and factorizations

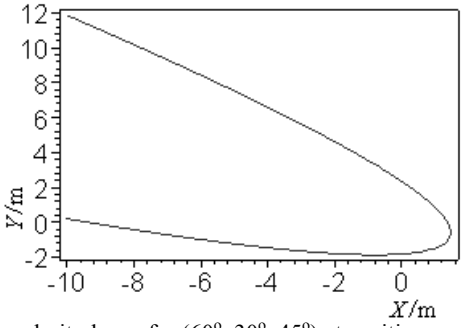
$$f(X, Y) = aX^2 + 2bXY + cY^2 + 2dX + 2eY + f = 0 \quad (3)$$

where $a = f_8 + f_4 Z$, $b = (f_{10} + f_6 Z)/2$, $c = (f_9 + f_5 Z)$, $d = (f_{14} + f_{11} Z + f_2 Z^2)/2$, $e = (f_{15} + f_{12} Z + f_3 Z^2)/2$, $f = (f_{16} + f_{13} Z + f_7 Z^2 + f_1 Z^3)$.

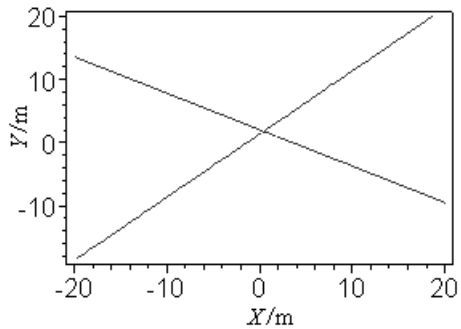
Eq. (3) is a quadratic polynomial expression with respect to *XY*. Furthermore, coefficients, *a*, *b*, *c*, *d*, *e*, *f*, are all functions of design parameters R_a, R_b, β_0 , Euler angles (ϕ, θ, ψ) and *Z*, which is a variable indicating the position of *Z*-plane. Therefore, for any given set of design parameters R_a, R_b, β_0 , Euler angles (ϕ, θ, ψ) and *Z*, singularity locus equation of 6/6-SPS Stewart manipulators in *Z*-plane is always a quadratic polynomial expression.

The property of the singularity locus of the 6/6-SPS Stewart manipulator with respect to *Z*-plane can be analyzed by two geometrical invariants, *D* and δ , of Eq. (3), $D = acf - ae^2 - b^2 f - d^2 c + 2bde$, $\delta = ac - b^2 = -\sin^4\theta (Z - k)^2$, where k is also a function of design parameters R_a, R_b, β_0 and Euler angles (ϕ, θ, ψ) . Further research shows that for any given set of design parameters R_a, R_b, β_0 and Euler angles (ϕ, θ, ψ) , *D* is a quartic expression while δ a quadratic expression with respect to the single variable *Z*. Generally, there are four real roots when *D*=0 and $\delta \neq 0$, and Eq. (3) degenerates into four pairs of intersecting lines, respectively. For the same reason, there is one real root k of multiplicity two when $\delta = 0$ and $D \neq 0$, and Eq. (3) degenerates into a parabola. With the exception of the above-mentioned five special values of *Z*, $D \neq 0$ and $\delta < 0$ for any value of *Z*, therefore Eq. (3) indicates a set of hyperbolas. So it can be concluded that when $\theta \neq 0^\circ$, singularity loci of the manipulator in parallel *Z*-planes are always quadratic expressions, which contain four pairs of intersecting lines, a parabola and infinite hyperbolas. This conclusion holds for the property of singularity loci of the 6/6-SPS Stewart manipulator in parallel θ-planes presented in [15].

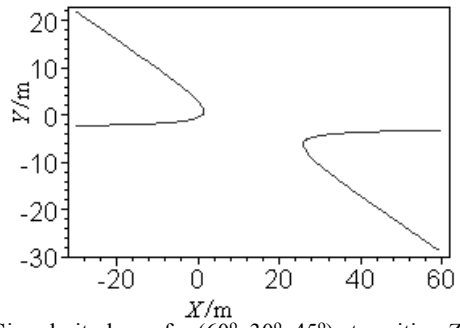
As mentioned above, for any given set of design parameters R_a, R_b, β_0 and Euler angles (ϕ, θ, ψ) , when $\theta \neq 0^\circ$, singularity loci of the 6/6-SPS Stewart manipulator in parallel *Z*-planes are always quadratic expressions including four pairs of intersecting lines, a parabola and infinite hyperbolas. In order to demonstrate the aforementioned theoretical result, a 6/6-SPS Stewart manipulator will be studied. Design parameters of the Stewart manipulator under investigation are given as $R_a = 2$ m, $R_b = 3/2$ m, $\beta_0 = \pi/2$, $(\phi, \theta, \psi) = (\pi/3, \pi/6, \pi/4)$. Further inspection shows that when $Z_1 = 3\sqrt{3}/8$, Eq. (3) degenerates into a parabola, as shown in Fig. 3(a)., and when *Z* is equal to the following values, $Z_2 = 0.23457$, $Z_3 = -0.84845$, $Z_4 = 0.98888$, $Z_5 = 1.35662$, Eq. (3) degenerates into four pairs of intersecting lines, respectively, as shown in Fig. 3(b)., with the exception of the above-mentioned five special values of *Z*, singularity loci of the 6/6-SPS Stewart manipulator in parallel *Z*-planes are always hyperbolas shown in Figs. 3(c). and 3(d)..



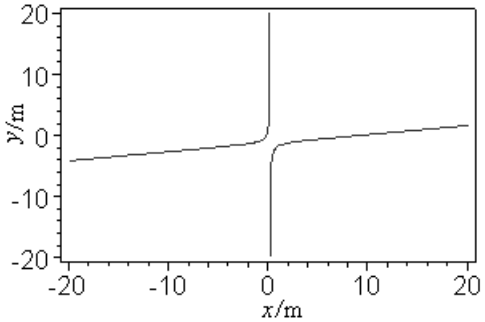
(a) Singularity locus for (60°, 30°, 45°) at position $Z_1 = 3\sqrt{3}/8$



(b) Singularity locus for (60°, 30°, 45°) at position Z_2=0.23457



(c) Singularity locus for (60°, 30°, 45°) at position Z= –1



(d) Singularity locus for (60°, 30°, 45°) at position Z= 0

Fig. 3. Singularity loci of the 6/6-SPS Stewart manipulator in parallel Z-planes for orientation (60°, 30°, 45°)

4. SINGULARITY ANALYSIS WHEN $\theta = 0^\circ$

It can be seen that when $\theta = 0^\circ$, the mobile platform is parallel to the base one. Meanwhile Eq. (2) can be reduces as follows after some rearrangements

$$Z^3 \cos(\phi + \psi) = 0 \tag{4}$$

1) $\theta = 0^\circ$ and $Z = 0$, the mobile platform and the base one are coincident. In this special configuration, the manipulator has three degrees of freedom: two rotational freedoms and one translational freedom. They belong to the first and the second special-linear-complex singularity [13].

2) $\theta = 0^\circ$ and $(\phi + \psi) = \pm\pi/2$, for the 6/6-SPS Stewart manipulator, it is the singularity proposed by Fichter [3].

5. PROPERTY IDENTIFICATION OF THE SINGULARITY LOCI

Based on the analyses described above, it can be safely concluded that singularity loci of the 6/6-SPS Stewart manipulator in parallel Z-planes contain infinite hyperbolas, four pairs of intersecting lines and a parabola when $\theta \neq 0^\circ$. From analytic geometry, there are five different kinds of quadric surfaces with hyperbola sections. They are hyperbolic cylinder, hyperbolic paraboloid, hyperboloid of one sheet, hyperboloid of two sheets and conic surface. But none of them can contain hyperbolas, a parabola and four pairs of intersecting lines, simultaneously. Therefore, the singularity locus equation of the manipulator in three-dimensional space is a special irresolvable polynomial expression of degree three whose cross sections in parallel Z-planes contain a parabola, four pairs of intersecting lines and infinite hyperbolas.

6. CONCLUSIONS

Based on the analyses described above, properties of singularity loci of the 6/6-SPS Stewart manipulator for all different orientations can be finally concluded as follows

1)The singularity locus equation of the manipulator in three-dimensional space is a special irresolvable polynomial expression of degree three, whose cross sections in parallel Z-planes contain a parabola, four pairs of intersecting lines and infinite hyperbolas.

2) Graphical representations of the singularity locus of the manipulator are quite complex and various for different orientations. The most complex graphic of the singularity locus is a trifoliate surface with two holes.

3) For the 6/6-SPS Stewart manipulator's singularity, it also has some special cases where six lines associated with six extensible links of the manipulator can intersect one common line, and even for the same orientation of the manipulator, there are two and more positions of the manipulator in which six lines of the manipulator all intersect one common line, simultaneously.

7. REFERENCES

[1] D. Stewart, "A Platform with Six Degrees of Freedom", Proc. of the Institution of Mechanical Engineers, London, UK, 1965, 371-378.

[2] K.H. Hunt, "Structural Kinematics of in-Parallel-Actuated-Robot-Arms", ASME Journal of Mechanical Transmissions Automation and Design, 1983, 105(4): 705-712.

[3] E. F. Fichter, "A Stewart Platform-Based Manipulator: General Theory and Practical Construction", The International Journal of Robotics Research, 1986, 5(2): 157-182.

[4] J. Merlet, "Parallel Manipulator Part 2: Singular Configurations and Grassmann Geometry", Technical Report, INRIA, Centre de Sophia Antipolis, Valbonne, France, 1988.

[5] J. Merlet, "Singular Configurations of Parallel Manipulators and Grassmann Geometry", The International Journal of Robotics Research, 1989, 8(5). 45-56.

[6] J. Merlet, "Singular Configuration and Direct Kinematics of a New Parallel Manipulator", IEEE International Conference on Robotics and Automation, Nice, France, 1992, 338-343.

[7] C.M. Gosselin, J. Angeles, "Singularity Analysis of Closed-Loop Kinematic Chains", IEEE Transactions on Robotics and Automation, 1990, 6(3): 281-290.

[8] B.M. St-Onge, C. M. Gosselin, "Singularity Analysis and Representation of the General Gough-Stewart Platform", The International Journal of Robotics Research, 2000, 19(3): 271-288.

[9] R.D. Gregorio, "Analytic Formulation of the 6-3 Fully-Parallel Manipulator's Singularity Determination", Robotica, 2001, 19.663-667.

[10] R.D. Gregorio, "Singularity-Locus Expression of a Class of Parallel Mechanisms", Robotica, 2002, 20.323-328.

[11] Z. Huang, L. F. Kong , et al, Theory of Parallel Robotic Mechanisms and Control, Beijing: China Machine Press, 1997.

[12] Z. Huang, Y. S. Zhao, et al., "Kinematics Principle and Geometrical Condition of General-Linear-Complex

Special Configuration of Parallel Manipulators", Mechanism and Machine Theory, 1999, 34(8).1171-1186.

[13] Z. Huang, L.H. Chen, et al, "The Singularity Principle and Property of Stewart Manipulator", Journal of Robotic Systems, 2003, 20(4), 163-176.

[14] Z. Huang, L.H. Chen, et al , "Singular Loci Analysis of 3/6-Stewart Manipulator by Singularity-Equivalent Mechanism", IEEE International Conference on Robotics and Automation, Taipei, Taiwan, 2003,1881-1886,

[15] Y. Cao, Z. Huang, "Property Identification of the Singularity Loci of the Stewart Manipulator", The Tenth IASTED on Robotics and Application 2004/RA 447-017, Honolulu, Hawaii, USA, 2004.

[16] V.K. Chan, E.-Uphoff Imme, "Investigation of the Deficiencies of Parallel Manipulators in Singular Configurations through the Jacobian Nullspace", IEEE International Conference on Robotics and Automation, Seoul, Korea, 2000.

Risk based Dynamic Temperature Controlling in Large Grain Depot using one-wire Bus Digital Sensor Network*

Bin Xu
College of Computer Science & Information Engineering,
Zhejiang Gongshang University, Hangzhou 310035, China
Email: xubin@mail.zjgsu.edu.cn

ABSTRACT

Temperature monitoring and controlling are essential to protect the grain safely in grain depot. However, it is very difficult to control the temperature in large grain depots due to the large volume of temperature sensors and various inspecting situations. One-wire bus digital temperature sensors such as DS18B20 can be a good choice to reduce the total deployment cost while keeps high availability and correctness. A risk based dynamic temperature monitoring and controlling protocol for one-wire bus digital sensor network is proposed in this paper to monitor and control the environment temperature dynamically. The result from the simulation and real practice in a large grain depot shows that the significant benefit of the supposed protocol in largely reducing communication volume and timely temperature data collection.

Keywords: Dynamic Temperature Monitoring and Controlling, Risk based Protocol, One-wire Bus Digital Sensor Network.

1. INTRODUCTION

Sensor networks have a variety of applications in habitat monitoring, environmental monitoring, condition based maintenance, seismic detection, military surveillance, and inventory tracking [9].

Among all the focus on sensor network, energy efficiency is a dominant consideration. Sensor nodes in the network only have a small and finite source of energy; therefore many solutions have been proposed to optimize energy usage [10, 11, 12].

Rahul Shah et al. [10] proposed using sub-optimal paths occasionally to increase the lifetime of the sensor network substantially. The protocol is also a destination initiated reactive protocol. Instead of maintaining one optimal path, a set of good paths are maintained and chosen by means of a probability, which depends on how low the energy, consumption of each path is. As a result, no single path could get its energy depleted because different paths are chosen at different times. This ensures the graceful degradation of the network in low-energy networks because energy is burnt more equally in all nodes.

The research by Wei Ye et al. [11] focuses on reducing the major sources of energy wastage, including Collisions, overhearing, control packet overhead and idle listening, while achieving good scalability and collision avoidance capability.

* This research work is partly supported by Zhejiang Gongshang University.

Unlike in traditional networks where all nodes require equal opportunity to transmit, all sensor nodes try to achieve a single common task. S-MAC proposed using three techniques to reduce energy consumption. Firstly the nodes go to sleep periodically so that they do not waste energy by listening to an empty channel or when a neighboring node is transmitting to another node. This helps in avoiding the overhearing problem too. Secondly, nearby nodes form virtual clusters to synchronize their wake-up and sleep periods to keep the control packet overhead of the network low. Finally, message passing is used to reduce the contention latency and control overhead.

Younggang Zhao et al. [12] proposed a mechanism, which collects a residual energy scan (eScan) of the network, so as to grasp an aggregated picture of the energy levels in the different regions of the sensor network. They also propose using incremental updates to scans so that when the state of a node changes, it does not have to send its entire scan again thereby saving energy. [13]

Temperature is an important factor in industry and temperature sensors are widely used in monitoring and control the environment temperature [2, 3, 4, 5, 6, 7, and 8]. In order to monitor the temperature of a large grain depots, thousands sensors will be used. Cheap and practicable sensors should be used while DS18X20 [1] may be a good candidate.

In this paper, a risk-based temperature monitoring protocol based on warning window mechanism [14] is proposed while DS18B20 [1] is chosen to be studied. The proposed protocol is evaluated in simulation and has been tried in a large grain depot. The feedback is significant due to the largely reduced communication and timely data collection.

1.1 Storage and its functions of DS18B20

Each DS18B20 [Fig. 1.] [1] Contains a unique 64–bit code stored in ROM. The least significant 8 bits of the ROM code contain the DS18B20's 1-Wire family code: 28h. The next 48 bits contain a unique serial number. The most significant 8 bits contain a cyclic redundancy check (CRC) byte that is calculated from the first 56 bits of the ROM code. The 64-bit ROM code and associated ROM functions [Fig. 2.] [1] allow the DS18B20 to operate as a 1-Wire device using the 1-wire bus protocol.

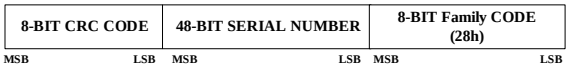


Fig. 1. 64-bit ROM of DS18B20 [1]

The rest of this paper is organized as follows. We introduce

the traditional temperature monitoring protocol in Section 2. We identify two factors impacting temperature monitoring in Section 3. The risk based dynamic temperature monitoring protocol is proposed in Section 4. With the historic temperature data, we compared the new protocol with the traditional one and the new protocol proved to be efficiently and adaptive with the circumstance. The conclusion is stated in Section 6.

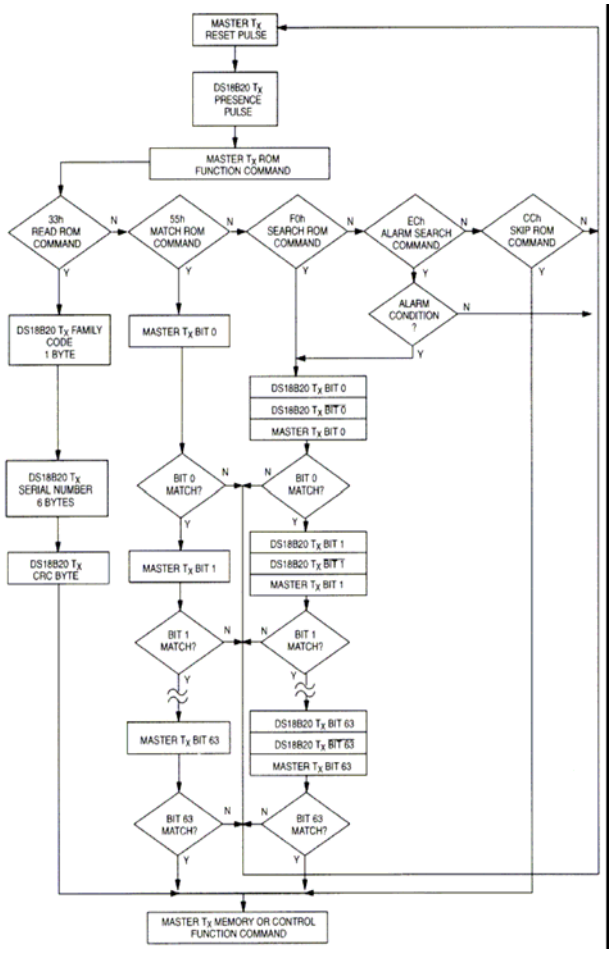


Fig. 2. ROM functions flow chart of DS18B20 [1]

2. TRADITIONAL TEMPERATURE MONITORING

When using traditional temperature monitoring protocol, all the sensors that connected to one single-chip machine will be called and visited in each cycle of temperature checking one by one.

After the single-chip machine has issued "Convert T" command, the sensors need around 750ms to make the temperature conversion. The above diagram enables each sensor connected to convert the temperature in parallel; therefore, no multiple 750ms will be wasted in conversion in such way. After the sensors have finished the temperature conversion, the temperature data will be stored in the scratchpads. The monitor transfers the temperature data per sensor. A cycle is finished when all sensors have been visited.

According to the traditional temperature data fetching protocol, the duration of a single cycle and the communication volume for each sensor can be,

$$\textbf{Duration=751.4ms+6.8ms*N} \quad (1)$$
$$\textbf{Communication per sensor = 96 bits}$$

Where N is the number of the sensors connected to one hypogenous monitor.

The cycle duration varies from different number of connected sensors. For instance, when there is 800 sensors connect to one monitor machine, the duration of one cycle will be around 6200ms. Even if there are only 100 sensors, a check cycle will last for more than 1400ms.

Therefore, it's impossible to get the timely environment temperature with traditional temperature monitoring protocol. The duration of one single temperature checking is long and some temperature change is ignored due to the large interval of a cycle.

According to Eq. (1), the temperature data transition and analysis will be time-consuming when there are large amount of monitoring sensors. In large grain depot, large amount of data is crowded in the 1-wire bus; as a result, the up-level monitoring machines are busy in getting the temperature data and doing involved analysis. This is the main reason why we conduct the research on dynamic temperature monitoring and risk based dynamic temperature controlling.

3. IDENTIFICATION OF TWO FACTORS IN TEMPERATURE MONITORING

We'd like to identify some important factors including interval and probability in temperature monitoring before modeling the dynamic temperature monitoring and risk based temperature controlling.

3.1 Interval of temperature monitoring

As shown in Fig. 3., the temperature varies from time to time during the monitoring. However, the curve is smooth in grain depot as the temperature of each monitor site is not change rapidly all the time. Moreover, when the curve is reviewed in a small interval, such as 1-2 minutes interval, it can be very flat. Smaller interval requires more monitor cycles while larger interval ignores some temperature change. Since there is a tradeoff should be made, the interval must be well determined in industry.

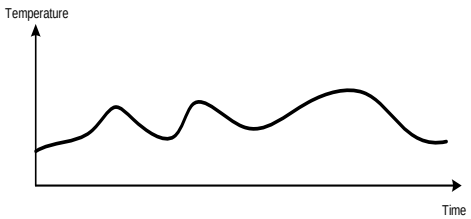


Fig. 3. Temperature varies from time to time

3.2 Probability of temperature change

Regarding to each monitor site, while the temperature remains the same most of the time, it may change from NºC to N+1ºC, N-1ºC, N+2ºC, N-2ºC or other degree after a certain period [Fig. 4.]. As shown in Fig. 4., if the temperature becomes large than T-Up, which is the top

bound, the temperature controlling system should make a serial of actions to reduce the temperature. On the other hand, if the temperature becomes less than T-Bottom, which it the bottom bound, the temperature controlling system also must make a serial of actions to increase the temperature.

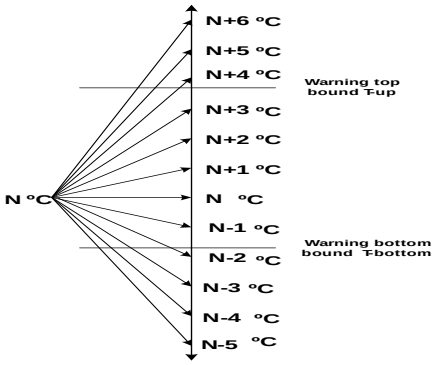


Fig. 4. Temperature changes within a certain period

Fig. 4. can also be expressed by Table 1., where there are three columns, namely source degree, target degree and the probability. The probability is how often the temperature will be changed from source degree to target degree. As the situation varies from different monitoring sites, the Location should be aware so as to handle with different warning area and different probability of temperature change. For instance, the temperature will be changed from NºC to $N\pm m$ºC at the probability of P (L, N, $N\pm m$) where L refers to different monitor sites.

Table 1. The temperature changes with probability at *l* site

From temperature (ºC)	To temperature (ºC)	Probability (%)
N	N	P(L, N, N)
N	N+1	P(L, N, N+1)
N	N+2	P(L, N, N+2)
N	N+3	P(L, N, N+3)
N	N+4	P(L, N, N+4)
N	N-1	P(L, N, N-1)
N	N-2	P(L, N, N-2)

4. RISK BASED DYNAMIC TEMPERATURE MONITORING AND CONTROLLING

4.1 Warning Window Mechanism

After the DS18B20 has performed a temperature conversion, the temperature value will be compared to the trigger values stored in TH and TL. An alarm flag inside the device will be set when the result of a temperature measurement is higher than TH or lower than TL. This flag is updated with every temperature measurement. As long as the alarm flag is set, the DS18B20 will respond to the alarm search command. This allows many DS18B20s to be connected in parallel doing simultaneous temperature measurements. If somewhere the temperature exceeds the limits, the alarming device(s) can be identified and read immediately without having to read non-alarming devices [1].

This mechanism has been introduced and used in one of the author's research paper [14]. In this paper, a risk-based temperature monitoring protocol is proposed on the basis of this warning window mechanism.

4.2 Risk based Temperature Monitoring Work Flow

Main workflow of risk based dynamic temperature monitoring can be stated as that in Fig. 5.

All the temperature data will be stored in the central control machine. When the system is started, these data will be fetched and be formatted to be analyzed. Probability generation will be made on the formatted historic temperature data. The probability data will be stored in the probability table. All these activities are suggested to be done by a central control machine with high performance in computing and disk operation. Meanwhile, the temperature sensors are initiated and deliver their initial temperature data to the hypogenous monitors. The central control machine will get all these data via the hypogenous monitors, and it will then determine the interval for all the sensors connected. The hypogenous monitors will store all the value of interval for the sensors connected to it. The initiate phase is finished after all these tasks have been done.

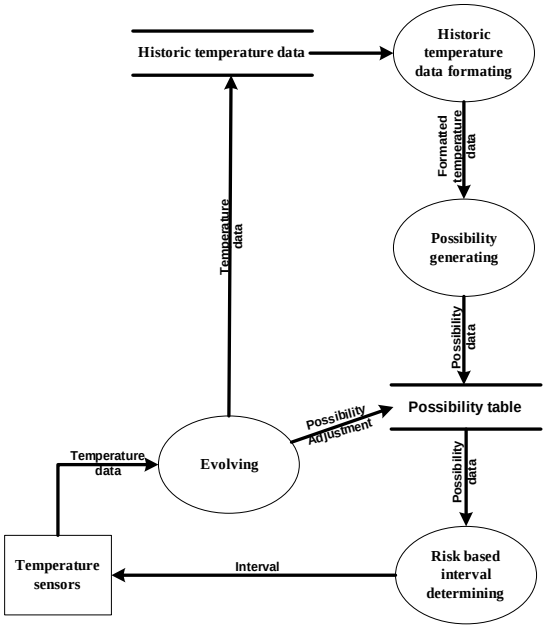


Fig. 5. Risk based temperature-monitoring workflow

During the working phase, the temperature sensors are visited only when the temperature is changed. The temperature data will then be delivered to the central control machine via the hypogenous monitor. If there is some adjustment to the probability table, updated possible data will be transferred to the hypogenous monitor.

4.3 Historic Temperature Data Formatting

The historic temperature data has been already recorded in files. In order to get the probability of the temperature change per sites, at least site, time and temperature should be captured from the historic data,

- Site and Time: where and when the temperature data was gotten.

- Temperature: The value of the temperature.

To save the time in analyzing, the data should be sorted in the order of site and time [Table 2.]. All the data will be analyzed per site and a remaining interval table can then be generated as shown in next part.

Table 2. Historic temperature data

Site	Time	Temperature (ºC)
Site 1	2006.2.1 10:02:0	12.1
	2006.2.1 10:22:0	13.6
	…	…
Site 2	2006.2.1 9:52:0	12.2
	…	…

4.4 Remaining Interval Table Generating

With the data from Table 2., the remaining interval of temperature at each value can be generated, as which is shown in Table 3.. In order to get the precise probability, the minimal, medium and max interval of temperature remaining at each temperature value should be calculated.

Table 3. Remaining interval analysis of the temperature

Temperature (ºC)	Minimal interval	Medium interval	Max interval
12.1	20 seconds	2 Minutes	1 hour
12.2	22 seconds	3 Minutes	2 hour
…	…	…	…
13.6	25 seconds	4 Minutes	3 hour

Since the temperature data is consecutive data, the temperature is assumed to be changed continuously. Supposing there are two adjoining records in Table 2. {site A, T_1, and V_1} and {site A, T_2, and V_2} for the same site A, the temperature is assumed to be remained the same for n cycles and then changed continuously from V_1 to V_2 in the last cycle [Fig. 6.]. Cycle duration is a normal duration for one turn when all the potential temperature data of every sensor have been visited. Though the temperature data is consecutive data, it is enough to have the precision to be 0.1 ºC.

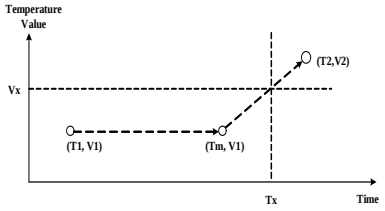


Fig. 6. Consecutive changing of temperature value in last monitoring cycle (a relative short duration)

One row in Table 2. can be removed if the temperature value is equal to that of its prepositive row so as to be more concise. As which is shown in Fig. 6., when the temperature value of one row equals to V_x ($V_1<V_x<V_2$ or $V_1>V_x>V_2$), the corresponding remaining interval $Interval(V_x)$ equals to,

$$Interval(Vx) = \frac{One\ Cycle\ Duration}{(V2 - V1) * \Pr ecision} \quad (2)$$

While $Interval(V1) = (Tm - T1) = T2 - T1 - One\ Cycle\ Duration$.

Giving the example shown in Table 2., under the condition that one cycle duration is 30 seconds, according to the two rows for Site 1, we know that the temperature of Site 1 remains 12.1 ºC from 2006.2.1 10:02:00 to 2006.2.1 10:21:30, and then increase to 13.6 ºC continuously in last 30 seconds. The remaining interval value will be updated to Table 3..

4.5 Risk based Interval Determining

PERT technique can be used to determine the monitor interval when the remaining interval table is ready. The expectation and standard deviation can be calculated as following,

- ***Expectation***

μ = *(minimal interval + 4 * medium interval + max interval) /6*. (3)

- ***Standard Deviation***

σ = *(max interval – minimal interval) /6* (4)

Such that, the final curve adheres to Gaussian distribution - $N(\mu, \sigma^2)$. The probability of the temperature remain the same for more than interval Ti is equal to,

$$P(t >= Ti) = 1 - \Phi(\frac{t-\mu}{\sigma}) \quad (5)$$

Eq. (3), (4) and (5) enable the industry to make tradeoff between precision in dimension of time and the performance of the system.

4.6 Evolving with the Temperature Data

In order to reflect the current situation of monitoring environment, the dynamic temperature data will be gathered and the remaining interval will be updated so that the interval decision will be adaptively suitable in monitoring.

The remaining interval analysis table can also be recorded in files or databases after it have been established. If so, the step of historic temperature data formatting and remaining interval table generating can be replaced with just fetching the stored table.

5. SIMULATION AND COMPARISON

5.1 Simulation

The historic temperature data was gathered and analyzed. We found that most of the time (around 85% of the time), most of the sensors (around 90% of the sensors) sensing the same temperature. However, those sensors near the windows, door or the aisle would sense dynamic temperature. A simulation was made according to such situation and the results showed that the communication volume was much less than the traditional one.

5.2 Comparison

There are thousands of sensors using in large grain depot, fortunately most of the sensors are working under stable situation where the temperature remain the same most of the time. Only those near the door, facing the windows, or close to the aisles may have shorter interval remaining the same temperature. The protocol presented in this paper enables the system to reduce monitoring times for most of the sensors.

Eq. (5) shows that the risk of potential temperature data ignoring or delay along with the increase of the monitor interval. It seems that the proposal brings in new risk in data precision, but such risk is also there for the traditional ones. Every time a sensor notices that there is a temperature change, we cannot guarantee that the change happened right before the sensing. Actually, we cannot ensure whether there is a lower or higher temperature had been lost before the sensing.

6. CONCLUSION AND DISCUSSION

The duration will last longer when there are more temperature sensors connecting to one monitor, which will then reduce the precision as some temperature change is ignored within the interval. Instead of using traditional monitoring protocol, we have proposed a new risk based temperature monitoring protocol, which adopts the warning window. The protocol have been described and evaluated formally. A simulation has been used to evaluate the benefit of new proposed protocol. The benefit is significant as the interval was reduced and the communication volume was reduced largely.

7. REFERENCES

[1] Semiconductor. DS18B20 programmable resolution 1-wire digital thermometer [Z]. Dallas Semiconductor Corporation, 2002, 1-26.

[2] L. Dubois, J. P. Sozanski, et al, “Temperature control and thermal dosimetry by microwave radiometry in hyperthermia,” IEEE Transactions on Microwave Theory and Techniques, Oct. 1996,44:1755–1761.

[3] L. Dubois, C. Vanoverschelde, et al, “Temperature control by microwave radiometry with narrow bandwidth,” European Journal of Physics D: Applied Physics, 2000,AP 9: 63–68.

[4] P. Kabacik, M.E. Bialkowski, “The temperature dependence of substrate parameters and their effect on microstrip antenna performance,” IEEE Transactions on Antennas and Propagation, June 1999,47:1042–1049.

[5] W.M. Zhao, Y.K. Zhang, et al, “The development and design of full digital temperature and humidity control of grain depot,” Journal of Process Automation Instrumentation, June 2001,22(6): 40-42.

[6] L.G. Chen, J.L. Liu, “Optimization of the addressing code of DS18X20 at measuring multiple temperatures,” Journal of Transducer Technology, 2001, 20(9): 55-57.

[7] L.G. Chen, “Optimization of Addressing Code of Digital Temperature Sensor Bass the 1-wire Bus at Measuring Multiple Temperatures,” Journal of Instrument Technique and Sensor, 2004, 1.41-43.

[8] L.G. Chen, “Addressing Code of DS18X20 at Measuring Multiple Temperatures and Online System Maintains,” Chinese Journal of Sensors and Actuators, 2005,18(1). 157-164.

[9] D. Estrin, R. Govindan, et al, “Next Century Challenges: Scalable Coordination in Sensor Networks,” In Proceedings of the Fifth Annual International Conference on Mobile Computing and Networks, August 1999, Seattle, Washington.

[10] R.C. Shah, J. Rabaey, “Energy Aware Routing for Low Energy Ad Hoc Sensor Networks,” IEEE Wireless Communications and Networking Conference, Orlando, FL, 2002,17-21

[11] W. Ye, J. Heidemann, et al, “An Energy-Efficient MAC Protocol for Wireless Sensor Networks,” In Proceedings of the 21st International Annual Joint Conference of the IEEE Computer and Communications Societies, New York, NY, USA, June, 2002.

[12] Y.J. Zhao, R. Govindan, et al, “Residual energy scan for monitoring sensor networks,” IEEE Wireless Communications and Networking Conference, 1: 17-21, 2002, 356 –362.

[13] M.Tubaishat, S. Madria, “Sensor networks: an overview,” IEEE Potentials, Apr-May 2003,22(2),.20 – 23.

[14] B.Xu, L.G. Chen, “Toward Efficient Temperature Monitoring and Controlling in Large Grain Depot,” To be appeared in Proceedings of International Workshop on Wireless Ad Hoc and Sensor Networks 2006, New York, June 2006.

On-line optimal control of switched reluctance motor based on the Least Square Method

Xilian Wang, Yihuang Zhang, Yuefeng Yang
Dept. of Electrical Engineering,Beijing Jiaotong University,
Beijing, 100044, China
Email: wangxl@dq.njtu.edu.cn
And
Xudong Wang
Dept. of Electrical Engineering, Harbin University of Science & Technology,
Harbin, 150040, China

ABSTRACT

On-line control schemes of angles are presented which optimize the steady-state performance of drive as measured by torque per ampere. A method of realization on-line control using EPROM is discussed. Based on the Least Square Method (LSM) a mathematic model of forecast current value after varied turn-on or turn-off angle is established. The flowchart of searching optimum turn-on or turn-off angle is detailed. Experimental results have shown the necessity and validity of the on-line control.

Keywords: SRM, EPROM, the Least Square Method, torque per ampere.

1. INTRODUCTION

Switched Reluctance Motor (SRM) has a wide application prospect in variable speed drive system, for it has many advantages, such as simplicity, firm, low cost, easy control, and high efficiency and so on. Its controllable parameters include phase winding current, turn-on angle θ_{on}, and turn-off angle θ_{off}. Commonly, SRM operates in current chopping mode at low speed and angle position control mode at high speed [1]. Traditional angle position control strategies usually calculate on-off angles off-line or linearly vary them along with variations of speed and load based on ideal design inductance mode. However, as the parameters of motor change, the model based on inductance curve may not precisely represent motor's electromagnetism state. In the off-line calculation some factors as magnetic saturation, iron losses, and leakage flux cannot be considered completely. Especially because of the asymmetry air gap caused by manual programming or abrasion of motor gear, the actual inductance curve is different from that in perfect inductance design, and torque output is affected and then the motor efficiency. So the turn-on angle and turn-off angle calculated off-line will not be optimal in actual operation. So it is necessary to adjust turn-on angle and turn-off angle on-line under the premise to maintain driving performance of SRM [2，3，4]. On-line control schemes of angles are presented which adjust turn-on and turn-off angle to optimize the steady-state performance of drive as measured by torque per ampere. And in this article, EPROM is used to adjust turn-on and turn-off angles on-line. Mathematic model of forecast current value after varied turn-on or turn-off angle is established based on the Least Square Method (LSM). The flowchart of searching optimum turn-on or turn-off angle is detailed.

2. REALIZATION OF ON-LINE CONTROLLING TURN-ON AND TURN-OFF ANGLES

According to pulse signal generated by position sensor, microprocessor controls turn-on or turn-off time of phase windings to dominate on-off angles of SRM. The stator discussed in this article has three phases but only one sensory component [5]. When the rotor rotates one cycle, there is only one pulse (P) generated. The mechanical angle between adjoining pulses is 2π. Using the principle of locking phase frequency doubling, 2π is divided into N equal parts, which are in turn marked P_0, P_1, ... P_m, ... P_{N-2}, P_{N-1} ($0 \leqslant m \leqslant N-1$) and position signal P is used as synchronous pulse. Then pulses of different positions can reflect rotor position relative stator. The outputs of the counter are linked to address bus of EPROM. Each counter output value corresponding to the mechanical angle of rotor will be reflected to one unit of EPROM, in other word, each selected unit of EPROM represents only one mechanical angle. data information as 1s or 0s is stored in EPROM as driving signal to control phase windings of SRM. Therefore which phase will be conducted is determined according to the amount of counter, so as to program EPROM. 1s is writed in the relative bit of unit selected by position where winding should be conducted, while 0s in others, then the control of on-off phase winding of SRM can be realized.

For convenience, same turn-on angles θ_{on} and different turn-off angles θ_{off} are stored in units which have same address bus A_N ... $A_{K-1}$and different A_K ... A_{L-1}, different turn-on angles θ_{on} and same turn-off angles θ_{off} are stored in units that have different address bus A_N ... $A_{K-1}$and same A_K ... A_{L-1}. Thus storage table of different θ_{on} and θ_{off} is established in fact. In the operation we connect A_N ... A_{L-1} to the parallel ports of the micro processor (or One-chip computer). And we can easily change the turn-on angle θ_{on} by changing the value of A_N ... A_{K-1}, and change the turn-off angle θ_{off} by changing the value of A_K... A_{L-1}, thereby the on-line control of on-off angles can be realized easily.

3. THE OPTIMIZED CONTROL OF TURN-ON AND TURN-OFF ANGLES

3.1 SRM driving control system

Due to the variations of parameters, there may be difference between actual and designed phase inductance profiles. It is necessary to use a on-line controller which can meet the requirement of driver and optimize performance of SRD. Drive efficiency, torque per ampere and torque ripple may all be used to describe performance of SRD. Drive efficiency depends on many factors, such as structure of converter, switch frequency of chopper, speed of motor and so on. Torque per ampere concentrate on the performance of motor itself. Torque ripples are mainly considered in low speed and application of accurate control. At present, it is difficult to make all these performance optimal at the meantime [3]. Torque per ampere is the target pursued in any drive applications, and its optimization will increase efficiency of drive system, so it is a hot topic of SRM drives system recent years. In the paper, on-line control of on-off angles is presented to optimize the steady-state performance of SRD as measured by torque per ampere.

General control strategy to realize the objection is to find minimum current value by varying turn-on angle and turn-off angle on precondition of given speed and load torque, thus torque per ampere reaches its maximum. The control method is realized by using double close-loop of speed and current. Block diagram of SRD control system which is used to find optimal θ_{on} and θ_{off} is shown in Fig.1.

The driving system adds the optimized control of θ_{on} and θ_{off} on the basis of traditional SRM driving system. The aim is to find the turn-on and turn-off angles that make the current minimum when the load is constant. The control depends on two input values, speed and current.

3.2 Optimized control scheme of turn-on and turn-off angles based on the Least Square Method

Torque sensors are not installed in normal drive system. But to most of load, torque is considered constant when speed attain constant. So, to find maximum torque-current ratio we can find minimum current by varying on-off angles under constant speed, thus θ_{on} and θ_{off} at this moment are optimal on-off angles corresponding to maximum torque-current ratio which are marked with $\theta_{on}{}^{*}$ and $\theta_{off}{}^{*}$.

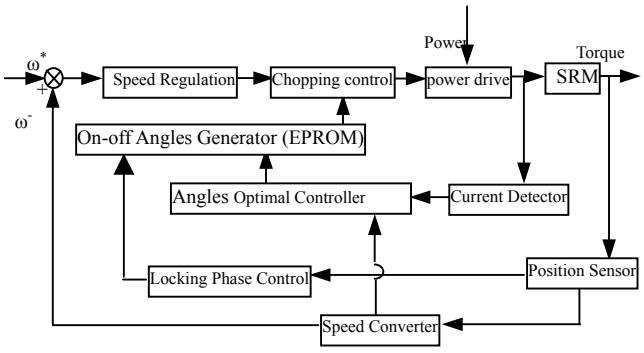


Fig. 1. Block diagram of SRM driving control

When the system reaches the steady state, we sample nearby the optimized turn-on and turn-off angle calculated off-line, obtain sampling point of several groups of turn-on and turn-off angle and relevant virtual current value, and find the valley current value through compare. Then we use the Least Square Method to fit the sampling point, and forecast the winding current if we increase (or decrease) the turn-on and turn-off angle. Increase (or decrease) the turn-on or turn-off angle if the forecasted current drop, and otherwise to decrease (or increase). After that, speed controller correspondingly regulates current I to make system reach steady-state operation again, then sample again and compare to current minimum value to gain new valley current value and the turn-on and turn-off angle at this time. Keep new sampling points to replace old ones, fitting, forecasting and comparing until get the turn-on and turn-off angle when the current is minimum in certain conduct range and within certain error range. And this is the optimized turn-on and turn-off angle. Above is the main thought train of this paper.

To enhance the speed to find the optimized value, we did not search the optimized θ_{on} and θ_{off} at the same time. Instead we at first search the optimized turn-on angle in fixed point which turn-off angle calculated offline is optimized, and then find the optimized turn-off angle. The current values relevant to only 3 turn-on (or turn-off) angles are used in the fitting to reduce the computation time. And use the 3 current values to create fit curve based on Least Square Method. Suppose the curve as:

$$I = P(\theta) = a_0 + a_1\theta + a_2\theta^2 = \sum_{j=0}^{2} a_j\theta^j \qquad (1)$$

Residual Error Expression:

$$\varphi = \sum_{k=n-2}^{n}\left[P(\theta_k) - I_k\right]^2 = \sum_{k=n-2}^{n}\left(\sum_{j=0}^{2} a_j\theta_k{}^j - I_k\right)^2 \qquad (2)$$

To minimize the residual error, as evaluate to extreme value of φ, make

$$\frac{\partial\varphi}{\partial a_i} = 0 \quad ,i = 0,1,2 \qquad (3)$$

And then:

$$\begin{cases} 3a_0 + a_1\sum_{k=n-2}^{n}\theta_k + a_2\sum_{k=n-2}^{n}\theta_k{}^2 = \sum_{k=n-2}^{n} I_k \\ a_0\sum_{k=n-2}^{n}\theta_k + a_1\sum_{k=n-2}^{n}\theta_k{}^2 + a_2\sum_{k=n-2}^{n}\theta_k{}^3 = \sum_{k=n-2}^{n} I_k\theta_k \\ a_0\sum_{k=n-2}^{n}\theta_k{}^2 + a_1\sum_{k=n-2}^{n}\theta_k{}^3 + a_2\sum_{k=n-2}^{n}\theta_k{}^4 = \sum_{k=n-2}^{n} I_k\theta_k{}^2 \end{cases} \qquad (4)$$

In this formula a_0、a_1、a_2 is fitting curve coefficient. θ_k Corresponds to the Kth adjusted turn-on (or turn-off) angle and I_k to sampling current value.

Substitute sampling value (I_{n-2},θ_{n-2})、(I_{n-1},θ_{n-1})、(I_n,θ_n) into Eq. (4) and resolve to get the value of a_0、a_1、a_2. Substitute a_0、a_1、a_2 and the turn-on (or turn-off) angle θ_{n+1} into Eq.(1). We can forecast the current value corresponding to the (n+1)th adjusted turn-on (or turn-off) angle.

3.3 Flowchart to find the optimal turn-on and turn-off angle

Because one position sensor is installed, the perigon 2π has been divided into N equalization. So every counting pulse represents a minimum unit $\theta_{\min}$ of angle adjusting value. There is the larger N and the more particular angle. Enhancing the speed to find the optimal value, in the process we can at first use $k\theta_{\min}$ （k is a random integral）to change the turn-on and turn-off angle, and gradually reduce the range of angle variation. The detailed control flowchart to find the optimized turn-on angle $\theta_{on}{}^{*}$ and turn-off angle $\theta_{off}{}^{*}$ is as Fig.2 shows.

As the way to find $\theta_{on}{}^{*}$ is the same to $\theta_{off}{}^{*}$, in the figure θ represents θ_{on} or θ_{off}. We will use $\theta_{on}{}^{*}$ as an example to illustrate in the following text.

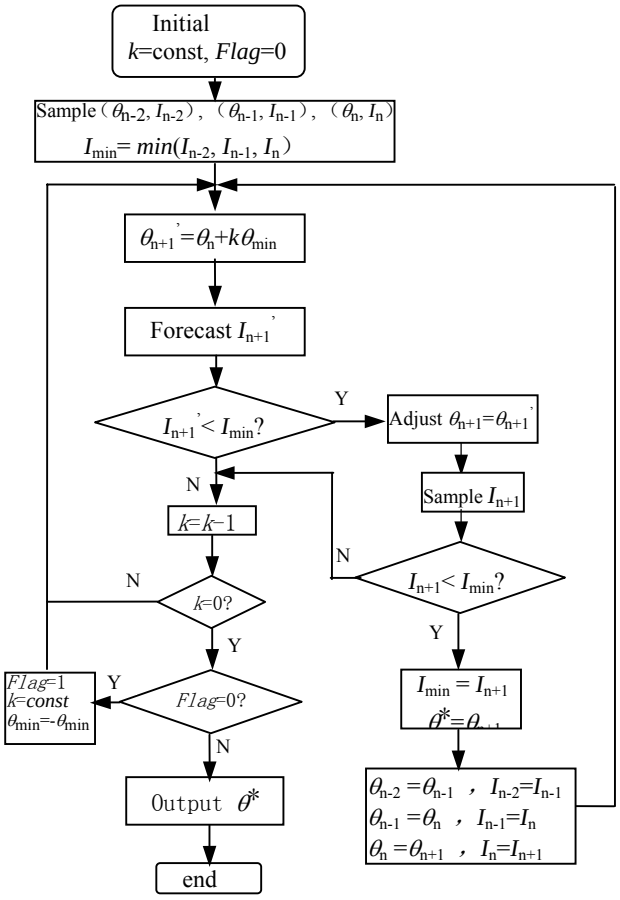
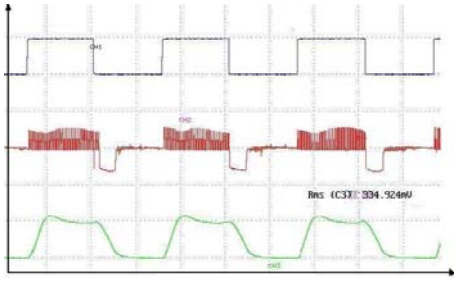


Fig. 2. Flowchart of optimizing θ_{on} and θ_{off}

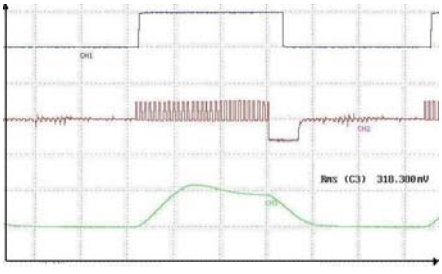
At first using speed controller to run the estimate working point（ω、T), as to ensure the rotating speed to be constant when the loads are constant and the system to reach the steady state. Then fix the turn-off angle θ_{off} at the optimized place θ_{off} * calculated off-line, and according to the optimized turn-on angle calculated off-line, using the output value of counting device that represents SRM position and parallel export of micro processor address the EPROM unit corresponding to different turn-on angle θ_{on}, and to change the winding



t (5ms/div)

Fig. 3. Waves according to off-line

turn-on angle. Accordingly the speed controller adjusts current I to make the system operating steady again. Sampling new current value and changing the turn-on angle twice nearby the optimized position is calculated off-line. Then make it steady and sample new current value again to find the valley current value $I_{\min}$ of the third sampling by comparing. Use the Least Square Method to forecast the current value when the turn-on angle is θ'_{n+1} after the range of $k\theta_{\min}$ variation increases (or decreases) the turn-on angle. Adjust the turn-on angle to θ'_{n+1} if the forecast value is less than present valley current value. Otherwise forecast the current value adjusted after the range of $(k-1)\theta_{\min}$ variation is adjusted. Change the range of $k\theta_{\min}$ variation to decrease (or increase) turn-on angle if the estimate current value after the turn-on angle is adjusted by minimum unit $\theta_{\min}$ is still larger than present value $I_{\min}$. Repeat the process above until the current value when use the minimum unit $\theta_{\min}$ to increase or decrease the turn-on angle at certain conduct range is less than present minimum value. And consider the turn-on angle here is the optimized value $\theta_{on}{}^{*}$. The



t (5ms/div)

Fig. 4. Waves according to the LSM

fix θ_{on} at the located optimized position of $\theta_{on}{}^{*}$ and find $\theta_{off}{}^{*}$ in the same way. Thus optimal turn-on and turn-off angles under given speed and torque are attained. Using them to control turn-on and turn-off of each phase winding can optimize performance of drive as measured by torque per ampere.

4. EXPERIMENT RESULTS

To validate the validity of this scheme to find optimized performance indexes as measured by torque per ampere based on the Least Square Method, we run an experiment on a three-phase 12/8 pole SRM. We test the winding current on the conditions of optimized on-ff angle calculated off-line; add control part to find optimization in many kinds of steady-state rotating speed. The waveforms before and after optimization of the turn-on and turn-off angles are shown as Fig.3 and Fig.4 when the rotating speed is constantly at 500 rpm. The topside square wave shape in these two figures is in the moment when the winding calculated off-line should be on or off. The wave shape in the middle is the voltage of one phase winding. And the bottom wave shape is the current of corresponding phase winding.

The situation when optimized turn-on angle calculated off-line is -3° and turn-off angle is 17° is as Fig.3 shows. At this time the virtual current value of the wave is 334.924mV. And the actual current value is 13.4A after conversion. The wave shape after adding finding optimization control is as Fig.4 shows. The square wave is in the moment when the winding calculated off-line should be on or off, too. In the figure we can clearly recognize the voltage wave changes in the actual on or off moment. The turn-on angle is about -3.3° and the turn-off angle is about 16.2°. And at this time the virtual current value of the winding current wave is 318.380mV. And the actual current value is 12.74A after conversion. It is 5% less than that without finding optimization.

5. CONCLUSION

The traditional angle and position control of switched reluctance motor is just to calculate turn-on and turn-off angles off-line or vary the angle linearly according to the variation of loads and rotating speed. Considering that, the control scheme based on the Least Square Method is presented that is to adjust turn-on and turn-off angles and make the SRM performance indexes optimal as measured by torque per ampere. The scheme is easy to realize, and the process of finding optimization is not easy to be trapped in part surge. The experiment results also prove that the optimized turn-on and turn-off angles calculated off-line are not always really optimized. The improvement of system performance testifies that the finding optimization process is necessary. And for other control scheme of varying on-off angle by software, it just needs to add the finding optimization part.

6. REFERENCES

[1] T.J.E. Miller, “Switched Reluctance Motors and Their Control,” *Magna Physics Publishing and Clarendon Press.* Oxford, 1993

[2] J. Reinert, R.B. Inderka, et al, Optimizing performance in switched-reluctance drives, *IEEE Industry, Applications, Megnetics*, 2000, 6(4):63-70

[3] Pillay, Pragasen, et al, Novel switched reluctance motor drive with optical graphical programming technology, *IEEE Transactions on Industrial Electronics*, 2000, 47(4): 915-919

[4] Omekanda, M. Avoki, A new technique for multidimensional performance optimization of switched reluctance motors for vehicle propulsion, *IEEE Transactions on Industry Applications*, 2003,39(3): 672-676

[5] X.D. Wang, X.L. Wang, “Novel position sensor of Switched reluctance motor based on phase locked principle,” *Journey of Harbin Institute of Technology*. China, vol.4, 1999, 11-14.

Quantum-Behaved Particle Swarm Optimization for designing H∞ structured specified controllers

Maolong Xi[1 2] Jun Sun[1] Wenbo Xu[1]
[1]School of Information Technology Southern Yangtze University, Wuxi, China, 214122
[2]Wuxi Institute of Technology, Wuxi, China, 214121
Email: ximaolong0517@sina.com Email: sunjun_wx@hotmail.com Email: xwb@sytu.edu.cn

ABSTRACT

The conventional output feedback designs of H_∞ optimal control are very complicated and not easily implemented for practical industrial application because of the order of the controller is much higher than that of the plant. In this paper, a quantum-behaved particle swarm optimization (QPSO) for the design of a structure specified H_∞ optimal problem is proposed from sub-optimal perspective. This method is very advantageous for practical control systems. Two examples involving a SISO phase-locked-loop motor speed system with lead/lag type controller and a MIMO super maneuverable F18/HARV fighter aircraft with a PI-type controller are given to illustrate the design procedure and exhibit the effectiveness of the proposed method via a comparison study with an existing GA approach.

Keywords: H∞ optimization，H∞ control, QPSO.

1. INTRODUCTION

In the past decades, H∞ optimal control theories have received a great deal of attention and widely applied to treat control system design problem, such as min-max model matching [1,2] and robust optimization [3~5]. In general, the H optimal output feedback control problem can be solved by two methods. One is based on Youla's parameterization with Nehari theorem or so called A-A-K technique from the frequency domain perspective [6]. The other is based on observer-based state feedback with game theory to solve the Riccati-like equations from time domain viewpoint [7]. However, base on the above two approaches, the order of the derived H∞ optimal controller is much higher than that of the plant. So it is not easy to implement the H∞ optimal controller in practical applications directly. Therefore, search algorithms are employed to treat this problem from sub-optimal point of view such as the gradient-based searching algorithm and genetic algorithm, and preferable performance is obtained using GA in [8,9]. However, the standard genetic algorithm is not a global search optimization, it employs mutation, crossover, and selection as the primary operators for optimization and the selection of them decides the performance of the GA.

In this work, quantum-behaved particle swarm optimization (QPSO) is employed to treat our H∞ optimization problem. QPSO is proposed by Jun Sun et al. and have been proved to be a global search strategy in [10,11]. It has simple equations and only one parameter to control the convergence to the global point. Hence, it is very suitable to treat the H∞ optimization problem.

2. PROBLEM FORMULATION

Consider a control system as shown in Fig. 1., where P(s) is an ni-inputs and no-outputs nominal plants, △P(s) is the plant perturbation, K(s) is the controller, r (t) is the reference input, u (t) is the control input, e (t) is the tracking error, d (t) is the external disturbance and y (t) is the output of the system.

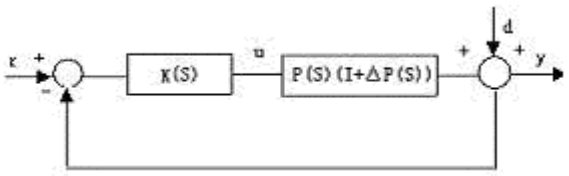


Fig. 1. The control system with plant perturbation and external disturbance

Without loss of a generality, the plant perturbation △P(s) is assume bounded by a known stable function matrix $W_1(s)$, i.e.,

$$\overline{\sigma}(\Delta P(j\omega)) \le \overline{\sigma}(\Delta W_1(j\omega)), \qquad \forall \omega \in [0,\infty)$$

Where $\overline{\sigma}(A)$ denotes the maximum singular value of a matrix A.

The robust H_∞ optimal disturbance attenuation problem for the system in Fig. 1 lies in how to find a controller K(s) to stabilize the closed-loop system and attenuate the external disturbance. That is to say to find all stabilizing controllers, which minimize

$$\|G_{wz}\|_\infty \equiv \left\| \begin{matrix} W_1 S \\ W_2 T \end{matrix} \right\|_\infty = \sup_\omega \overline{\sigma}(G_{zw}(j\omega)) \qquad (1)$$

Where $W_2(s)$ is a stable weighting function matrix specified by designer, S and T are the sensitivity function and complementary sensitivity function respectively. Fig. 2 shows the standard control configuration of S/T mixed sensitivity minimization problem.

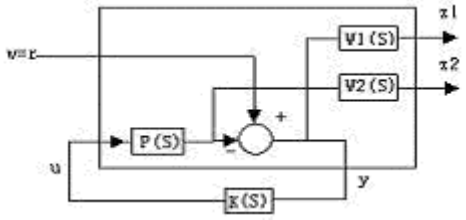


Fig. 2. S/T mixed sensitivity optimization form

The plant G(s) has the following form:

$$\begin{bmatrix} z_1 \\ z_2 \\ \cdots \\ y \end{bmatrix} = \begin{bmatrix} G_{11} & G_{22} \\ G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix} \qquad (2)$$

Where $G_{11} = \begin{bmatrix} W_1 \\ 0 \end{bmatrix}$, $G_{12} = \begin{bmatrix} -W_1 P \\ W_2 P \end{bmatrix}$, $G_{12} = I$, $G_{12} = -P$.

In practice, by using the conventional methods, the optimal solution of the H_∞ problem is very difficult to obtain. To overcome this problem, in this work, we investigate the robust H_∞ optimal disturbance attenuation problem from sub-optimal perspective. On the assumption that γ_{opt} is the optimal value of $\|G_{zw}\|_\infty$ over all stabilizing controller and γ_{sub} is the sub-optimal one that is close to the optimal one in the sense of the H_∞ norm. Then the H∞ sub-optimal control problem is to find all stabilizing controllers K with

$$\|G_{zw}\|_\infty < \gamma_{sub} . \quad (3)$$

In this paper, the structured- specified controller has the following form:

$$K(s) = \frac{N_k(s)}{D_k(s)} = \frac{\alpha_m s^m + \alpha_{m-1} s^{m-1} + \cdots + \alpha_0}{s^n + \beta_{n-1} s^{n-1} + \cdots + \beta_0} \quad (4)$$

Which is assigned with some desired order m and n to minimized the performance criterion (4).

3. QUANTUN-BEHAVED PARTICLE SWARM OPTIMIZATION

Quantum-behaved particle swarm optimization is stochastic optimization algorithm that was originally motivated by the thinking model of an individual of the social organism. In [10,11], Jun Sun et al considered a social organism is a system far more complex than that formulated by particle swarm optimization (PSO) , and a linear evolvement equation is not sufficient to depict it at all. In practice, the evolution of man's thinking is uncertain to a great extent somewhat like a particle having quantum behavior and they introduce quantum theory into PSO and propose a Quantum-behaved PSO algorithm. The experiment results indicate that the QPSO works better than standard PSO on several benchmark functions and it is a promising algorithm in [12].

In QPSO, at every iteration, each particle records its *pbest* (best information of individual) and compares its pbest with those of all other particles in its neighborhood, and population gets the gbest.(best information of the population) .Then its Learning Inclination Point p can be given by Equation (5) after random numbers φ_1 and φ_2 which are generated between 0 and 1.

$$p = (\varphi_1 * pbest + \varphi_2 * gbest\ )/(\varphi_1 + \varphi_2) \quad (5)$$

A mainstream thought point is employed to evaluate the creativity of a particle and the point or Mean Best Position (*mbest*) is defined as the center-of –gravity *gbest* position of the particle swarm which is formulated as equation (6)

$$mbest = \sum_{i=1}^{M} pbest_i / M = \left(\sum_{i=1}^{M} pbest_{i1} / M, \sum_{i=1}^{M} pbest_{i2} / M, \ldots, \sum_{i=1}^{M} pbest_{id} / M \right) \quad (6)$$

Where *M* is the population size.

Therefore, QPSO has iterative equation of following form:

$$x(t+1) = p \pm \beta * \left| mbest - x(t) \right| * \ln(1/u) \quad (7)$$

Where β is called Creativity Coefficient, u is a random number between 0 and 1. $\pm$ is decided by a random number between 0 and 1 in every iteration, when the number is bigger than 0.5, - is used otherwise + is used.

The process for implementing the global version of QPSO is as follows:

Step1 for a given P(s), specify the weighting function matrices W1(s) and W2(s), the controller K(s), the QPSO parameter β and the population size. Initialize a population of particles with random positions on d dimensions in the controller parameter feasible space.

Step2 for each particle, evaluate the desired optimization fitness function in *d* variable.

Step3 compare particle's fitness evaluation with particle's *pbest*. If current value is better *pbest*, then set *pbest* location equal to the current location in *d* dimensional space.

Step4 compare fitness evaluation with the population's overall precious best. If current value is better than *gbest*, then reset *gbest*.

Step5 change the position of particle according to equations (7)

Step6 loop to step2 until a criterion is met, usually a sufficiently good fitness or a maximum number of iterations.

As a result, QPSO has simple form and only one parameter β works on individual particle's convergence speed; therefore it is very suitable to treat our optimization problem.

4. TEST CASE

In this section, two different examples are given to illustrate the proposed design procedures and a comparison study with existing algorithm is carried out to illustrate the effectiveness of the proposed approach. Here, β in QPSO is declined linearly from 1 to 0.45 and population size is 20 in example1, 30 in example2.

4.1 Example 1:

Consider a phase-locked-loop motor speed control system [13] as in Fig. 1., where

$$P(s) = \frac{68.76}{s(1 + 0.05s)}$$

Suppose the system suffers from the external disturbance $d(t) = 0.1e^{-0.1t}\sin(0.8\pi t)$ and the plant perturbation

$$\Delta P(s) = \frac{0.5}{s^2 + 0.2s + 8}$$

In fact the plant perturbation is unknown but bounded by the following known stable function:

$$W_2(s) = \frac{0.6}{s^2 + 0.2s + 8}$$

To treat the robust H_∞ disturbance attenuation problem, the weighting function $W_1(s)$ is chosen as

$$W_1(s) = \frac{0.5s + 0.05}{s^2 + 0.2s + 6.32}$$

Here, we select a lead/lag type controller to tune the system to achieve the robust H_∞ optimal design objective in (1).

$$K(s) = \frac{\alpha_1 s + \alpha_0}{s + \beta_0}$$

The control parameter vector is define as $\theta = [\alpha_1 \quad \alpha_0 \quad \beta_0]^T$, the bound are $\Theta = \left\{ \theta_i \middle| -20 \le \theta_i \le 20, i = 1,2,4 \le \theta_3 \le 20 \right\}$ in [9].

After 50 generations, an appropriate controller

$$K(s) = \frac{0.8211s + 1.8372}{s + 19.998}$$

With the H_∞ norm 0.9140 of (1) is obtained near (2.82) rad/sec. Let the reference input r (t) be a square wave with amplitude ± 1 and period 4π, the output of the system with the derives controller under the plant perturbation and external noise is shown in Fig (3).

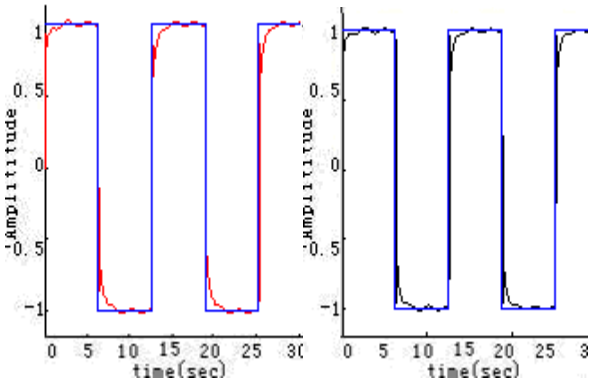


Fig.3.The output of the system in example1under QPSO. Fig.4.The output of the system in example1 in [9]

4.2 Example 2

Consider a longitudinal control system of the super maneuverable F18/HARV fighter aircraft [14] in horizontal flight at an altitude of 15000(ft) with Mach number M=0.24, airspeed $V_T = 238.7(ft/s)$, angle of attack $\alpha = 25(\deg)$ and pitch angle $\vartheta = 25(\deg)$. The trim value or path angle is $\gamma = \vartheta - \alpha = 0(\deg)$ and the trim pitch rate is $q = 0(\deg/s)$. The longitudinal dynamics of the system can be described in the following state-space form:

$$\dot{x} = Ax + Bu = Ax + B_v v$$

$$y = Cx$$

Where matrices A, B_v, C are:

$$A = \begin{bmatrix} -0.075 & -24.05 & 0 & -36.16 \\ -0.0009 & -0.1959 & 0.9896 & 0 \\ -0.0002 & -0.1454 & -0.1677 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$B_v = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$v = \begin{bmatrix} -0.0230 & 0 & -0.0729 & 0.0393 & -0.0411 & 0.1600 \\ -0.0002 & -0.0001 & -0.0004 & 0 & -0.0003 & -0.0003 \\ -0.0067 & -0.0007 & -0.0120 & -0.0006 & 0.0007 & 0.0005 \end{bmatrix} u$$

Suppose the system is encountering with the external disturbance $d_1(t) = d_2(t) = d_3(t) = 0.01e^{-0.2t}\cos(3162.3t)$ and the plant uncertainty

$$\Delta P(s) = \frac{0.012s^2 + 1.2s + 1.2}{s^2 + 20s + 100} I_{3\times 3}$$

Note that the plant uncertainty $\Delta P(s)$ is unknown in fact but bounded by the following stable function matrix:

$$W_2(s) = \frac{0.0125s^2 + 1.2025s + 1.25}{s^2 + 20s + 100} I_{3\times 3}$$

to treat the robust H disturbance attenuation problem, a weighting function matrix is chosen as

$$W_1(s) = \frac{0.25s^2 + 0.025}{s^2 + 0.4s + 10^7} I_{3\times 3} \quad .$$

In this example, a PI type controller is selected to treat the H_∞ optimal disturbance attenuation problem

$$K(s) = \frac{\begin{bmatrix} \alpha_{1_{11}} & \alpha_{1_{12}} & \alpha_{1_{13}} \\ \alpha_{1_{21}} & \alpha_{1_{22}} & \alpha_{1_{23}} \\ \alpha_{1_{31}} & \alpha_{1_{32}} & \alpha_{1_{33}} \end{bmatrix} s + \begin{bmatrix} \alpha_{0_{11}} & \alpha_{0_{12}} & \alpha_{0_{13}} \\ \alpha_{0_{21}} & \alpha_{0_{22}} & \alpha_{0_{23}} \\ \alpha_{0_{31}} & \alpha_{0_{32}} & \alpha_{0_{33}} \end{bmatrix}}{s}$$

Let us define the controller parameter vector θ as $\theta = [\alpha_{1_{11}}, \alpha_{1_{12}}, \alpha_{1_{13}}, \alpha_{1_{21}}, \alpha_{1_{22}}, \alpha_{1_{23}}, \alpha_{1_{31}}, \alpha_{1_{32}}, \alpha_{1_{33}}, \alpha_{0_{11}}, \alpha_{0_{12}}, \alpha_{0_{13}}, \alpha_{0_{21}}, \alpha_{0_{22}}, \alpha_{0_{23}}, \alpha_{0_{31}}, \alpha_{0_{321}}, \alpha_{0_{33}}]^T$

For some practical requirement, all of parameters are assumed in the space region

$$\Theta = \left\{\theta_i \middle| -20000 \le \theta_i \le 20000, i = 1, \cdots, 18\right\}$$

And the stable region according probabilistic decision [9]:

$$\Theta: \begin{bmatrix} 6500 \le \theta_1 \le 8000 & -17000 \le \theta_{10} \le -14000 \\ -8000 \le \theta_2 \le -5000 & -3000 \le \theta_{11} \le 0 \\ -12000 \le \theta_3 \le -9000 & 0 \le \theta 1_2 \le 3000 \\ 5000 \le \theta_4 \le 8000 & -13000 \le \theta_{13} \le -10000 \\ 4000 \le \theta_5 \le 6000 & -6000 \le \theta_{14} \le -3000 \\ -3000 \le \theta_6 \le 0 & 1200 \le \theta_{15} \le 15000 \\ -5000 \le \theta_7 \le -2000 & 2000 \le \theta_{16} \le 5000 \\ 0 \le \theta_8 \le 3000 & -18000 \le \theta_{17} \le -15000 \\ -5000 \le \theta_9 \le -2000 & 4000 \le \theta_{18} \le 7000 \end{bmatrix}$$

After 50 generations, proper controller parameters could be obtained, and the corresponding control parameter are0.6827

$$K(s) = \frac{\begin{bmatrix} 17119 & 5476 & -7425 \\ -7851 & -4857 & 3347 \\ -1665 & -10803 & 11203 \end{bmatrix} s + \begin{bmatrix} 13425 & -18186 & 81 \\ 5396 & -17144 & 9582 \\ 407 & 10004 & 10975 \end{bmatrix}}{s}$$

4.3 Comparison study

We can derive that QPSO receives H∞ norm better than [9] and these norms are found at the same 50 generation in example1. Fig. 3. and Fig. 4. shows the output response of the phase-locked-loop motor speed control system for a square wave under two algorithm. Two controllers have the same tracking response and the population size of QPSO is 20 while there is 100 in [9]. In example2, the population size of the proposed algorithm is 30 while 100 in [9], let the reference inputs $r_1(t) = 0$, $r_2(t) = 1 - e^{-6t}$, $r_3(t) = 1 - e^{-3t}$, Fig. 5.and Fig. 6 show the closed loop response of the F18 with the derived controller under plant uncertainties and external disturbances with two methods, the lines in Fig. 5.is smoother than in Fig. 6.. Fig. 7. and Fig. 8. illustrate singular value line of sensitivity for F18 with controller derived from proposed algorithm.

	GA in[8]	QPSO
Example1	H_∞ norm=0.8323	H_∞ norm=0.9140
	Generation=200	Generation=50
	Population size=100	Population size=20
	Prameter(Operator)=3	Prameter(Operator)=1
Example2	H_∞ norm=0.8194	H_∞ norm=0.6827
	Generation=200	Generation=50
	Population size=100	Population size=30
	Prameter(Operator)=3	Prameter(Operator)=1

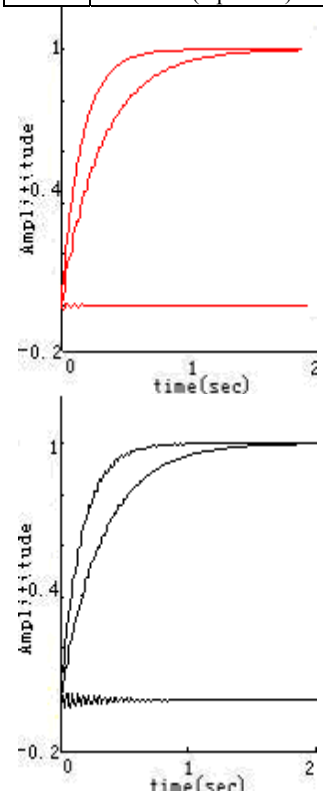


Fig.5.The output of the system in example2 under QPSO

Fig.6.The output of the system in example2 in [9]

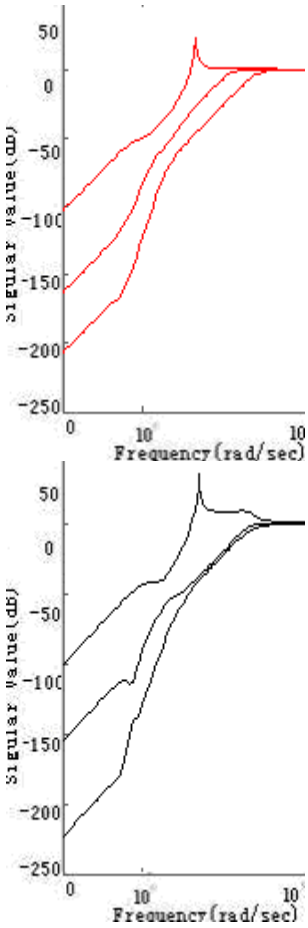


Fig.7.Singular value of sensitivity of F18 with PI controller in example2 under QPSO

Fig.8.Singular value of sensitivity of F18 with PI controller in example2 in [9]

Also the proposed algorithm has only one parameter while three operators in GA of [9], we adjust the convergence speed easier and more feasible in practice applications. Table 1. illustrate comparison of parameters and results in QPSO and GA in [9].

Table 1. Value comparison in QPSO and paper[8]

5. CONCLUSION

A simple H∞ optimal control design via quantum-behaved particle swarm optimization has been proposed in this work. . Simulation results as showed by the comparison study with GA indicate that QPSO can offer an effective and simple method to design structured specified H_∞ controllers which can not be treated by conventional techniques. Furthermore, it can also give very good tracking response with only one parameter. Actually, the proposed algorithm can also be applied to other optimal control designs.

6. REFERENCES

[1] H. Kwakernaak, ”min-max frequency domain performance and robustness optimization of linear feedback system”, IEEE Transactions onAutomatic Control.1985,10,no.AC-30, 994-1004.

[2] T. Alamo, D. Munoz, et al, “Constrained Min–Max

Predictive Control: Modifications of the Objective Function Leading to Polynomial Complexity" IEEE Transactions on Automatic Control, May 2005,50(5): 710-714.
[3] G.L. Wang, Y.F. Li, et al, "Robust control design for stiffness measurement of flexible joint" Machine Learning and Cybernetics, Proceedings of 2005 International Conference on 2: 1160 - 1163.
[4] B. Kouvaritakis, J. A. Rossiter, et al, "Efficient robust predictive control," IEEE Transactions on Automatic Control, Aug. 2000.,45(8).1545–1549,
[5] Y.I. Lee, B. Kouvaritakis, "A linear programming approach to constrained robust predictive control," IEEE Transactions on Automatic Control, Sep. 2000,45(9). 1765–1770.
[6] M. Vidyasagar, Control System Synthesis: A factorization approach Cambridge MA.MIT Press, 1985.
[7] A. Stoorvogel, The H-inf Control Problem:A State Space Approach Prentice Hall, 1992.
[8] I. Kitsios, T.Pimenides, et al, " A genetic algorithm for designing H structured specified controllers", Proceedings of the 2001 IEEE International Conference on Control Applications, Sep.2001,1196-1202.
[9] B.S Chen , Y.M.Cheng, "A structured- specified H∞ optimal control design for practical applications: A Genetic Approach", IEEE Transactions on Control Systems.Technology,vol,6(6),1998,707-719.
[10] J. Sun, B. Feng, et al, "Particle Swarm Optimization with Particles Having Quantum Behavior," IEEE Proceedings of Congress on Evolutionary Computation, 2004.,325-331
[11] J. Kennedy, R. Eberhart, "Particle Swarm Optimization". Proceedings of. IEEE international Conference On Neural Network, 1995,1942-1948.
[12] J.Sun, W.B. Xu, et al, "A Global Search Strategy of Quantum-Behaved Particle Swarm Optimization" Cybernetics and Intelligent Systems Proceedings of the 2004 IEEE Conference, 111-116.
[13] B.C. Kuo, automation controls systems. Englewood Cliffs, NJ:Prentice-Haoo,1991.
[14] P. Voulgaris, L. Valavani, "High performance H_2 and H_∞ designs for supermaneuverable F18/HARV fighter aircraft", AIAA Guidance Navigation Control Conference,1989, l4-16.

PID Parameters Turn and Simulate Based on RBF Neural Networks

Hui Li [1, 2], Dejiang Zhang[2], Jun Lin[1], Wenxue Li[2]
[1.] School of Electrical &Electronic Engineering, Changchun University of Technology, Changchun, Jilin 130012, China
[2.]Key Lab of Geo-Information Exploration & Instrumentation， Jilin University , Changchun ,Jilin 130026, China)
Email: lucylihui@126.com

ABSTRACT

This paper adopts RBF Neural Networks to control the nonlinear and time-varying object accurately. The function of RBF is regarded as the base in middle layer. The input vectors are mapped to dormant space. The adjustable parameters in the networks are the weight of linear combination because the mapping is linear from connotative layer to the space of output solution. This method is adopted to simulate to control the single-input and single-output object. The simulative result shows that this method has the adaptive characters.

Keywords: Radial Basis Function Networks, PID Control, Neural Networks, Nonlinear, Time-varying object.

1. INTRODUCTION

In the process of modern industrial control, many mechanisms of control process are very complex. Under the conditions of some complex perturbation and variational environment, the process parameters and model structure will be changed. In such cases, if the conventional PID controller is used, its parameters will be hard to adjust to the changes of external environment automatically. In order to enable controllers have more adaptability and enable controller parameters be adjusted automatically, neural networks can be used with control method. This paper presents a neural network based on the radial base of PID parameters optimized design method, that constructing neural network PID controller, achieving controller parameters automatically adjust. The result of simulation shows that this method is effectual. [1,2].

2. THE PRINCIPLES OF PID PARAMETERS SETTING OF RBF NETWORKS

The PID controller diagram of automatic adjusting RBF is shown in Fig.1.

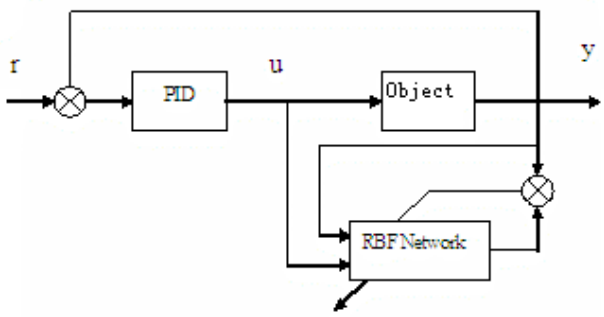


Fig. 1. The diagram of RBF network turning PID

2.1 PID controller

The increment PID controller is usually used in the process of industrial control which control arithmetic is

$$u(k) = u(k-1) + k_p(e(k) - e(k-1)) + k_i e(k) + k_d(e(k) - 2e(k-1) + e(k-2)) \quad (1)$$

Where k_p , k_i , k_d represent the coefficient of the proportion, integral, respectively.

2.2 RBF neural networks identifier [3, 4, 5, 6, 7]

Radial base function (RBF) neural network is a part of the neural network approach, its basic idea is : making radial base function (RBF) as a middle layer hidden modules base, a layer of space implied that the vector can be imported directly to shine upon hidden space. When RBF determine the focal point after the shine upon relations ascertained. The mapping from the hidden space to the output solution space is linear, that the network is implicit export linear weighted and export modules, and what the network can adjust parameters is the linear combination of the right. Technical parameters adjustment using linear adjustment, so the RBF network is faster than BP network learning characteristics, and its approach capable is very good, in theory, as general BP network, the RBF network can approach arbitrary nonlinear shine upon with arbitrary accuracy , so we select RBF as a system to identify vehicles.

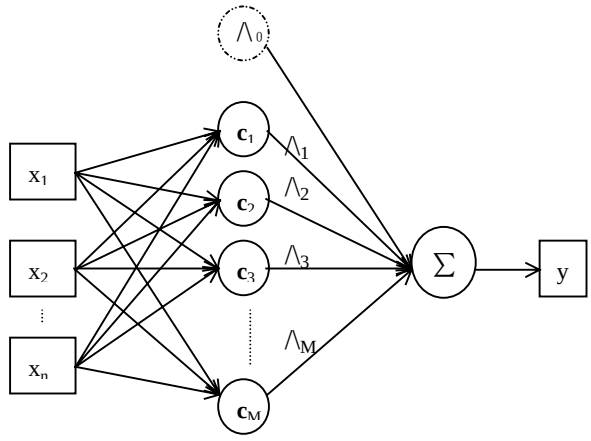


Fig. 2. The network model of based on RBF

Typical RBF network structure showing as Fig.2, it is a

simple two-tier network to the former. The adjacent layer nerves are interconnected. The single-output and number-input network structure is discussed in this paper, so the results can be easily extended to many more of the exports.

RBF can be achieved through two tiers: the input layer is nonlinear shine from the $x \to a_i = \varphi_i(\| x - c_i \| / \sigma_i)$, the output layer is linear shine from the $a_i \to y$, that is:

$$y_i = \sum w_{ji} a_i (j = 1,2,\cdots,r) \quad (2)$$

x is the n dimensions input vectors, a_i is the output of latent layer nodes(The number is q).y is the output vectors of r dimension. c_i is the centre point of the number i RBF. w_{ji} is the $n \times m$ dimensions matrix from a to y. σ_i is a parameter chosen freely. $\| x - c_i \|$ is the of the vector $x - c_i$.

φ_i is the base function of RBF. There is a maximal value at the c_i .When $\| x - c_i \|$ is increased, φ_i attenuates to zero. For the present input $x \in R^n$, a little part units are activated near the centre. The nonlinear function is the gauss function in RBF networks. $\varphi_i = \exp\left[-\left(\frac{\| x - c_i \|}{\sqrt{2}\sigma^2} \right)^2 \right]$.

The study arithmetic of RBF network connection power is:

$$w_{ji}(l+1) = w_{ji}(l) + \beta[y_j^d - y_j(l)]a_i / \alpha^T \alpha \quad (3)$$

β is the study rate. The range is $0 < \beta < 1$ commonly.

2.3 The design of the algorithm program

In RBF neural network, the importing can be identified, exporting variables of neural network basing on the target. The target given is single-input and single-output nonlinear system.

$$y(t) = F[y(t-1),\cdots,y(t-m),u(t-1),\cdots,u(t-(n-m))] \quad (4)$$

If the y of output track whit the change of the input r, so the performance function is

$$J = \frac{1}{2}[r(t+1) - y(t+1)]^2 \quad (5)$$

For the increment PID controller, the error of controlling is

$$e(k) = r(k) - y(k) \quad (6)$$

The inputs of PID are

$$xc(1) = e(k) - e(k-1) \quad (7)$$

$$xc(2) = e(k) \quad (8)$$

$$xc(3) = e(k) - 2e(k-1) + e(k-2) \quad (9)$$

Controlling arithmetic is adopted Eq.(11)，the complete target of NN network is

$$E(k) = \frac{1}{2}e(k)^2 \quad (10)$$

k_p, k_i, k_d:

$$\Delta K_p = -\eta \frac{\partial E}{\partial K_p} = -\eta \frac{\partial E}{\partial y}\frac{\partial y}{\partial u}\frac{\partial u}{\partial K_p} = \eta e(k)\frac{\partial y}{\partial u}xc(1) \quad (11)$$

$$\Delta K_i = -\eta \frac{\partial E}{\partial K_i} = -\eta \frac{\partial E}{\partial y}\frac{\partial y}{\partial u}\frac{\partial u}{\partial K_i} = \eta e(k)\frac{\partial y}{\partial u}xc(2) \quad (12)$$

$$\Delta K_d = -\eta \frac{\partial E}{\partial K_d} = -\eta \frac{\partial E}{\partial y}\frac{\partial y}{\partial u}\frac{\partial u}{\partial K_d} = \eta e(k)\frac{\partial y}{\partial u}xc(3) \quad (13)$$

3. STEPS OF PID ADAPTIVE ALGORITHM

In actual control systems and simulation process using the following steps to do the PID auto adjustment which can identify RBF network of the PID.

1) Set the initial rules of PID controller and the initial parameters of neural network identifying devices.

2) Get $r(t), y(t)$ from sampling, then calculate the $e(k), e(k-1), e(k-2)$.

3) Normalize $xc(1), xc(2), xc(3)$.

4) Calculate the output of controller according to the algorithms, the export target at the same time to identify and RBF neural network devices.

5) Based on the output of the RBF neural networks Identifier and Eq. (3),the value of RBF is revised.

6) Calculate $\frac{\partial \hat{y}(k+1)}{\partial u(k)}$, and revises the parameter of PID controller according Eq. (11), Eq. (12) amd Eq. (13).

7) Return to the step 2).

4. SIMULATION EXPERIMENT AND RESULTS

In order to confirm the validity of the method which the paper proposed, an experiment of simulation can be carried. Given the target controlled is

$$y(k) = \frac{a(k)y(k-1)}{1 + y^2(k-1)} + u(k-1) \quad (14)$$

The coefficient of $a(k)$ is changed along with time slowly, $a(k) = 1.2(1 - 0.8e^{-0.1k})$ .The structure of RBF networks is adopted as the Fig.2.The original value of the power coefficient is the stochastic data in the range of[-0.5,0.5].The PID control response curves and error curves of RBF neural networks identifier are shown in Fig.3 and Fig.4 under the input of the unit step. The PID control response curve of RBF neural networks identifier is shown in Fig.5 under the input of the stochastic square wave.

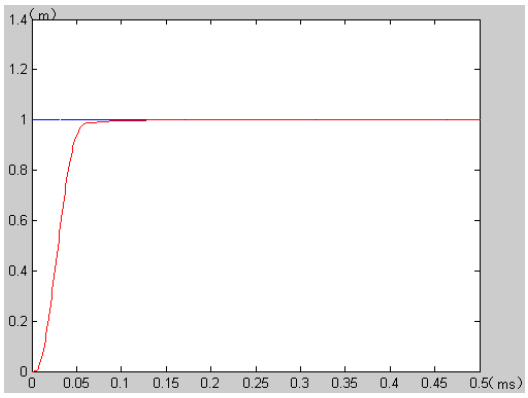


Fig. 3. The response curve of RBF PID control under the input of the unit step

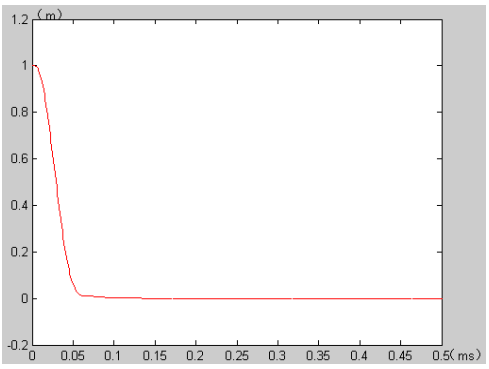


Fig. 4. The error curve of RBF PID control under the input of the step

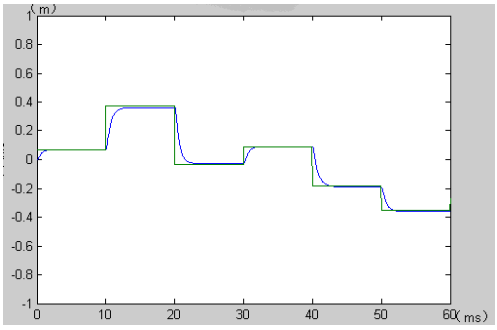


Fig. 5. The response curve of RBF PID control under the input of the stochastic square wave

5. CONCLUSIONS

Since the RBF auto-adapted PID controller algorithm is simple, it has strong auto-adapted performance. It solves the question that PID control system can not be able to study by itself. The simulation result indicates that this controller algorithm can get very good control effect in complex nonlinear system, therefore will have the widespread application prospect in the process of modern industrial control.

6. REFERENCES

[1] H.J. Hu, “Stable and Adaptive PID Control Based on Neural Network”, Journal of Beijing University of Aeronautics and Astronautics,27(2),2001,153-156.

[2] Y.H. Tan,”Adaptive Control Based on BP Neural Networks”,Control Theory and Applications, 11(1),1994,85-88.

[3] L.N. Xu, “Neural Network Control”, Electrical Industry publishing company, 2003.

[4] Y.H. Tao, “PID Control and Application”,Beijing: Machine Industry Press , 2003.

[5] S. Cong, “Survey of Structures, Functions of Artificial Neural Networks and Their Applications in Intelligent Systems”. Information and Control, 30(2), 2001,97-103.

[6] Y. Xue, “On-line self-learning neural network control for pneumatic robot position system”, Proceedings of 2003 international conference on machine learning and cybernetics, 2(5).676-680.

[7] H. Huang, S.W. Lou et al, “Research on autonomous artificial neural network”, Proceedings of 2003 international conference on machine learning and cybernetics, 2(5).1122-1126.

International Trade Customers mining based on Neuro-Fuzzy Decision Tree

Jian-na Zhao, Zhi-peng Chang
School of business and Administration, North China Electric Power University
Baoding, Hebei 071003, China
Email: zhjnzhf@163.com webigren@sina.com

ABSTRACT

Fuzzy decision trees are powerful, top-down, hierarchical search methodology to extract human interpretable classification rules. However, they are poor in classification accuracy. In this paper, neural networks-fuzzy decision tree (Neuro-FDT) is constructed by using the method of Rajen B.Bhatt and Gopal; (See [1]) a fuzzy decision tree structure with neural like parameter adaptation strategy. The method improves FDT's classification accuracy and extracts more accuracy human interpretable classification rules. The fuzzy rules enable a decision-maker to adjust sale strategy according to different customers. The decision-maker may give some special policies to higher profit customers. This paper is a new attempt in international trade. The results of the research indicate that the Neuro-fuzzy decision tree technique is very valid in data base of international trade and it will have a good application prospect in this context

Keywords: Fuzzy Decision Tree, Customers Mining, International Trade, Neuro-Fuzzy Decision Tree.

1. INTRODUCTION

With the rapid development of network technique and globalization of trade, the business activities of international trade transfer gradually to the electronic virtual market. The process of the business operation produce the deeply change. As a trader, marketing and strategy of price are based on the information of customer and the change of market.

Comparing with traditional international trade, Electronic Commerce in International Trade owns more dependable customer's data. If we can analyze these data correctly, we will find much useful knowledge, and discover more business opportunities and finally win in drastic competition. The flowing process of electronic commerce decides that every foreign trade company owns a great deal of information of customers. If this information of customers applies for data mining, we can extract useful knowledge from Neuro-fuzzy decision tree to decide company strategy. So, data mining technology is very important in data base of international trade.

2. NEURO-FUZZY DECISION TREE

2.1 Fuzzifying the training data

In a classification problem, training data can be in the form of either categorical or numerical. When the data are numerical, they need to be fuzzified into linguistic terms. In fact, the fuzzification is a process of conceptualization, which is often used by people to reduce information overload in decision making process. For instance, the credit risk rank data may be perceived in linguistic terms such as high, medium, and low. Their membership functions can be approximately determined based on foreign trade experts' opinion or people's common perception. Alternatively, the membership function may be derived from statistical data [2]. Fuzzy clustering based on self-organized learning can also be used to determine membership functions [3].

Here a simple algorithm is used to generate triangular membership function on numerical data. Assume attribute A has numerical value x .The numerical values of attribute A for all objects $u \in U$ then can be represented by $X = \{x(u), u \in U\}$ .we want to cluster X to k linguistic terms $T_i, i = 1,...,k$. Each linguistic term T_i has a triangular membership function as follows:

$$\mu_{T_1}(x) = \begin{cases} 1, & x \le m_1, \\ (m_2 - x)/(m_2 - m_1), & m_1 < x < m_2, \\ 0, & x \ge m_2, \end{cases}$$

$$\mu_{T_k}(x) = \begin{cases} 1, & x \le m_k, \\ (x - m_{k-1})/(m_k - m_{k-1}), & m_{k-1} < x < m_k, \\ 0, & x \le m_{k-1}, \end{cases}$$

$$T_i(x) = \begin{cases} 0, & x \ge m_{i+1}, \\ (m_{i+1} - x)/(m_{i+1} - m_i), & m_{i-1} < x < m_{i+1}, 1 < i < k \\ (x - m_{i-1})/(m_i - m_{i-1}), & m_{i-1} < x < m_i, \\ 0, & x \le m_{i-1}, \end{cases}$$

The slops of the triangular membership function are selected in the way that adjacent membership functions cross at the membership value 0.5. In this case, the only parameters need to be determined are the set of k centers $M = \{m_i, i = 1,...,k\}$. The centers m_i can be calculated by using Kohonen's feature-maps algorithm. At time 0, the centers $m_i[0]$ are initially set to be evenly distributed on the range of X .The centers are then adjusted iterative in order to reduce the total distance of X to M , defined as

$$D(X, M) = \sum_{x \in X} \min_i \|x - m_i\|. \quad (1)$$

Each iteration at time t consists of three steps:

1. Randomly draw one sample x from X , denoted as $x[t]$;
2. Find the closest center to $x[t]$, i.e. find c such that $\|x[t] - m_c[t]\| = \min_i \|x[t] - m_i[t]\|$
3. Adjust $m_c[t+1] = m_c[t] + \eta[t](x[t] - m_c[t])$ and

keep $m_i[t+1] = m_i[t]$ for $i \neq c$, where t is iteration time $\eta[t]$ is a monotonic decreasing scalar learning rate.

The iteration will continue until $D(X, M)$ converges. [4][6]

It is obvious that the three linguistic terms can be described as S_{01}, S_{02} and S_{03}. The second column of Table 1 shows the membership degree of the attribute S_0 belonging to the three membership functions. Similarly we can find membership function for others attributes and result shows in table1.Using this method, we can get training data shows in Table1

Table1**.** Training Set With Fuzzy Representation

NO	ATTRIBUTE										classification		
	S_0			$S_1(S_4)$			S_2		S_3				
	S_{01}	S_{02}	S_{03}	S_{11}	S_{12}	S_{13}	S_{21}	S_{22}	S_{31}	S_{32}	H	M	L
1	1.0	0.0	0.0	0.5	0.5	0.0	0.0	0.0	0.0	1.0	0.8	0.6	0.0
2	0.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0	0.8	0.2	0.0	0.0	1.0
3	0.0	0.9	0.1	0.0	0.9	0.1	0.1	0.9	0.7	0.3	0.0	0.0	1.0
4	0.9	0.1	0.0	1.0	0.0	0.0	0.8	0.2	0.4	0.6	0.0	0.8	0.2
5	0.8	0.2	0.0	0.6	0.4	0.0	0.0	1.0	0.0	1.0	1.0	0.7	0.0
6	0.0	0.7	0.3	0.8	0.2	0.0	0.1	0.9	0.2	0.8	0.3	0.6	0.1
7	0.2	0.7	0.1	0.3	0.7	0.0	0.2	0.8	0.3	0.7	0.9	0.1	0.0
8	0.0	0.1	0.9	0.7	0.3	0.0	0.5	0.5	0.5	0.5	0.0	0.0	1.0
9	0.0	0.7	0.3	0.0	0.3	0.7	0.7	0.3	0.4	0.6	0.2	0.0	0.8
10	0.0	0.3	0.7	0.0	0.0	1.0	1.0	0.1	0.1	0.9	0.0	0.0	1.0
11	0.0	1.0	0.0	0.0	0.2	0.8	0.2	0.8	0.0	1.0	0.7	0.0	0.3
12	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.6	0.4	0.7	0.3	0.2	0.8
13	0.9	0.1	0.0	0.0	0.3	0.7	0.0	1.0	0.9	0.1	0.0	0.3	0.7
14	0.7	0.3	0.0	1.0	0.0	0.0	1.0	0.0	0.2	0.8	0.4	0.7	0.0
15	0.2	0.6	0.2	0.0	1.0	0.0	0.3	0.7	0.3	0.7	0.7	0.2	0.1
16	0.9	0.1	0.0	0.2	0.8	0.0	0.1	0.1	0.9	1.0	0.0	0.0	1.0

Here, $S_0 \rightarrow$ Commodity category $S_1(S_4) \rightarrow$ Payment risk rank $S_2 \rightarrow$ Count of bargaining $S_3 \rightarrow$ Credit rank H $\rightarrow$ High M $\rightarrow$ Medium L $\rightarrow$ Low

2.2 Algorithm of fuzzy decision tree

The fuzzy decision tree approach considered here results in the construction of a series of fuzzy rules characterized in linguistic terms. The inductive method proposed by Yuan and Shaw [7], and which is employed in the current paper, is based on measures of cognitive uncertainties. In particular, it focuses on the minimization of classification ambiguity in the presence of fuzzy evidence. This section provides a brief description of the functions used in the fuzzy decision tree method.

A membership function $\mu(x)$ of a fuzzy variable Y defined on X, can be viewed as a possibility distribution of Y on X, that is $\pi(x) = \mu(x)$ for all $x \in X$ the possibility measure $E_a(Y)$ of ambiguity is defined as

$$E_a(Y) = g(\pi) = \sum_{i=1}^{n} (\pi_i^* - \pi_{i+1}^*) \ln[i] \qquad (2)$$

where $\pi^* = \{\pi_1^*, \pi_2^*, ..., \pi_n^*\}$ is the permutation of the possibility distribution $\pi = \{\pi(x_1), \pi(x_2), ..., \pi(x_n)\}$[13] sorted so that $\pi_i^* \geq \pi_{i+1}^*$ for $i = 1, ..., n$ and $\pi_{n+1}^* = 0$ [8] [4].The ambiguity of attribute A is then given as

$$E_a(A) = \frac{1}{m} \sum_{i=1}^{m} E_a(A(u_i)) \qquad (3)$$

where $E_a(A(\mu_i)) = g(\mu_{T_s}(\mu_i) / \max_{1 \leq j \leq s}(\mu_{T_j}(\mu_i)))$, with T_j the linguistic scale used within an attribute for m cases. When there is overlapping between linguistic terms of an attribute or between classes, then ambiguity exists. As described by Kosko [5] the fuzzy subsethood $S(A,B)$ measures the degree to which A is a subset of B and is given by

$$S(A,B) = \frac{\sum_{\mu \in U} \min(\mu_A(\mu), \mu_B(\mu))}{\sum_{\mu \in U} \mu_A(\mu)} \qquad (4)$$

Given fuzzy evidence E, the possibility of classifying an object to class C_i can be defined as

$$\pi = (C_i \mid E) = \frac{S(E, C_i)}{\max_j S(E, C_j)} \qquad (5)$$

where $S(E, C_i)$ represents the degree of truth for the classification rule. With a single piece of evidence (a fuzzy value for an attribute), then the classification ambiguity based on this fuzzy evidence is defined as

$$G(E) = g(\pi(C \mid E)) \qquad (6)$$

The classification ambiguity with fuzzy partitioning $P = \{E_1, ..., E_k\}$ on the fuzzy evidence F, denoted as $G(P \mid F)$, is the weighted average of classification ambiguity with each subset of partition:

$$G(P \mid F) = \sum_{i=1}^{k} w(E_i \mid F) G(E_i \cap F) \tag{7}$$

where $G(P \mid F)$ is the classification ambiguity with fuzzy evidence $w(E_i \mid F)$, and where $w(E_i \mid F)$ is the weight $(w(\square))$ which represents the relative size of subset $E_i \cap F$ in F:

$$w(E_i F) = \frac{\sum_{\mu \in U} \min(\mu_{E_i}(\mu), \mu_F(\mu))}{\sum_{j=1}^{k} (\sum_{\mu \in U} \min(\mu_{E_j}(\mu), \mu_F(\mu)))} \tag{8}$$

In summary, attributes are assigned to nodes based on the lowest level of ambiguity. A node becomes a leaf node if the level of subsethood (based on the intersection of the nodes from the root) is higher than some truth value , when assigned to the whole of the decision tree—in this case when the classification from the leaf node is to the decision class with the largest subsethood value [7][9]

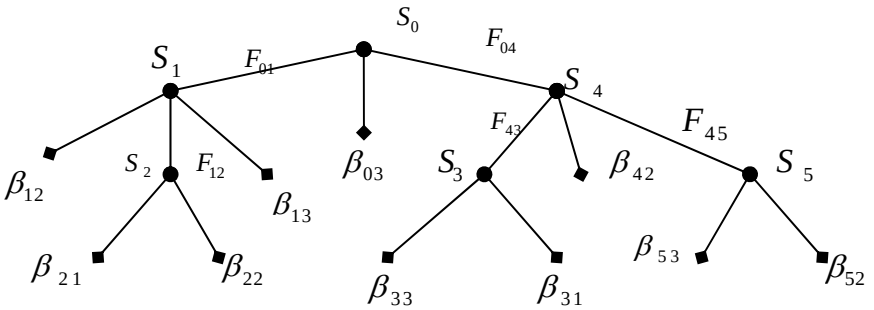


Fig.1. Fuzzy Decision Tree by Using Our Proposed Algorithm to Train Table1

2.3 Neuro-fuzzy decision tree construction

Due to the size and performance of FDT is severely affected by fuzziness control parameter (α -cut) and leaf selection threshold (β_{th}), however, guide rules of selecting α and β_{th} are very hard to find in the existing fuzzy decision tree literature. Neuro-FDT incorporates the merits of neural learning algorithms into the feedback cycle of hierarchical FDT. The method significantly improves the classification accuracy of FDT without compromising the comprehensibility, the FDT structure has been kept intact during the parameter adaptation stage. Ratenb.B Bhatt and M. Gopal proposed back propagation learning to be applied directly on FDT structure by traversing back from each leaf node to root node. Neuro-FDT includes one forward cycle of FDT induction and then several back propagation iterations of tuning the FDT parameters (membership functions and leaf certainties). This strategy doesn't disturb the hierarchical structure of FDT and effectively tune the tree parameters, while preserving the interpretability.

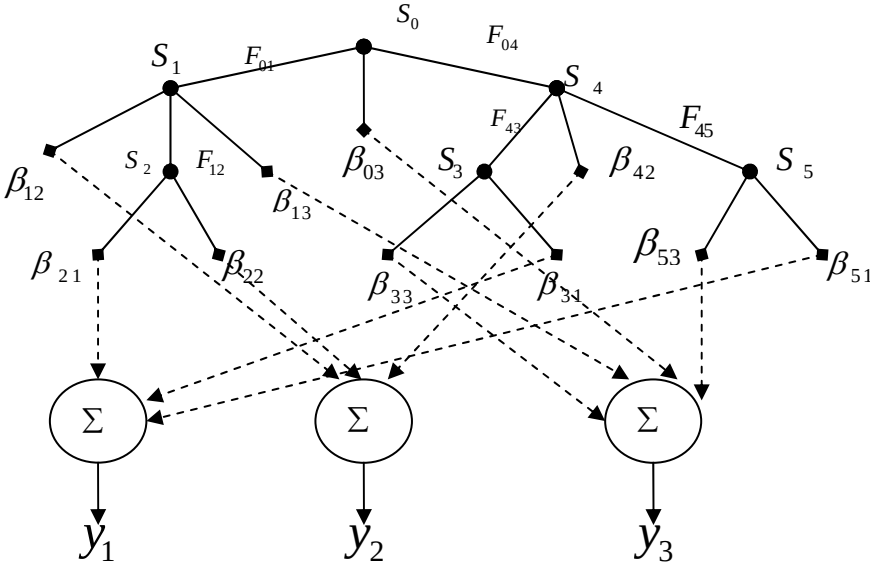


Fig.2. Fuzzy Decision Tree for Direct Back Propagation

Fig.2 shows the basic Neuro-FDT structure with three summing nodes added to it to carry out inference. From all the leaf nodes, certainty factors corresponding to class 1 (high profit) are summed up to calculate y_1. Same way,

certainty factors corresponding to class 2 (medium profit) are summed up to calculate y_2 and certainty factors corresponding to class 3 (low profit) are summed up to calculate y_3.

For an arbitrary pattern, the firing strength of lth class at mth leaf node is given by

$$\mu^{i}_{path_m} \times \beta_{ml} \tag{9}$$

Each $path_m$ ($m = 0,...,5$) is defined on the premise space composed of input features available in traversing from root node to mth leaf node. Where $\mu^{i}_{path_m}$ is membership degree of $path_m$.Which can calculated as

$$\mu^{i}_{path_m} = \prod_j \mu F_{jm}(s^i_j) \tag{10}$$

β_{ml} ($0 \le \beta_{ml} \le 1$; $l = 1,...,q$) is the degree of certainty, with which $path_m$ can classify the class l. In Fig.2, $path_5$ can classify 'down profit' with the certainty of β_{53} and 'high profit' with the certainty of β_{51}. F_{jm} is jth variable's membership function available on mth path.

Firing strengths of all the leaf nodes for a particular class l are summed up to calculate the prediction certainty y^i_l ($l = 1,...,q$) of ith pattern through FDT,

$$y^i_l = \sum_{m=0}^{5} \mu^{i}_{path_m} \times \beta_{ml} \tag{11}$$

where $0 \le y^i_l \le 1$. For example, in Fig.2, prediction certainties for 'high profit' (y_1), 'medium profit' (y_2) and 'low profit' (y_3) are to be calculated by

$$y^i_1 = \sum_{m=0}^{5} \mu^{i}_{path_m} \times \beta_{m1}$$

$$y^i_2 = \sum_{m=0}^{5} \mu^{i}_{path_m} \times \beta_{m2}$$

$$y^i_3 = \sum_{m=0}^{5} \mu^{i}_{path_m} \times \beta_{m3}$$

For example, for Neuro-FDT of Fig.2,

$$y^i_1 = \sum_{m=0}^{5} \mu^{i}_{path_m} \times \beta_{ml}$$

$$= \mu^{i}_{path_2} \times \beta_{21} + \mu^{i}_{path_3} \times \beta_{31} + \mu^{i}_{path_5} \times \beta$$

To fuzzify input attributes, the method select Gaussian membership functions out of many alternatives, due to its differentiable property. For ith pattern membership degree of $path_m$ can be calculated by

$$\mu^{i}_{path_m} = \prod_j \mu F_{jm}(s^i_j) = \prod_j \exp\left(\frac{(s^i_j - c_{jm})^2}{2\sigma^2_{jm}}\right) \tag{12}$$

where c_{jm} and σ_{jm} are center and standard deviation (width) of Gaussian membership of jth input attribute on mth path of F_{jm}.

The method defines as the error function of the FDT a differentiable function like the mean-square-error E,

$$E = \frac{1}{2n} \sum_{l=1}^{q} \sum_{i=1}^{n} (d^i_l - y^i_l)^2 \tag{13}$$

where n the total is number of training patterns, d^i_l and y^i_l is the desired class of ith pattern through Neuro-FDT, respectively.

The necessary condition for the minimization of error is that its differentiations with respect to the parameters Gaussian center locations, Gaussian widths, and certainty factors are all vanish. The leads to the parameter update rule

$$\theta^{\tau+1} = \theta^{\tau} - \eta \frac{\partial E}{\partial \theta} \tag{14}$$

For FDT structure with Gaussian membership functions, we obtain the following update rules for the adaptation of the parameters centers, widths, and certainty factors.

$$\beta^{\tau+1}_{ml} = \beta^{\tau}_{ml} + \frac{\eta}{n} \sum_{i=1}^{n} (d^i_l - y^i_l) \mu^{2}_{path_m} \tag{15}$$

Using the pseudo code of the Neuro-FDT strategy to train and get degree of certainty β shown below Table1. (The details see [1]).

Table1. Degree of certainty β of Neuro-FDT

β_{03}	1.00	β_{31}	1.00
β_{12}	0.92	β_{33}	1.00
β_{13}	1.00	β_{42}	0.86
β_{21}	1.00	β_{51}	1.00
β_{22}	0.91	β_{53}	0.87

2.4 Fuzzy Rules

In Fig.2, We can extract a human interpretable weighted fuzzy classification rule of form "If " $path_m$ Then $leaf_m$ ".Five fuzzy classification rules shown below can be extracted from proposed Neuro-FDT. These rules conform to actual foreign trade experiences. According to these rules, we can predict which customers may bring higher profit to foreign trade company and adjust sale strategy accord to different customers. We may give some special policies to higher profit customers. [11]

If (S_0 is F_{01} ∧ S_1 is F_{12}) Then y = High Profit (β_{21})
and y =Medium Profit (β_{22})

If (S_0 is F_{04} ∧ S_4 is F_{45}) Then y = Low Profit (β_{53})
and y =High Profit (β_{51})

If (S_0 is F_{04} ∧ S_4 is F_{43}) Then y = Low Profit (β_{33})
and y =High Profit (β_{31})

If (S_0 is F_{01}) Then y = Medium Profit (β_{12})
and y =Low Profit (β_{13})

If (S_0 is F_{04}) Then y = Medium Profit (β_{42})

3. FURTHER EXAMINATION OF NEURO-FDT'S CREDIBILITY

This research will use classification error rate to evaluate the result of classification. We select 100 examples for testing. Through examination of group of test examples, we can find accurate rate of forecast is 80%. So the Neuro-FDT has high classification accurate rate and can find high profit customer for international trade company. It is shown in Table2, 3.

Table2. Classification Accurate Rate for FDT

		Result of Prediction		Sum	Accuracy
		High	Low		
Total	High	50	10	60	80%
	Low	10	30	40	
%	High	83.33%	16.67%	100%	
	Low	25%	75%	100%	

Table3. Classification Accurate Rate for Neuro-FDT

		Result of Prediction		Sum	Accuracy
		High	Low		
Total	High	55	5	60	90%
	Low	5	35	40	
%	High	91.67%	8.33%	100%	
	Low	12.5%	87.5%	100%	

4. CONCLUSIONS

This paper presents a new FDT for customer mining of International Trade. Most of the existing data mining techniques are not so efficient for international trade e-commerce databases. FDT is based on minimum classification ambiguity to select expanded attributes. Proposed method has some advantages such as fairly simple but powerful strategy for improving the classification accuracy of FDT without compromising the comprehensibility and high efficiency of the generated useful rules to foreign trade company.

This paper is based on many domestic and international research findings and application of FDT to mining useful rules and help foreign trade company to decide sale strategy. Empirical study proves that Neuro-FDT has good result of prediction.

5. REFERENCES

[1] Rajen B.Bhatt, M.Gopal, Neuro-Fuzzy Decision Trees, World Scientific Publishing Company, 2005, pp. 63-78.

[2] M.Rcivanlar, H.J. Trussell, Constructing Membership Functions Using Statistical Data, Fuzzy Sets and System. (1986) Vol 18, pp.1-14.

[3] T.Kohonen, Self-Organization and Associative Memory, Springer, Berlin. 1988.

[4] Yufei. Yuan, Michael J. Show, Induction of Fuzzy Decision Trees, Fuzzy Sets and System, 1995 Vol 69, pp. 125-139.

[5] Wang Xi-Zhao, Ling Wei—De, A Comparison of two Heuristics for Fuzzy Decision Tree Generation, Journal of Hebei University, 2000, Vol 20, pp. 1-7.

[6] Han Jiawei, kamber M, Data mining Concepts and Techniques, Morgan Kaufmann Publishers, Inc, 2001.

[7] Yuan Y, Shaw M, Induction of fuzzy decision trees, Fuzzy Sets and Systems, 1995, Vol69, pp. 125–39.

[8] Miller GA, The magical number seven, plus or minus two: some limits on our capacity for processing information. The Psychological Review, 1956, pp. 81–97.

[9] Wang X, Chen B, Qian G, Ye F, On the optimization of fuzzy decision trees Fuzzy Sets and Systems, Vol12, pp. 117–25

[10]Kosko B, Fuzzy entropy and conditioning, Information Science, 2000, Vol30, pp. 165–74.

[11]Li Yang, Pan Yu, Mining Data Application for E-Commerce of International Trade, Journal of Nanjing University of Technology, 2004, Vol 26, pp. 55-57

Genetic Algorithm for Multivariable Optimal Control Of Activated Sludge Process*

Zaiwen Liu, Lifeng Cui, Yuanfeng Zhang, Hong Xue, Siying Lu
School of Information Engineering, Beijing Technology and Business University
Beijing, 100037, China
Email: liuzw@th.btbu.edu.cn
And
Chunzhi Zhang
Beijing Vocational and Technical Institute of Industry
Beijing, 100042, China
Email: cslz@public.bta.net.cn

ABSTRACT

A genetic algorithm for multivariable optimal control with the lowest operational cost by limiting total substrate discharge in activated sludge process was discussed. Based on satisfying the requirements of precision, binary coding was used to express units, and 20 bits of binary digits express *DO* (Dissolved Oxygen), Q_w (sewage discharge amount) separately. An improved Euler formula was applied to solve the state equation and the solution was brought in fitness function. In order to transform the restricted problems into unrestricted ones, some penalty items should be added in fitness function to punish the unfeasible solutions. The optimal control can be imitated by manual or semiautomatic control based on the changing principles of optimal control variable *DO* and Q_w in the wastewater treatment system. This algorithm avoids the difficulty of guessing iteration initial value and increases calculation efficiency, and the results of this research are practical and realistic.

Keywords: Genetic Algorithm, Optimal Control, Activated Sludge Process, Wastewater Treatment, Mathematic Model, State Equation.

1. INTRODUCTION

Optimal control, which is the highest control mode in control process, is applied widely in some of industrial process. By solving the state equation and objective function (performance index or index function) under constraint conditions, the system can be operated and performance indexes can be optimized furthest.

If control variables aren't restricted or permissible control is employed, it can be discussed through classical calculus of variation and L.S. Pontryagin's maximum theory. However, control has been restricted in many practical problems. Now, one good solution about restricted optimal control is to transform it into an unrestricted problem by penalty function. There are many traditional algorithms to unrestricted optimal, such as gradient method (fastest descent method), Newton method and its various deformation, conjugate gradient method, changing dimension method (including DFP method, Pearson method), Powell method etc., but only local solution can be obtained through these algorithms. Furthermore, restricted optimal control problem is to be resolved by means of penalty function method, which asks penalty factor to be infinite. If some traditional algorithms are adopted at that time, pathological form problem is easy to happen. This restricts the application of traditional algorithms further [1, 2].

Because of the problems mentioned above, genetic algorithm is brought in the optimal control of process control, such as sewage treatment process. Genetic algorithm, based on natural selection and group heredity mechanism, is an effective searching method to resolve optimal problem. It is an algorithm looking for superior solution in overall situation [3].

Comparing with traditional optimal algorithms, genetic algorithm is of these characteristics: strong robust, parallel processing, wide application etc., furthermore, genetic algorithm can be applied to some complicated nonlinear problems. It is only function values, not differential coefficient that needs to be obtained in Genetic algorithm, consequently, in the resolution of restricted optimal control problem with penalty function method, pathological form problem are not caused even the value of penalty factor is very big. That decreases the calculation greatly.

* This study is supported by Beijing Natural Science Foundation (*No.*4062011) and Beijing Education Committee (*No.*200589)

2. MATHEMETIC MODEL OF WASTEWATER TREATMENT PROCESS FOR MULTIVARIABLE OPTIMAL CONTROL

As the water quality and water quantity of inflow change every moment, the wastewater treatment process and industrial effluent are operated in a state of instability. This is to cause the ceaseless changes of other parameters and outflow quality. In recent years, automatic control of sewage disposal plant has been spread and extended increasingly, however, its optimal control has not been realized till today.

At present, there are two problems existing commonly in sewage treatment plant: the density or the fluctuation of substrate (organic material) from outflow is high and the operation fee is expensive. This paper researches on optimal control by connecting these two problems, considering sewage discharge amount (Q_w) and dissolved oxygen (*DO*) concentration in aeration basin as control variables, the sum of operation cost (energy consumption) including surplus sludge treatment, inverse sludge and aeration basin as performance indexes, water quality of outflow as the constraint. Obviously, Q_w and *DO* are the most important input parameters in the operation of activated sludge process

except water quality and water quantity of inflow. Most of the operation cost in activated sludge process is involved in the performance indexes of this research. Based on meeting the requirement of outflow water quality, the most important aim is to decrease the operation cost.

2.1 Foundation of Basic State Equation

The metabolism of tiny organism and dissolved oxygen of inflow in secondary sedimentation tank can be ignored in this research. In order to control the discharged sewage amount quantificationally, sludge should be discharged from aeration basin directly. Fig. 1. shows the system.

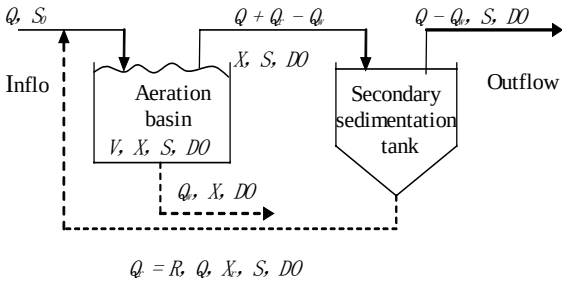


Fig. 1. Activated sludge process map

Where Q, Q_r, Q --- flow quantity of inflow, inverse sludge and sludge discharge (m^3/d)

X, X_r --- MLSS and inverse sludge density (*mg/L*)

So, S --- substrate density (BOD) of inflow and outflow (aeration basin)

DO--- dissolved oxygen density (*mg/L*)

V--- available capacity of aeration basin (m^3)

R--- inverse sludge ratio

According to literature [4], let dissolved oxygen density *DO* and sludge discharge amount Q_x be control variables, let substrate density *S* and tiny organism density of aeration basin be state variables, so the basic state equation of activated sludge process are as the equations:

$$\frac{dX}{dt} = X\left(\frac{Y_K S}{K_S + S} + K_d\right)\frac{DO}{K_0 + DO} - \frac{Q_W X}{V} \qquad (1)$$

$$\frac{dS}{dt} = \frac{Q(S_0 - S)}{V} - \frac{X_K S}{K_S + S} - \frac{DO}{K_0 + DO}$$

In this equation, *X, S, Q_w, DO* all change following time.

They are functions of time *t* and can be expressed as *X(t), S(t), $Q_w(t)$* and *DO(t)* [4].

2.2 Foundation of Basic Performance Indexes

Obviously, the operation cost of sewage treatment system is relative to control variables *DO* and Q_w. Let everyday operation cost be *Jc*, which consists of disposal cost of surplus sewage J_1, inverse sewage cost J_2, oxygen supply cost J_3:

$$J_C = J_1 + J_2 + J_3 \qquad (2)$$

In nonstable state, the operation cost of activated sludge process for one day can be expressed as by function [5]:

$$J_C = \int_0^1 \left\{ AQ_W X + \frac{BX(Q - Q_W)}{X_T - X} + \frac{C_1(D_S - DO)}{D_S - DO} \right. \qquad (3)$$
$$\left. \left[\frac{VXDO}{K_0 + DO}\left(\frac{akS}{K_S + S} + 1.42K_d\right) + QDO \right] \right\} dt$$

In order to fit for the restricted conditions of organic discharge gross, one state variable *Z(t)* is added to state equation. *Z(t)* represents that organic discharge gross of outflow everyday changes following time *t*.

$$Z(t) = \int_0^t QSdt \qquad (4)$$

Because the initial value of *Z Z(0)=0*, the end value *Z(1)* represents organic discharge gross in every day

$$Z(1) = \int_0^1 QSdt \qquad (5)$$

The restricted condition can be expressed as:

$$Z_S - Z(1) \geq 0 \qquad (6)$$

Z_s --- organic discharge gross allowed in every day. Let

Z_s*=150 kg (BOD)/d*

Which represents that average outflow *BOD* is 15mg/L.

3. GENETIC ALGORITHM OF OPTIMAL CONTROL

3.1 Characters of Genetic Algorithms

Genetic algorithm, based on natural selection and group heredity mechanism, is an effective searching method to resolve optimal problem. It imitates the propagation, amphimixis and mutation in natural selection process and natural heredity process. Every possible solution is considered as every unit (chromosome) of colony; furthermore, every unit is coded in the form of character string. After that, every unit is evaluated in terms of expected objective function and an adaptive degree is put forward. Some units are always produced randomly when genetic algorithm iterates.

According to the adaptive degree of units, which can be operated genetically through genetic operator, superior units can be saved, inferior ones are eliminated, a group of new units can be obtained. Because inheriting excellent character from the former generation, these new units are superior to the former ones and the algorithm develops toward the direction of better solutions. Genetic algorithms need not complicated calculations to resolve complex optimal problems. The optimal solutions can be obtained only by several kinds of operators (choosing, crossover, mutation etc.). It is an algorithm looking for superior solution in overall situation [6, 7].

3.2 Design of Genetic Algorithms

1) Coding of Parameters
Because binary coding has the advantages: easy operation for coding and decoding and easy realization for crossover, mutation etc. This paper applies binary coding to express units. Based on satisfying the requirements of precision, 20 bits of binary digits are employed to express *DO*, Q_w separately.

2) Enactment of Initial Colony
Initial solution colony is made of N feasible solutions by random method.

3) Calculation of Fitness function
In order to obtain fitness function of every unit, state equation (1) needs to be calculated first and $x(t)$, $s(t)$ are obtained. Put *DO*、Q_W to solve $x(t)$、$s(t)$ with the modified Euler formula, space state equation (1) can be expressed as:

$$\begin{cases} y' = f(x, y, z) & y(x_0) = y_0 \\ z' = g(x, y, z) & z(x_0) = z_0 \end{cases} \tag{7}$$

Where *x* express time, *y means X , and z means S.* Improved Euler form (forecast formula) is applied to state equation ,and can be expressed as:

$$\begin{cases} \tilde{y}_{n+1} = y_n + hf(x_n, y_n, z_n) \\ \tilde{z}_{n+1} = z_n + hg(x_n, y_n, z_n) \end{cases} \tag{8}$$

Its revised formulas are as following:

$$\begin{cases} y_{n+1} = y_n + \frac{h}{2}[f(x_n, y_n, z_n) + f(x_{n+1}, \tilde{y}_{n+1}, \tilde{z}_{n+1})] \\ z_{n+1} = z_n + \frac{h}{2}[g(x_n, y_n, z_n) + g(x_{n+1}, \tilde{y}_{n+1}, \tilde{z}_{n+1})] \end{cases}$$

Where:

$$h > 0,\ x_n = x_0 + nh,\ y_n \approx y(x_n),\ z_n \approx z(x_n).$$

The solution is brought in fitness function. The restricted conditions of end value (organic discharge gross in every day) are calculated in penalty strategy. In order to transform the restricted problems into unrestricted ones, some penalty items should be added in fitness function to punish the unfeasible solutions.

$$V(DO,Q_W)=F(DO,Q_W)+G\times P(DO,Q_W) \tag{10}$$

G --- penalty coefficient

4) Selection
Sort algorithm of standard geometry is adopted in selective function. Selective probability P_i is defined for every unit:

$$p_i = \frac{q(1-q)^{n-i}}{1-(1-q)^n} \tag{11}$$

P_i --- The selective probability of unit *i*

n --- The number of unit in colony
q --- Optimal probability of selection

The selective probability of a chromosome only lies on its sequence by big and small in the colony. The offspring number of his chromosome is not affected by the situation that adaptive value of chromosome is bigger or smaller than that of others, so that over-centralized state of chromosomes is avoided after selection to a certain extent.

5) Crossover
One-point crossover technology is used in it. Based on crossover probability P_c, one crossover point is chosen randomly for every couple of selected units. Two new units (*child 1, child 2)* are born after the part exchange between the couple selected.

$$child1=parent1_{1\sim curr\_site}+parent2_{curr\_site+1\sim sz}$$

$$child2=parent2_{1\sim curr\_site}+parent1_{curr\_site+1\sim sz} \tag{12}$$

curr_site --- crossover position

6) Mutation
Parameter P_m is defined as aberrant probability which represents that units of expectation value $P_m \cdot n$ is selected for aberrant operation in this colony. Every parent generation selected is operated aberrantly (reverse every bit of the unit) and one generation is obtained:

child： $Child=parent_1$

$$_{Mutate\_site-1}+/parent_{mutate\_site}+parent_{mutate\_site+1\sim sz} \tag{13}$$

mutate site --- aberrant position

7) Repeat steps 3)-6) until the optimal solutions are obtained to satisfy precision requirement or largest aberrant algebra.

4. SIMULATION OF NUMERICAL VALUE

Give an example for the simulation: the optimal control of sewage treatment process with an aeration basin whose available capacity is $2500m^3$.

Let water quantity of inflow *Q* and density of inflow substrate (BOD) change as sine wave following time everyday:

$$\begin{aligned} Q &= Q(t) = \overline{Q} + 5000\sin(2\pi t) \\ S_o &= S_o(t) = \overline{S}_o + 0.05\sin(2\pi t) \end{aligned} \tag{14}$$

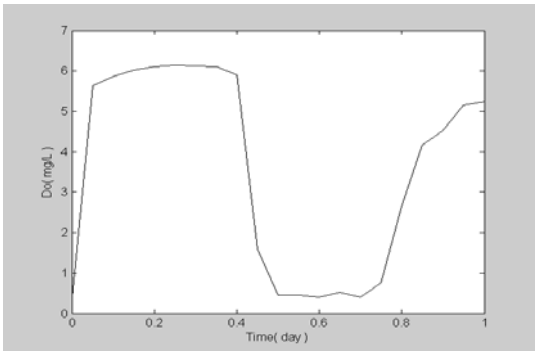
Where $\overline{Q}$ --- Average value of inflow quantity, let $\overline{Q}$ =10000 m^3/d.

$\overline{S}_o$ --- Average value of substrate density of inflow, let

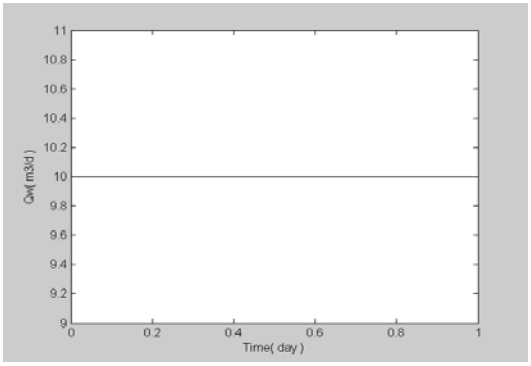
$$S_0 = 0.150kg/m^3$$

Let colony size *N*=100, aberrant probability be 0.05, crossover probability be 0.8, largest aberrant algebra be 40, X_0=2000.0*mg/L*， S_0=9.96 *mg/L*. The result of simulation is shown in Fig. 2. in which Fig. 2. (a), (b) are optimal curves of dissolved oxygen density *DO* and sludge discharge flow Q_w separately, and Fig. 2. (c), (d) are changing curves of *MLSS* density and substrate density *X* of aeration basin *S*

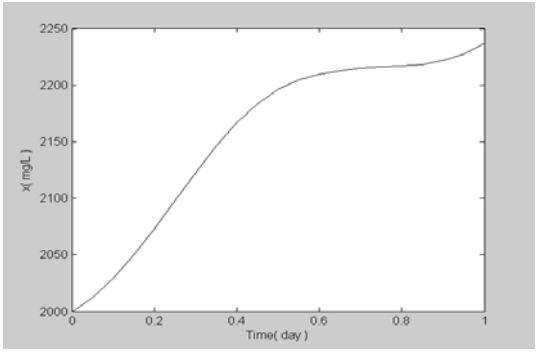
following optimal control. Consequently, *DO* have to be controlled between 0.5~6.08 *mg/L*, so that the substrate discharged everyday can be controlled below 150*kg* (BOD)/d and the lowest operation cost of sewage treatment can be realized.



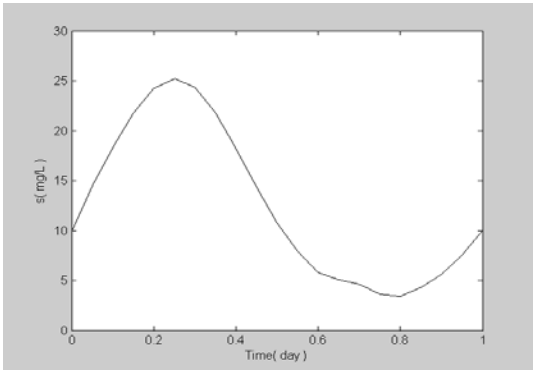
(a) Optimal curves of *DO* density



(b) Sludge discharge flow Q_w



(c) Changing curves of *MLSS* density



(d) Substrate density *X*

Fig. 2. Curves of simulation result

5. CONCLUSIONS

Iterating initial value needs to be estimated for optimal control by traditional gradient method. When step width parameter selected is small, calculation time is long and it is searching optimization locally. By contrast, when step width parameter selected is big, iteration is divergent and the solution cannot be obtained. Genetic algorithm is searching optimization of overall situation in the whole parameter space. It is only one fitness function that is needed, not the differential coefficient or other assistant information. The last simulation proves that it is feasible and effective to adopt genetic algorithms further.

In automatic control system of sewage treatment, the optimal values of Q_w and *DO* are calculated through genetic algorithms for real-time optimal control. If the sewage treatment plant is unable to realize optimal control, the optimal control can be imitated by manual or semiautomatic control based on the changing principles of optimal control variables Q_w and *DO* in the sewage treatment system. Meanwhile, the changing principles can be compared with the one of optimal curve *X* and *S* to regulate the control variables Q_w and *DO*. Consequently, the results of this research are very practical and realistic.

6. REFERENCES

[1] X.S. Xie, Optimal control theory and its application, Peking: Beijing Tsinghua University Press, 1986.

[2] E.Ayesa, B.Goya, et al, “Selection of operational strategies in activated sludge process based on optimization algorithm”. Water Science & Technology, 1998, 37(12): 327-334.

[3] Z.M. Yang, “Application of genetic algorithms in automatic control fields”, Information and control, 2002, 29 (4).

[4] Y.Z. Peng, B.Z. Wang, “Multivariable optimal control of activated sludge process I, influence to operation cost by basic theories and DO density”, Environmental science transaction, 1998, 18(1): 11-14.

[5] Y.Z. Peng, B.Z. Wang, “Multivariable optimal control of activated sludge process II, restrict the discharge gross of organic material”, Environmental science transaction, 1998, 18 (1): 20-23.

[6] X.P. Wang, L.M. Cao, Genetic algorithms—theory, application and software realization. Xian: Xian Jiaotong University Press, 1995.

[7] G.L. Chen, X.F. Wang, et al, Genetic algorithms and its application, Beijing: People’s Posts and Telecommunications Press, 1996.

The Obstacle Avoidance Application of the EAPF based on Consensus in the Robot soccer

Fei Yu, Dongmei Lv, Song Yuan, Ximei Liu
College of Automation and Electronic Engineering, Qingdao University of Science and Technology
Qingdao , Shandong Province, China
Email: hahaldm@163.com

ABSTRACT

This article introduces an approach of the obstacle avoidance based on consensus in the robot soccer and analysis the disadvantage of the traditional artifical potential field (APF). The agents can avoid the obstacles by introducing gyroscopic force to replace the repulsion force. We solve the navigation and obstacle avoidance of robot with the optimal potential function which is derived by combining EAPF with GA and satisfy the real-time optimization.

Keywords: EAPF based on Consensus, gyroscopic force, gradient function, optimal time.

1. INTRODUCTION

The Robot World Cup Initiative (RoboCup) is an international research initiative that uses the domain of robot soccer to foster AI and intelligent robotics research. The system is composed of vision subsystem, decision-making subsystem, and wireless communication subsystem and robot dolly system. It is a decision-making control system through closed loop vision. It captures the information of obstacle through the vision and generates the optimal path through the decision-making system. The path planning, a kind of nonlinear question is the key question in the navigation of the robot. The common path planning methods are as follows: gradient method, APF, A* search method, enumeration method, Monte Carlo analysis and so on. All of those methods are only based on the path minimum optimization and short of robust in optimization.

Collision avoidance plays an important role in the context of managing multiple vehicles. Many traditional control methods for collision avoidance rely on a potential-based approach, but there are many disadvantages .This article introduces a gyroscopic force to replace the repulsion force, so that the agents can avoid the obstacles. We solve the navigation and obstacle avoidance of robot with the optimal potential function which is derived by combining EAPF with GA and satisfying the real-time optimization.

2. EAPF

2.1the traditional APF

Collision avoidance plays an important role in the context of managing multiple vehicles. Many traditional control methods for collision avoidance rely on a potential-based approach, such as in the Navigation Function Method (NFM) or harmonic potential fields.The conception of APF was first put forward by Khatib[1,2,3] and applied in many aspects, then later many literatures introduce the APF and de some improvement on APF responding to many different questions. The ideas presented in the paper are inspired by the NFM. The general idea of the NFM is to create a global potential field to accomplish some control objective, such as getting a vehicle to travel from its initial location to some target point while not colliding with any obstacles. To create this global potential field, an attracting potential might be placed at the target point while repelling potentials are placed at the locations of obstacles to push an approaching vehicle away from the obstacle. Then the vehicle navigates using the gradient of the potential field as a force field.

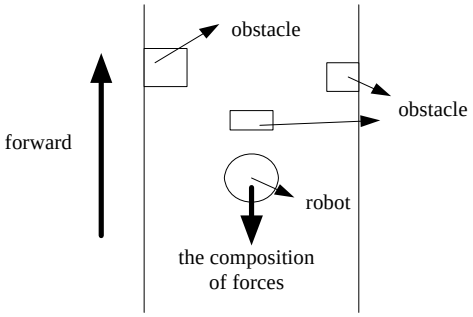


Fig. 1. The sketch map of the robot's force

The breakthrough of the NFM was that it could be used to show the existence of trajectories that avoid collision with any obstacles. However, this method has a few drawbacks: i) global information is needed regarding the location and shape of all the existing obstacles, ii) corresponding to any obstacle, there exists a neighborhood that can trap the vehicle for relatively long time, iii) the NFM is often computationally impractical, iv) the original NFM only considers the case of a single vehicle.

Instead of relying on repelling potentials for obstacle avoidance, as in the NFM, the control law we present relies on gyroscopic forces.

2.2the Improvement of EAPF

Gyroscopic forces can be thought of as steering forces since they always act perpendicular to the direction of motion. Fortunately, gyroscopic forces can be used for obstacle avoidance without affecting the global potential function that was constructed to foster some control objective. Actually, gyroscopic forces do not even change the energy of the system, a well -known and easily verified fact.

It turns out that gyroscopic forces alone have some difficulty preventing collision in large groups of vehicles. Therefore we also introduce a type of braking force that allows the vehicles to slow down when they are getting too close to one another or an obstacle. Intuitively, if the vehicle is moving too fast towards an obstacle it will not have enough time to turn to avoid the obstacle. Therefore the braking force is used to slow the vehicle so that it can turn in time to avoid the obstacle. With the gyroscopic force, the

braking force still does not change the global potential function.

Since the control of each vehicle is localized, the computations can as well be localized, which is very important to ensure scalability of our control law to groups that contain a large number of agents. The methods that we present are equally applicable to either 2- or 3-dimensional motion.

(1) Consensus Problems in Network Systems [4, 5]

Let $G = (\Gamma, E)$be a graph with a nonnegative adjacency matrix A=[$a_{ij}(t)$]that specifies the interconnection topology of a network of dynamic systems, sensors, or agents. The set of nodes is denoted by Γ = {1, 2. . . , n}. For complex networks, we refer to $|\Gamma|$ and |E| as the scale and size of the network, respectively.

A consensus algorithm can be expressed in the form of a linear system

$$\dot{x}_i = u_i(x_i, x^{(i)}) \qquad \forall i \in \Gamma \tag{1}$$

$$u_i(x_i, x^{(i)}) = \frac{1}{dg(x_i)/dx_i} \sum_{j \in N_i} \phi(x_j, x_i), \quad \text{for all } i \in \Gamma \tag{2}$$

Note1: (1) converge to an aligned state $x^* = (\mu....\mu)^T$ with identical elements equal to

$$\mu = \bar{x}(0) = \frac{1}{n}\sum_i c_i \tag{3}$$

This explains why in the term "average-consensus" was first coined to refer to the distributed algorithm [6] in (1). In a more compact form, system (1) can be expressed as

$$\dot{x}(t) = -L(t)x(t) \tag{4}$$

Where *L* is the Laplacian matrix of graph G and is defined as

$$L = \Delta - A \tag{5}$$

Where $\Delta = diag(A \bullet 1)$ is the *degree matrix* of G with diagonal elements $d_i = \sum_j a_{ij}$ here is, $1 = (1.....1)^T$ denotes the vector of ones that is always a right eigenvector of L corresponding to $\lambda_1 = 0$. If G is connected, then λ_2 >0. Apparently, the analysis of consensus problems in networks reduces to spectral analysis of Laplacian of the network topology. Particularly λ_2 is the measure of speed of convergence of the consensus algorithm. λ_2 is named the algebraic connectivity of the graph by Fiedler due to the following inequality:

$$\lambda_2(G) \le v(G) \le \eta(G) \tag{6}$$

Where $v(G)\eta(G)$ are node-connectivity and edge connectivity of graph respectively .According to this inequality, *a* network with a relatively high algebraic connectivity is necessarily robust to *b*oth node failures and edge-failures. A lower bound on this degree of robustness is [λ_2].

(2) The obstacle avoidance based on consensus

Suppose we have a group of fully actuated vehicles obeying second-order translational dynamics. Since each vehicle will implement the same control law, we need only to develop the control for one vehicle. We desire a feedback control law to drive the vehicle to a target point with x_T out colliding with any obstacles or other vehicles. A detection shell, a ball of radius, is $r_{\det}$ .given to the vehicle so that the vehicle can respond to any obstacle within this shell. For the purpose of designing the control law, let us refer to an obstacle of vehicle *i* as being *either* an external object that vehicle *i* is trying to avoid *or* a neighboring vehicle within the detection shell of vehicle *i*. In addition, we assume all obstacles are convex. The dynamics of the vehicle are given simply by

$$\ddot{x} = u, x \in R^3 \tag{7}$$

The control *u* consists of three parts as follows:

$$u = F_p + F_d + F_g \tag{8}$$

Where F_p is a potential force which assigns to the vehicle a potential function with the minimum at the target x_T ; F_d is a damping force; F_g is a gyroscopic force. If the system matrix is time-independent and symmetric, $A(t) = A = A'$ for all t, it is convenient to rewrite (1) as

$$\dot{x} = -\nabla V(x) \tag{9}$$

Where $\nabla V(x)$ denotes the gradient of V at x: $\nabla V(x) = (\frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial x_2}.....\frac{\partial V}{\partial x_n})'$. It is clear that the equilibrium points correspond to stationary points of V and that away from these points the potential is strictly decreasing with time.

$$\begin{aligned} F_p &= -\nabla V(x) \\ F_d &= -D(n)x \\ F_g &= s(n, \dot{x})\dot{x} \end{aligned} \tag{10}$$

where n denotes the vector from the vehicle to its nearest obstacle, *V* is a (potential) function, the matrix *D* is symmetric and positive-definite, and the matrix *S* is skew-symmetric.

One suitable choice for the potential function is a simple quadratic $V(x) = \frac{1}{2}\|x - x_T\|^2$ .The matrix *S* in the gyroscopic force term $F_g = s(n, \dot{x})\dot{x}$ is chosen to be an infinitesimal rotation. Since the matrix *S* is skew- symmetric, $< s(n, \dot{x})\dot{x}, \dot{x} >= 0 \Rightarrow F_g \bullet \dot{x} = 0$ .Therefore F_g does not do any work.

The damping term, F_d in (8) can be thought of as having two positive definite components:

$$F_d = -D(n)x = -(D_{bc} + D_b(n))\dot{x} \tag{11}$$

where D_{bc} is a constant matrix that represents any natural dissipation, such as drag, and $D_b(n)$ is an imposed *braking* term that varies with the relative distance between the vehicle and its nearest obstacle. It is this braking force $D_b(n)\dot{x}$ along with the gyroscopic force that the vehicle uses for collision avoidance. As previously mentioned, each

vehicle is given a detection shell of radius r_{dec} within which it can determine the relative location between itself and its nearest obstacle [7]. Even if more than one obstacle is within the vehicle's detection shell, the vehicle only reacts to the nearest one. If there is no obstacle within the vehicle's detection shell then the gyroscopic and braking forces are zero. In addition, each vehicle does not react to obstacles "behind" them, even if the obstacle is within the detection shell. The magnitude of the gyroscopic and braking forces varies as a negative exponential of the distance between the vehicle and its nearest obstacle, for example:

$$D_b(n) = -c_1 \exp\left(-\|n\|\right) - c_2 \tag{12}$$

What is promising about using gyroscopic forces is that we have the ability to decouple, in a sense, the collision avoidance from the control objective. For instance, if we can prove the vehicle does not collide with any obstacles, then we can easily prove convergence of the vehicle to its target point using the energy $E(x,\dot{x}) = \frac{1}{2}\|\dot{x}\|^2 + V(x)$ as a Lyapunov function since gyroscopic forces do not change the energy.

Proof: time varying edge costs

$$c_{ij} = \alpha\left(g(x_j) - g(x_i)\right)\hat{\phi}\left(\vartheta(x_j) - \vartheta(x_i)\right) \quad \forall (i,j) \in E \tag{13}$$

Subset $\Lambda \in \varepsilon$ of edgsets defining trees$_{(\Gamma,Q)}$. For each $Q \in \Lambda$ Defining continuous and positive definite function

$$W_Q(\eta) = \sum_{(i,j)\in Q} c_{ij} \tag{14}$$

$$\eta_i = g(x_i) - g\left(\hat{x}(x(0))\right) \tag{15}$$

$\eta = 0$ Corresponds to $x = \hat{x}(x(0))1$, Lyapunov candidate function:

$$E(\eta,\dot{\eta}) = \frac{1}{2}\|\dot{\eta}\|^2 + V(\eta) \tag{16}$$

Find minimum spanning tree:

$$W(\eta) = \min_{Q\in\Lambda}\{W_Q(\eta)\} \tag{17}$$

So $W(\eta)$, continuous, positive definite and

$$\dot{V}(\eta) \le -W(\eta), \quad \forall \eta \in IR^n, E \in \varepsilon \tag{18}$$

3. The Application of Time-optimization Path Planning

3.1 The Fundation of the Fitness Function

Because every operation of GA does some genetic operation and successive generations according to the information of the fitness, the choice of the fitness function will has effect on the convergence speed directly. Moreover, the choice of the fitness function should consider the character of some specific questions, namely requiring time-optimization and obstacle avoidance.

It can be found from (19)、(20) and (21) that path minimum and velocity maximum are needed in order to realize time optimization ,.

$$t = \int_{initpos}^{target} \frac{ds}{v} \tag{19}$$

$$ds = \sqrt{X'^2(u) + Y'^2(u)}\ du \tag{20}$$

$$t = \int_a^b \frac{\sqrt{X'^2(u) + Y'^2(u)}}{V(u)} du \tag{21}$$

3.2 The Fitness Function of the Path Smooth

For driving with high speed, smooth path is necessary. The path smoothness is presented by the curvature *k.* When dealing with spline curves in two dimensions k-is given as follows:

$$K(u) = \frac{X'(u)Y''(u) - Y'(u)X''(u)}{[X'(u)^2 + Y'(u)^2]^{\frac{3}{2}}} \tag{22}$$

3.3 The Fitness Function in the limitation of the Acceleration

In order to make the robot not to slide, his accelerations must be within limits. The model of the robot as follows:

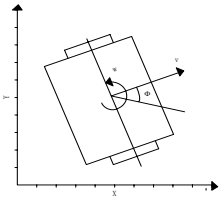


Fig. 2. The Model of robot

The overall acceleration can be decomposed to tangential acceleration and radial acceleration. The tangential acceleration is the derivative of velocity with respect to time and is caused with desire to increase or decrease speed

$$a_{\tan g} = \frac{dv}{dt} \tag{23}$$

The radial acceleration is caused by turning at certain speed and is the product of linear and angular velocity

$$a_{rad} = v \times w \tag{24}$$

Orientation: $w = \dot{\Phi}$. (25)

Since tangential and radial acceleration are orthogonal, the overall acceleration is the Pythagoras sum as follows:

$$a_{总} = \sqrt{a_{\tan g}^2 + a_{rad}^2} \tag{26}$$

4. RESULT ANALYSIS

By introducing gyroscopic force to replace the repulsion force, the agents can avoid the obstacles and overcome the disadvantage of the traditional APF. We solve the navigation and obstacle avoidance of robot with the optimal potential function which is derived by combining EAPF with GA and satisfying the real-time optimization which can be seen from the following:

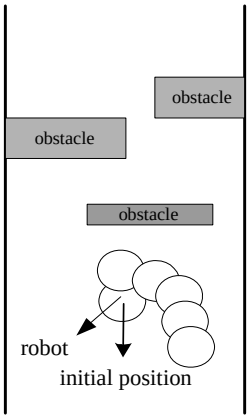


Fig. 3. The traditional sketch map

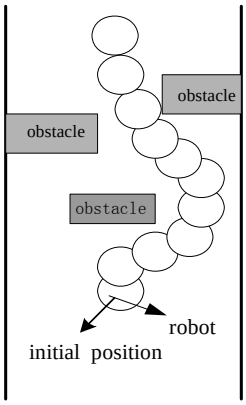


Fig. 4. The evolutional sketch map

5. CONCLUSION

Collision avoidance plays an important role in the context of managing multiple vehicles. The Robot World Cup Initiative is an international research initiative that uses the domain of robot soccer to foster AI and intelligent robotics research. Many traditional control methods for collision avoidance rely on a potential-based approach, but there are many disadvantages .This article introduces a gyroscopic force to replace the repulsion force, so that the agents can avoid the obstacles. We solve the navigation and obstacle avoidance of robot with the optimal potential function which is derived by combining EAPF with GA and satisfying the real-time optimization.

6. ACKNOWLEGMENT

This work was supported in part by the Natural Science Foundation of Qingdao 05-1-jc-88 and Subitem of the tenth five years' national science & technoloy key project J2001-1884-XJ.

7. REFERENCES

[1] O. Khatib, “Real-time obstacle avoidance for manipulators and mobile robots”. In Proceedings of the IEEE International Conference on Robotics and Automation, St.Louis, 1985, 500-505.

[2] O. Khatib, Real-time obstacle avoidance for manipulators and mobile robots. The International Journal of Robotics Research, 1986, 5(1):90-98.

[3] O. Khatib, A unified approach to motion and force control of robot manipulators: The operation space formulation. IEEE Journal of Robotics and Automation, 1987, 3(1):43-53.

[4] W. Ren, R.W. Beard, “Consensus seeking in multi-agent systems under dynamically changing interaction topologies,” IEEE Transactions on Automatic Contro*l*, May 2005.

[5] R. Olfati-Saber, R. Murray, Consensus problems in networks of agents with switching topology and time-delays, IEEE Transactions on Automatic Contro*l*, 2004, 49(9), 1520–1533.

[6] L. Xiao, S. Boyd, Fast linear iterations for distributed averaging, Systems and Control Letters 53 (1) (2004) 65–78.

[7] A. Jadbabaie, J. Lin, et al, Coordination of Groups of mobile autonomous agents using nearest neighbor rules, IEEE Transactions on Automatic Contro*l* , 2003,48 (6):988–1001.

The Data Warehouse Technology of DSS of Three Gorge Safe Monitor and Control System

Jing Li
School of Computer Science, Wuhan University
Wuhan, Hubei，430079, China
Email: leejingcn@msn.com

ABSTRACT

Data warehouse has been developed for the need of the increasing data to integrate applications at the data level. One key area involved in data warehouse is to transfer the source data into the data warehouse and to make it available to the user. A new data model, called multidimensional data model, is proposed in this paper. This model is successful used in DSS of the Three Gorge safe monitor and control system. The essential concepts, the key technology and the system structure about this new model are presented.

Keywords: Data Warehouse, Meta Data, MOLAP**,** OLAP, Decision Support.

.

1. INTRODUCTION

The three Gorge Safe Monitor and Control System (TGSMCS) has thousands of sensors to collect data [1]. As the quantity of data becomes numerous, the computer memory will be used up rapidly and be hard to search, therefore the data analysis can not be done easily. Moreover, if the previous data is restored regularly, it is can not used to analyze the project during long time period. Therefore, the safe situation of the Three Gorge's buildings can not be evaluated [2]. The use of the data warehouse has increased significantly in recent years and now plays an important role in many decision support processes. Different from the usual data warehouse structure, the Three Gorge Safe Monitor and Control's data warehouse is based on relation database, paralleling process and distribution techniques. The problems invoked by the numerous data are solved. And also it is convenient to use the valuable information.

A lot of research efforts have been made to integrate heterogeneous information sources in a distributed environment. The object-oriented approach is seen to be as a good solution for the heterogeneous system. An object framework for data warehousing is proposed by Qing Li [3]. A new structure of data warehouse based on data marts is proposed by Jeff Lawyer and Shamsul Chowdhury [4]. A new method called DWS-AQA (Data Warehouse Striping with Approximate Query Answering) is proposed by Jorge Bernardino, Pedro Furtado, and Henrique Madeira [5]. The goal is to use the processing and disk capacity normally available in large workstation networks to implement a data warehouse with little infrastructure cost.

The paper is organized as follows. In Section 2, the data warehouse technology is briefly discussed. Section 3 introduces the structure of data warehouse in DSS of TGSMCS. Section 4 proposes the organization of multi-dimension data. Section 5 gives some conclusions.

2. DATA WAREHOUSE TECHNOLOGY

2.1 Definition

In his book "construct data warehouse", W.H.Inmon, who is the founder of data warehouse, gives the definition of the data warehouse: "the data warehouse is orient to subject, integrative , standard and time-different data collection to support manage administration in DSS."[6,7]. This new type data is different from the traditional database objects. In the past tens of years, the database technology, especially OLAP server for automatic production, has decreased the tasks and speeded the data's collection[8]. The OLAP server is transaction-drive, application-orient and can save the records done by current transaction quickly and safely. As the society develops, the simple data operation does not satisfy ordinary use. And also, in order to offer support for decision-making, we may need the data for ourselves to analysis and to inference, which leads to the appearance of DSS. Initially the construction of DSS is based on the database. As the database technology grows up rapidly, particularly the fast develop of the relation database, the total performance of the database improves quickly. According to the traditional database, many DSS is set up, which play an important role in many applications.

2.2 Key Technology of Data Warehouse

Technically, the data warehouse can be divided into three parts: obtaining data, storing and managing data and representing data. The three parts of the data warehouse are presented as follows.

1) Extracting data

The Extracting data is the entrance to the warehouse. As the data warehouse is an independent data situation, it gets data from Online Transaction Processing System, external data source, offline data-storing medium. The obtaining data technically refers to interlink age, copy, increment, conversion, and attempter, monitor and so on. The data in the data warehouse don't need online transaction processing system to keep real time synchronization. So the data-extracting can be timing. But the running time, the mutual order among the multi-extracting is significant to validate the information in the data warehouse.

2) Storing and managing

The key of the data warehouse is to store and to manage the data. Different from the traditional database, the manage mode in the data warehouse decides its action form in response for outside. What products and technique should be taken to construct core of data warehouse, technique characteristics of data warehouse should be first analyzed.

The first problem of the data warehouse is how to store and to manage the mass data since the data involved are much more than that of the traditional transaction processing. These data will be cumulated rapidly as time passed by. Among all the existing technologies and products, only the RDBS can be used to solve this problem. After nearly 30 years of development, the RDBS is much superior to the other DBS in storing and managing data. Based on the data fragment technology, the RDBS can store a huge DB into a number of memory equipments and enhance their expansibility of managing mass data more and more. It's very ordinary to use RDBS to manage data with hundreds of GB. Moreover, it exists some manufacturers specializing in the mass data backup. So the data warehouse does not require a high standard of online backup.

The next problem of the data warehouse to solve is the parallel processing. In the traditional OLTP (Online Transaction Processing), the users visit the system shortly and intensively. For a multi-processor system, it is important to make a balanced distribution among the users requirement, which is the so-called concurrent operation. But in the data warehouse system, the users visit the system vastly and sparsely. When the visit frequency is not very high, the system will have the ability to schedule processors to serve the complex query request and the parallel process. Consequently the parallel process technology in the data warehousing is more important than ever.

The third problem of the data warehouse is to optimize the decision-support queries. This problem is special for RDBS, as the other data management environments do not have perfect capability of implementing basic generic enquiries. Technically, the optimization of the decision-support queries includes the DBS indexing mechanism, query optimizing devices, linking strategy, data classification and sampling and many other parts.

3) Data Representation

The data representation is a window of the data warehousing, which concentrates on the issues of multidimensional analysis, of numerical statistics and of data mining. The multidimensional analysis is an important manifestation of the data warehousing. As the MOLAP system is appropriative, the most common tools and products in the multidimensional analysis field are the ROLAP tools, which recently pay more attentions to providing front online analysis interface based on Web, rather than just promulgating data on the internet.

3. DESIGN DATA WAREHOUSE OF DSS OF TGSMCS

3.1 Data Organization

The data in a safety monitored Data Warehouse is first gained from an existing monitoring DB. Then they are organized into different levels in type of the basic data or into the different level integrated data according to the different requirements. The integrated data includes the recent basic data, the forward basic data, the mildly integrated data and the highly integrated data. Among these data, the recent basic data, which means recently received source data, should be paid more attention and be used timely and consequently. As the time passes by, it would be transformed into the forward basic data by the time control mechanism of the data warehouse, or stored into some storage medium such as CDs, magic tapes and so on. The mildly integrated data is periodically received from the recent basic data and can be classified into the data of a month, a week, a day, typically 2 hours, 8 hours, 16 hours, 20 hours, or directly every hour, according to the data gaining period control; The mildly integrated data is also can be classified into △1、△2、△3、△4、△5 according to the increment size control. The highly integrated data is on the top level where the data is refined to be quasi-decision data.

3.2 The Data Warehouse Structure of DSS of TGSMCS

The structure of safe monitoring data warehouse of DSS of the Three Gorges Projects is shown in fig. 1. It includes data source components, data warehouse management components, and warehouse storage and analysis tools. The data source, which is collected from the Safe Monitor Relation Database of Three Gorge, transfers the basic data to the data warehouse. The interesting data in the data source are queried by the management components and are then computed, stored and updated in the warehouse.

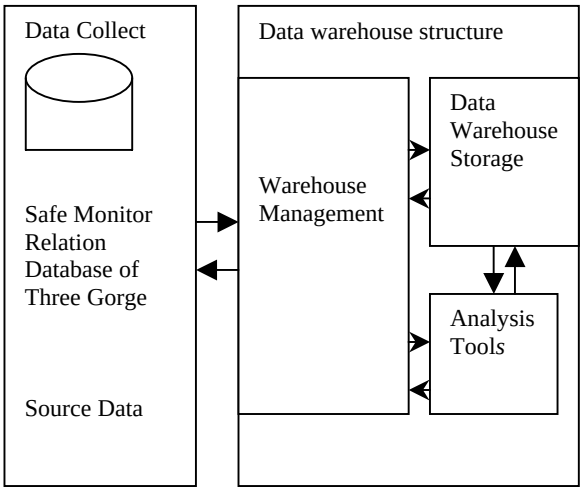


Fig. 1. The structure of data warehouse system

The structure of warehouse management is shown in Fig. 2.

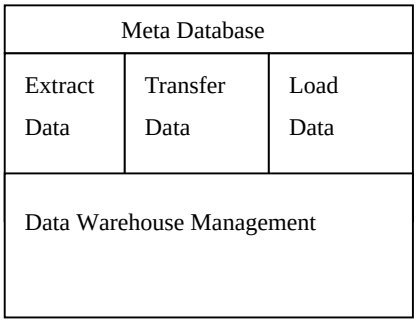


Fig. 2. Warehouse Management Structure

The Data warehouse storage and analysis tools are shown in Fig. 3.

The source data is the data collected automatically or collected manually by monitoring system in the original relational database. The system establishes layered analytic the entity of data warehouse outside the source

database, but the communication between it and source database in data and structure are closed.

The warehouse management first extracts, transfers and loads the data, then defines the process of data integration and transformation, and finally transfers data from the different level data sources into the data warehouse automatically and gives support to DSS. It is used to store data model, to define data structure, to manage data security, maintenance, backup, restoration and so on by metadata.

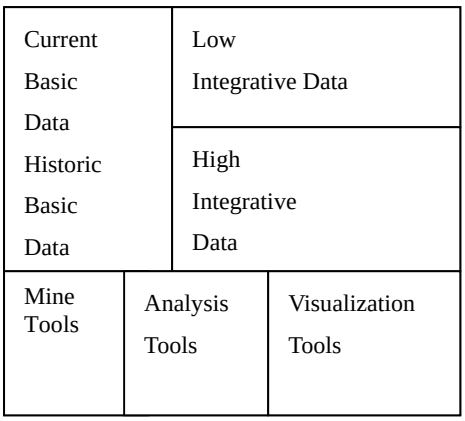


Fig. 3. Data warehouse storage and analysis tools

The data analysis tools include data mining tools, models for data analysis and visualization tools. The data mining tools mainly are used to discovery knowledge, rules and recent data relationships. The data analysis tools are formed by the current basic data layers from the sources according to the requirements of decision-making, and then formed by a low-grade and a highly integrative decision-making data according to the requirements of integrative decision-making. Since the monitoring data has some special characteristics, the monitoring data analysis uses many general models and some professional models, and organizes them by using the models.

The environment of the data warehouse tools is the most important part of the data warehouse. It permits the users to obtain the effective data and the application development and to establish quickly a decision support data warehouse application interface and the application software. The analysis methods include the statistical analysis, the time series analysis, several variable state spaces of index form flat to complex, the spectrum analysis, and so on. In addition, it includes the guide data analysis and the graphs data analysis, as well as the interactive programming language. The visualization tools are used to analyze the data and to show the results.

The data stored in the data warehouses are classified into the current data, the historical data, and the synthesis and analysis data in the different levels of the three parts. These data are formed the multidimensional data structure by decision-making of each parts.

3.3 Data Warehouse's Metadata

The metadata is a describing data. Moreover, it can also describe the environment of the data warehouse and functions as a procedure groupware. So it is used as the data dictionary of the database system. It includes two types: management metadata and user metadata. The former describes the subject of the data warehouse, the foundation data, the data transform, and so on. The latter can help user to know the data warehouse organization, to inquire information and to understand results. The data warehouse metadata must have good item names and definition, as well as standardized technical names and database formats. The data columns in the data warehouse are defined as the "code" type columns. They have all potential values and their meanings either encoded in metadata extensions or, if many values exit, considered building special data warehouse code/decode tables. The five function of the data warehouse are supported by the metadata as follows: describe the data warehouse content and the data position; define the data extraction and the transform rule; extract the data attempt rule; describe the data synchronization demand; weigh the data quality index. The metadata is also used to amend the tracing data and the synchronization capture historical data.

In the data warehouse in the DSS of TGSMCS, the metadata management includes also the defining data operation and the metadata definition.

1) Define data operation

The defining data operation includes two operations: data extraction and data transform. The data extraction is to extract the data from the foundational database and the data transform to transform the data into the data warehouse as object data. The whole process is shown as follows:

Extraction: getting some useful data for decision making.

Filtration: getting rid of useless content.

Validation: validating data quality from decision support.

Amalgamation: between recent extracting data and Data warehouse.

Integration: integrating data so as to create summary data.

Loading: loading new data to data warehouse.

2) Metadata definition

Defining extraction function: it describes every extraction and identifies the foundational system, points out the extraction period and the new type after extraction including whole warehouse replacement, whole warehouse addition and updating addition.

Defining extraction steps: it uses Job Control Language(JCL) to show filtration, validation and so on.

Defining extraction table mapping: it enables to map input file or database table to output file or database table. Two kind of property are proposed by the definition.

Defining extraction domain mapping: it builds relation between input domain and output domain for every extraction step.

Defining record filtration rules: it builds one filtration mechanism, using record filtration in the step of data extraction; Metadata is able to create source data so as to do data transform, etc.

The function of metadata is that source data is transformed data warehouse's data.

4. ORGANIZATION OF MULTIDIMENSION DATA

The data in the data warehouse must be organized into the multidimensional view of the data, of which some parts is the measured value of numbers. These numbers are related to the context of these data dimension, such as the measured time and the measured place. The dimension or the data

multi dimension can do the set operations (such as sorting, comparing, and the analyzing of the trend of time, projection, selection). The storage of the multidimensional database is stored in the multi-array of structural files and the database can establish the dimension index and the data management files with relative data.

4.1 Model of the Multidimensional Data

The multidimensional data model transfers the structures information into facts and dimensions. Once a TGSMCS item is defined, it can be described by a set of attributes called measures or fact attributes. The attributes are contained in cells or points with the data cube. These sets of measures are based on a set of dimension deriving from the granularity to represent the facts. These dimensions thus present the context for analyzing the facts. Further, the dimension attributes characterize the dimensions.

4.2 Analysis of the Multidimensional Data

The analysis of the multidimensional data can share the multi dimensional information and analyze them quickly. It has three characteristics: 1) Quick. It use specific format of data storage and lots of prepared operations which make analysis to achieve quickly; 2) Analyzable. It deals with all the logical and statistical analysis related to application, and uses OLAP technology to analyze data. And also it can use external analysis tools (such as the time sequence analysis tool, correlative analysis tool, data mining tool, the accident warning and so on); 3) Multidimensional. It is the key attribute of OLAP, which can provide the multi dimensional view and the multi dimensional analysis processing for the data analysis. And it supports the layer dimension and the multi layer dimensional analysis.

Multidimensional data analysis is that data in the system mainly realize multidimensional analysis. When the data is analyzed, some methods are used as follows: 1) 2-dimension cutting. That's to say to organize the data dimension according to the layers used to analyze data statically; 2) 3-dimension cutting. It analyzes data in the forms of cubic, cuts the carry through locally and analyzes relative statically; 3) Upward survey. It can realize the statistics of macro concept(such as month-year-years) and regularity data analysis in the long term; 4) Downward careful analysis, such as month-day-time, achieves viewing clearing and analyzing dynamically; 5) Rotating form changes the data indexing and the retrieval form of data (dimension viewing, analysis) in order to realize relative analysis and impact analysis.

4.3 The Multidimensional Structure of OLAP

The multidimensional structure of the OLAP describes the super-cubic structure as an object with triplex dimension or more dimensions. The measured value of data occurs at the cross point and the every part of data spaces has the same dimensional attributes. If the data is applied to the multidimensional database in this system, it can make the operations easy. If the meter structure is contracted, the density of data will increase. So large data structure can be detached into many multidimensional structures and then the dimension is detached into a cubic structure in terms of some application. Thus we can store the sparsely matrix effectively and reduce calculations. In synthesizing analysis, large multidimensional structure can be reorganized.

4.4 Multidimensional Database Storage

Unlike the relational database, which stores the data in the forms of records, the multidimensional database store the data in an n-dimensional array. Lots of sparsely matrixes exist in the system. The data can be viewed as time dimension. Thus, the speed of processing data has been increased and the response time has been reduced. Finally the enquired efficiency has been improved.

5. CONCLUSIONS

The Three Gorges Project is not only large in scale but also a long-time project as long as 17 years. The buildings have been constructed respectively in Phase I, Phase II and Phase III. The secure monitoring system of the Three Gorges Project is a distributed system with measuring point distributed widely and numerous monitoring objects. The application of the data warehouse improves the performance of the monitoring system of the Three Gorges Project and makes it convenient to monitor, analyze and evaluate TGSMCS.

6. REFERENCES

[1] Z.Z. Wu, B.F. Zhu, Three Gorges Secure Monitoring of Hydraulic Structure and Feedback Design, China waterpower press, China, 1999, pp.30-41.

[2] J. Li, H.M. He, "Design of KDD of Three Gorges secure monitoring Information System", Computer Engineering,china,Nov. 2001, Vol.27(11): 150-154.

[3] Z.Y. Peng, Q. Li, et al, "Using object deputy model to prepare for data warehousing. Knowledge and Data Engineering", sept. 2005, IEEE Transactions on volume 17(9), pp.1274-1278.

[4] J. Lawyer, S. Chowdhury, "Best practices in data warehousing to support business initiatives and needs", System sciences, 2004, proceedings of the 37th Annual Hawaii International conference on 5-8 Jan., 2004, pp.1-9.

[5] J. Bernardino, P. Furtado, et al, "DWS-AQA: a cost effective approach for very large data warehouses", Database Engineering and Applications Symposium, 2002, Proceedings International 17-19, July 2002, pp.233-242.

[6] W.H. Inmon, Building the DataWarehouse[M], 2nd Edition, NewYork:Wiley,2003.

[7] C.S. Park, M.H. Kim, et al, "Finding and efficient rewriting of OLAP queries using materialized views in data warehouses", Decision Support Systems,March 2002, Vol.32(43),pp.79-399.

[8] A. Ammoura, O.Zaine, "Towards a Novel OLAP Interface for Distributed Data Warehouse", Proc. Of DAWAK 2001, Springer LNCS2114, Munich, Germany, Sept, 2001, pp.174-185.

Detectability of Nondeterministic Discrete Event Systems

Shaolong Shu[1], Feng Lin[1,2], and Hao Ying[2]
[1]School of Electronics and Information Engineering
Tongji University, Shanghai, China (E_mail:shushaolong@hotmail.com)
[2]Department of Electrical and Computer Engineering
Wayne State University, Detroit, MI 48202, USA

ABSTRACT

The detectability problem is to determine or estimate the current and subsequent states of a system based on a sequence of observation when the initial state of the system is unknown. We assume that the system is nondeterministic in the sense that an event occurrence may lead the system to more than one state. The observation includes partial event observation and/or partial state observation. We defined four types of detectabilities: strong detectability, (weak) detectability, strong periodic detectability, and (weak) periodic detectability. Then we derived necessary and sufficient conditions for these detectabilities with and without state observations. Constructing an observer, which models the estimation of states under different observation, can check these conditions.

Keywords: Discrete Event Systems, State Estimation, Detectability, No determinism.

1. INTRODUCTION

Discrete event system theory has investigated the problems such as controllability, observability, co-observability, normality and diagnosability [1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 13]. Recently, we studied the detectability problem of estimating or determining the current and subsequent states of the system based on observation [12]. The motivation is that in many applications, it is important to know the state of a system. For example, in medical applications, it is important to diagnose a disease and know disease stage of a patient. State estimation is also important in industrial applications. For these and other reasons, in [12], we investigated the state estimation problem. We say that a system is detectable if one can determine its state by observing its outputs. The assumption made in [12] is that the system is deterministic in the sense that if the system is in a particular state and an event occurs, then the next state of the system is unique. This assumption will be relaxed in this paper to cover nondeterministic situations.

In general, there are two types of outputs: event output and state output. For event output, we assume, as in supervisory control theory of discrete event systems that some events are observable and others are not. For state observation, we assume that there is a many-to-one output mapping from the state set to a state output set. In [12], four types of observations are studied: (1) all events are observable and no state is observable, (2) all events are observable and some states are observable, (3) some events are observable and no state is observable, (4) some events are observable and some states are observable. Obviously, Cases 1 and 2 are special cases of Cases 3 and 4. For nondeterministic systems, there is no reason to consider Cases 1 and 2. So we will consider only Cases 3 and 4. In Section 3, we will study the case when there is no state observation and in Section 4, we will study the case when some state observation is available. In both cases, we assume that either full event observation or partial event observation is available.

2. NONDETERMINISTIC DISCRETE EVENT SYSTEMS

We use an automaton (also called state machine or generator) to model a discrete event system [2, 8, 10]:

$$G = (Q, \Sigma, \delta),$$

where Q is the set of discrete states, Σ is the set of events, and δ is the transition function describing what event can occur at a state and the resulting new state. We assume that the system is nondeterministic, that is, the transition function is given by $\delta : Q \times \Sigma \to 2^Q$. Another equivalent way to define the transition function is to specify the set of all possible transitions: $\{(q, \sigma, q') : q' \in \delta(q, \sigma)\}$. With a slight abuse of notation, we will use δ to denote the set of all possible transitions as well and write $(q, \sigma, q') \in \delta$ if (q, σ, q') is defined.

In many applications of discrete event systems, it is desirable to estimate the current and subsequent states of the system based on observation of some events and/or some states. As is [12], the event observation is described by the projection:

$$P : \Sigma^* \to \Sigma_o^*,$$

where Σ_o is the set of observable events. The output map describes the state observation

$$h : Q \to Y,$$

where Y is a (finite) output set. The discrete event system with the event and state observation is described by

$$(G, P, h, \Sigma_o, Y)$$

where $G = (Q, \Sigma, \delta)$. The question is whether we can estimate the current and subsequent states based on the event and state observations.

We will make two assumptions about the system. (1) The system $G = (Q, \Sigma, \delta)$ is deadlock free, that is, for any state of the system, at least one event is defined at that state: $(\forall q \in Q)(\exists \sigma \in \Sigma)\delta(q, \sigma)$ is defined. (2) All loops in G contain at least one observable event. The second assumption ensures that the system will not look like in a deadlock. As discussed in [12], these two assumptions are not restrictive and can be relaxed at the cost of more complex notations and explanations.

3. DETECTABILITIES WITHOUT STATE OBSERVATION

In this section, we define detectabilities and derive conditions for checking detectabilities for systems where no state observation is available, that is, Y is the empty set The problem we want to solve is as follows.

State Estimation Problem 1
Given a nondeterministic discrete event system
$$(G, P, h, \Sigma_o, Y)$$
we do not know the initial state of $G = (Q, \Sigma, \delta)$. We have no state observation, that is, Y is the empty set. Can we determine the current and subsequent states of the system after a finite number of event observations?

Depending on whether we want to determine the current and subsequent states for all trajectories or just some trajectories, we define strong detectability and (weak) detectability as follows.

Strong Detectability
A nondeterministic discrete event system
$$(G, P, h, \Sigma_o, Y)$$
with $Y = \text{the empty set}$ is strongly detectable if we can determine the current state and the subsequent states of the system after finite number of event observations for all trajectories of the system.

(Weak) Detectability
A nondeterministic discrete event system
$$(G, P, h, \Sigma_o, Y)$$
with $Y = \text{the empty set}$ is detectable if we can determine the current state and the subsequent states of the system after finite number of event observations for some trajectories of the system.

For detectability and strong detectability, we require that after some initial period, we can determine the current and subsequent states of the system for all future times. Since the system is nondeterministic, this requirement may be too strict; because in a nondeterministic system, the next state is not unique. Therefore, sometimes it is more reasonable to require that the state of the system can be determined only periodically, which lead to the following two definitions.

Strong Periodic Detectability
A discrete event system
$$(G, P, h, \Sigma_o, Y)$$
with $Y = \text{the empty set}$ is strongly periodically detectable if we can periodically determine the current state of the system for all trajectories of the system.

(Weak) Periodic Detectability
A discrete event system
$$(G, P, h, \Sigma_o, Y)$$
with $Y = \text{the empty set}$ is periodically detectable if we can periodically determine the current state of the system for some trajectories of the system.

The procedures to check detectability, strong detectability, periodic detectability, and strong periodic detectability are similar for the first few steps, which are steps to construct an observer of the system.

Step 1, we add an initial state q_o to G. For all states $q \in Q$, we add ε-transition from q_o to q. For observable transitions $\sigma \in \Sigma_o$, we write the transition as (q', σ, q). For unobservable transitions $\sigma \notin \Sigma_o$, we re-label the transition as (q', ε, q). Formally,
$$G_{p,nd} = (Q \cup \{q_o\}, \Sigma_o \cup \{\varepsilon\}, \delta_{p,nd}, q_o)$$
where
$$\delta_{p,nd} = \{(q_o, \varepsilon, q) : q \in Q\}$$
$$\cup \{(q', \varepsilon, q) : (q', \sigma, q) \in \delta \wedge \sigma \notin \Sigma_o\}$$
$$\cup \{(q', \sigma, q) : (q', \sigma, q) \in \delta \wedge \sigma \in \Sigma_o\}.$$
In other words, $\delta_{p,nd}$ is a mapping
$$\delta_{p,nd} : (Q \cup \{q_o\}) \times (\Sigma_o \cup \{\varepsilon\}) \to 2^{Q \cup \{q_o\}},$$
which can be easily extended to
$$\delta_{p,nd} : (Q \cup \{q_o\}) \times (\Sigma_o \cup \{\varepsilon\})^* \to 2^{Q \cup \{q_o\}}.$$

Step 2, we convert the nondeterministic automaton $G_{p,nd}$ into a deterministic automaton $G_{p,obs}$ called observer [1]. Since $G_{p,nd}$ has unobservable transition (q', ε, q), we need first define the unobservable reach from a subset of states $x \subseteq Q \cup \{q_o\}$ as follows:
$$UR(x) = x \cup \{q \in Q \cup \{q_o\} : (\exists q' \in x) q \in \delta_{p,nd}(q', \varepsilon)\}.$$
Then we can define
$$G_{p,obs} = Ac(X, \Sigma_o, \xi_p, x_{p,o})$$
where
$$X = 2^{Q \cup \{q_o\}},$$
$$\xi_p(x, \sigma) = UR(\{q \in Q \cup \{q_o\} : (\exists q' \in x)(q', \sigma, q) \in \delta_{p,nd}\}),$$
and
$$x_{p,o} = UR(\{q_o, q_1, \cdots, q_n\}).$$

Step 3, we mark the states in $G_{p,obs}$ that contain singleton state and denote the set by $X_m = \{x \in X : |x| = 1\}$.

Step 4, we remove states in X_m from $G_{p,obs}$ and denote the automaton remained after removing the states in X_m as $G_{p,obs}^{rem}$:
$$G_{p,obs}^{rem} = Ac(X - X_m, \Sigma, \xi_p|_{X - X_m}, x_{p,o}).$$

The reason for constructing $G_{p,obs}$ is that it describes the estimation of possible states of the system as follows.

Lemma 1
If the current estimate of possible states of G is $x \in X$ (that is, $x \subseteq Q \cup \{q_o\}$) and an event $\sigma \in \Sigma_o$ is observed, then the next estimate of possible states of G is
$$x' = UR(\{q \in Q \cup \{q_o\} : (\exists q' \in x)(q', \sigma, q) \in \delta_{p,nd}\}).$$
Proof: The set of states $\{q \in Q \cup \{q_o\} : (\exists q' \in x)(q', \sigma, q) \in \delta_{p,nd}\}$ consists of all states that can be reached from a state in x right after event $\sigma \in \Sigma_o$ is observed. The unobservable reach of this set,
$$x' = UR(\{q \in Q \cup \{q_o\} : (\exists q' \in x)(q', \sigma, q) \in \delta_{p,nd}\})$$
then consists of all states possible before the next observable event is observed. Therefore, x' is the estimate of possible states of G after the observation of $\sigma \in \Sigma_o$.

Lemma 2
1. The initial estimate of possible states of G is
$$x_{p,o} = UR(\{q_o, q_1, \cdots, q_n\}).$$

2. The estimate of possible states of G after transitions $\sigma_1\sigma_2...\sigma_k$ are observed is $\xi_p(x_{p,o}, \sigma_1\sigma_2...\sigma_k)$.

Proof:

1. Initially, before any event is observed, the system can be in any of $\{q_o, q_1, \cdots, q_n\}$. Therefore, the initial estimate of possible states of G is $x_{p,o} = UR(\{q_o, q_1, \cdots, q_n\}) = \{q_o, q_1, \cdots, q_n\}$.
2. By the definition of $G_{p,nd}$ and repeated applications of Lemma 1.

After the construction of $G_{p,obs}$ and $G_{p,obs}^{rem}$, we can check the four types of detectabilities using the following criterions.

Criterion for Checking Strong Detectability

A discrete event system (G, P, h, Σ_o, Y) with $Y = \text{the empty set}$ is strongly detectable if and only if all loops in $G_{p,obs}$ are entirely within X_m.

Proof: By Lemma 2, the observer $G_{p,obs}$ describes the estimate of states the system may be in. When $G_{p,obs}$ enters states in X_m, we know exactly which state the system is in. By our assumptions, the observer $G_{p,obs}$ is deadlock free. Since $G_{p,obs}$ is finite, after some finite observations, $G_{p,obs}$ must enter some loops. If all loops in $G_{p,obs}$ are entirely within X_m, then the current state and the subsequent states of the system are known no matter which trajectory the system follows, that is, the system is strongly detectable. On the other hand, if there exists some loops in $G_{p,obs}$ that are not entirely within X_m, then the system can follow those loops and hence not strongly detectable.

Criterion for Checking Detectability

A discrete event system (G, P, h, Σ_o, Y) with $Y = \text{the empty set}$ is detectable if and only if there are loops in X_m.

Proof: If there are loops in X_m (and hence X_m is not empty) then such loops are accessible from the initial state in $G_{p,obs}$. Any such loop can produce at least one infinite string by which the system can always stay within X_m. Hence the system is detectable by Lemma 2. On the other hand, if there are no loops in X_m, the system will eventually leave X_m along any trajectory of the system. Hence the system is not detectable.

Criterion for Checking Strong Periodic Detectability

A discrete event system (G, P, h, Σ_o, Y) with $Y = \text{the empty set}$ is strongly periodically detectable if and only if there are no loops in $G_{p,obs}^{rem}$.

Proof: The condition of no loops in $G_{p,obs}^{rem}$ ensures that the system cannot always stay in $X - X_m$. Therefore, the system must visit X_m periodically. This implies that we can periodically determine the current state of the system for all trajectories of the system by Lemma 2. On the other hand, if there are loops in $G_{p,obs}^{rem}$, then the system may stay in such loops forever. Hence the system is not strongly periodically detectable.

Criterion for Checking Periodic Detectability

A discrete event system (G, P, h, Σ_o, Y) with $Y = \text{the empty set}$ is periodically detectable if and only if there is loops in $G_{p,obs}$ which include at least one state belonging to X_m.

Proof: If there is loops in $G_{p,obs}$ which include at least one state belonging to X_m, then the system can stay in this loop, and we can periodically determine the current state of the system for some trajectories of the system by Lemma 2. If no such loops exist, then the system is not periodically detectable.

4 DETECTABILITIES WITH STATE OBSERVATION

We now consider the case where some state observation is available in addition to partial event observation. Obviously, with additional information obtained from state observation, we can do better in estimating the state of the system. However, in terms of technical approach, the problem is more complicated and more difficult to solve. We need to introduce additional notations and the algorithm is also more complex. To start with, let us first formally define our estimation problem.

State Estimation Problem 2

Given a discrete event system

$$(G, P, h, \Sigma_o, Y)$$

we do not know the initial state of $G = (Q, \Sigma, \delta)$. We have some state observations, that is, Y is not the empty set. Can we determine the current and subsequent states of the system after a finite number of observations?

For this state estimation problem with state observation, the definitions of strong detectability, (weak) detectability, strong periodic detectability and (weak) periodic detectability are similar to those of systems without state observation.

Strong Detectability

A discrete event system

$$(G, P, h, \Sigma_o, Y)$$

with $Y \neq \text{the empty set}$ is strongly detectable if we can determine the current state and the subsequent states of the system after a finite number of observations for all trajectories of the system.

(Weak) Detectability

A discrete event system

$$(G, P, h, \Sigma_o, Y)$$

with $Y \neq \text{the empty set}$ is (weekly) detectable if we can determine the current state and the subsequent states of the system after a finite number of observations for some trajectories of the system.

Strong Periodic Detectability

A discrete event system

$$(G, P, h, \Sigma_o, Y)$$

with $Y \neq$ the empty set is strongly periodically detectable if we can periodically determine the current state of the system for all trajectories of the system.

(Weak) Periodic Detectability

A discrete event system

$$(G, P, h, \Sigma_o, Y)$$

with $Y \neq$ the empty set is periodically detectable if we can periodically determine the current state of the system for some trajectories of the system.

The algorithms to check the above detectabilities are more cumbersome since we need to consider state observation. They are described in the following steps.

Step 1, we add an initial state q_o to G. We extend the event set from Σ to $(\Sigma \cup \{\phi\}) \times Y \cup \{\varepsilon\}$. For all states $q \in Q$, we add transitions from q_o to q with label (ϕ, y) if $h(q) = y$. Re-labeling the original transitions (q', σ, q) in G as follows: For observable transitions $\sigma \in \Sigma_o$, we re-label (q', σ, q) as $(q', (\sigma, h(\sigma)), q)$. For unobservable transitions $\sigma \notin \Sigma_o$, we re-label (q', σ, q) as $(q', (\phi, h(\sigma)), q)$ if $h(q') \neq h(q)$ and as (q', ε, q) if $h(q') = h(q)$. The reason for the re-labeling of unobservable event is that if $h(q') \neq h(q)$, then we know that some event has occurred because the system has changed states, although we do not know which event has occurred; if $h(q') = h(q)$, then nothing will be observed. Formally,

$$G_{ps,nd} = (Q \cup \{q_o\}, (\Sigma_o \cup \{\phi\}) \times Y \cup \{\varepsilon\}, \delta_{ps,nd}, q_o)$$

Where

$$\begin{aligned}\delta_{ps,nd} = &\{(q_o, (\phi, h(q)), q) : q \in Q\} \\ &\cup \{(q', (\sigma, h(q)), q) : (q', \sigma, q) \in \delta \wedge \sigma \in \Sigma_o\} \\ &\cup \{(q', (\phi, h(q)), q) : (q', \sigma, q) \in \delta \wedge \sigma \notin \Sigma_o \wedge h(q') \neq h(q)\} \\ &\cup \{(q', \varepsilon, q) : (q', \sigma, q) \in \delta \wedge \sigma \notin \Sigma_o \wedge h(q') = h(q)\}.\end{aligned}$$

In other words, $\delta_{ps,nd}$ is a mapping $\delta_{ps,nd}$: $(Q \cup \{q_o\}) \times (((\Sigma_o \cup \{\phi\}) \times Y) \cup \{\varepsilon\}) \to 2^{Q \cup \{q_o\}}$, which can be easily extended to

$$\delta_{ps,nd} : (Q \cup \{q_o\}) \times (((\Sigma_o \cup \{\phi\}) \times Y) \cup \{\varepsilon\})^* \to 2^{Q \cup \{q_o\}}$$

.

Step 2, we convert the nondeterministic automaton $G_{ps,nd}$ into a deterministic automaton

$$G_{ps,obs} = Ac(X, (\Sigma_o \cup \{\phi\}) \times Y, \xi_{ps}, x_{ps,o}),$$

where $X = 2^{Q \cup \{q_o\}}$, $x_{ps,o} = \{q_o\}$, and

$\xi_{ps}(x, (e, y)) = UR($

$\{q \in Q \cup \{q_o\} : (\exists q' \in x)(q', (e, y), q) \in \delta_{ps,nd}\})$ with

$e \in \Sigma_o \cup \{\phi\}$.

The reason for constructing $G_{ps,obs}$ is same as before, that is, to describe the estimate of possible states of the system.

Step 3, we mark the states in $G_{ps,obs}$ that contain singleton state and denote the set by $X_m = \{x \in X : |x| = 1\}$.

Step 4, we remove states in X_m from $G_{ps,obs}$ and denote the automaton remained after removing the states in X_m as $G^{rem}_{ps,obs}$:

$$G^{rem}_{ps,obs} = Ac(X - X_m, \Sigma_o, \xi_{ps}|_{X - X_m}, x_{ps,o}).$$

We can now check the four detectabilities for systems with some state observation using the following criterions. Because the proofs of these criterions are similar to those in Section 3, they will be omitted.

Criterion for Checking Strong Detectability

A discrete event system (G, P, h, Σ_o, Y) with $Y \neq$ the empty set is strongly detectable if and only all loops in $G_{p,obs}$ are entirely within X_m.

Criterion for Checking Detectability

A discrete event system (G, P, h, Σ_o, Y) with $Y \neq$ the empty set is detectable if and only if there are loops in X_m.

Criterion for Checking Strongly Periodical Detectability

A discrete event system (G, P, h, Σ_o, Y) with $Y \neq$ the empty set is strongly periodically detectable if and only if there are no loops in $G^{rem}_{p,obs}$.

Criterion for Checking Periodic Detectability

A discrete event system (G, P, h, Σ_o, Y) with $Y \neq$ the empty set is periodically detectable if and only if there is loops in $G_{ps,obs}$ which include at least one state belonging to X_m.

5.CONCLUSION

In the paper, we investigated the detectability of nondeterministic discrete event systems with or without state observation and discussed the criteria for detestability and periodic detectability accordingly. We obtained necessary and sufficient conditions for these detectabilities.

6. ACKNOWLEDGEMENT

This research was supported in part by NSF under grant INT-0213651, and by NIH under grant 1 R21 EB001529-01A1.

7. REFERENCES

[1] C. Cao, F. Lin, et al, “Why event observation: Observability Revisited,” Discrete Event Dynamic Systems: Theory and Applications, 1997，7(2)：127-149.

[2] C.G. Cassandras, S. Lafortune, “Introduction to Discrete Event Systems,” Kluwer, 1999.

[3] R. Cieslak, C. Desclaux, et al, “Supervisory control of discrete-event processes with partial observations,” IEEE Transactions on Automatic Control, March 1988，Vol. 33, 249-260.

[4] M. Heymann, F. Lin, “On-line control of partially observed discrete event systems,” Discrete Event Dynamic Systems: Theory and Applications, 1994，4(3): 221-236.

[5] M. Heymann, F. Lin, “Discrete event control of nondeterministic systems,” IEEE Transactions on Automatic Control, 1998， 43(1)： 3-17.
[6] R. Kumar, M. Shayman, “Formulas relating controllability, observability, and co-observability,” Automatica, 1998，34（2）： 211-215.
[7] S. Lafortune, F. Lin, “On tolerable and desirable behaviors in supervisory control of discrete event systems,” Discrete Event Dynamics Systems: Theory and Applications, May 1991， 1（1）： 61-92.
[8] F. Lin, W.M. Wonham, “On observability of discrete event systems,” Information Sciences, 1988， 44(3): 173-198.
[9] P.J. Ramadge, “Observability of discrete event systems,” In Proceedings of CDC, December 1986.
[10] R.J. Ramadge, W. M. Wonham, “Supervisory control of a class of discrete event processes,” SIAM J. Control and Optimization, 1987，25(1): 206-230.
[11] P.J. Ramadge, W.M. Wonham, “The control of discrete event systems,” Proceedings of the IEEE, January, 1989，77（1）： 81-98.
[12] S. Shu, F. Lin, et al, “Detectability of Discrete Event Systems,” submitted to IEEE Transactions on Automatic Control.
[13] W.M. Wonham, P.J. Ramadge, “Modular supervisory control of discrete-event systems,” Mathematics of Control, Signals, and Systems, 1988，1（1）： 13-30.

Design Task Planning and Dynamic Scheduling to Concurrent Engineering

Gongliang Jiang
Chongqing Key Laboratory of Communications Engineering, Chongqing Jiaotong University
Chongqing, 400074, China
Email:jianggongliang@hotmail.com

ABSTRACT

In order to take full advantages of the resources and reduce design time, it was studied design task decomposition and task dynamic scheduling according to the characteristic of product design process on CE (Concurrent Engineering). This paper discusses the principle of task decomposition during product design process and the tree construction of task decomposition. The method of modeling for design task scheduling is investigated. The paper proposes dynamic scheduling algorithm that arranges the tasks and distributes resource according to the priority.

Keywords: task decomposition, dynamic scheduling, CE.

1. NTRODUCTION

In the era of economy globalization, with increasing market competition, shortening product design and development cycle becomes the key factor of enterprise living and developing. CE has considered the problems of products during their whole life cycle in the primary design. It can shorten product development cycle with concurrent product design and correlative processes. Product design to CE needs not only building a computer environment adapted to concurrent work style but also rebuilding product design process and managing design effectively. It builds a dispersed concurrent design platform by integrating the members of multi-object groups from different areas and subsystems, and optimizes its construction to improve efficiency and utilize resource rationally [1, 2]. So, design task should be planned rationally and scheduled efficiently during the process of product design. The paper emphasizes on studying design task decomposition and dynamic scheduling algorithm in order to save resource and shortens design time.

2. ECOMPOSITION OF DESIGN TASK

The members of multi-object group from different departments and areas always complete design of complex product. It's the basic of completing concurrent design that task is disassembled into relatively independent and suitable granularity subtask according to the group characteristics.

2.1 Decomposition rules of design task

Task decomposition is relevant to product construction characteristic, design process, design and manufacture resource of enterprise or dynamic enterprise union, quality of designers, and so on. It can be disassembled according to the follow rules [4].

Rule 1: solvability

Disassembled task should be solvable.

Rule 2: independence

Disassembled tasks should be independent as possible to reduce the dependent between each other, and can be implemented concurrently.

Rule 3: completeness

A subtask should remain completeness on product construction or function as possible.

Rule 4: compatibility of task granularity

The control of task granularity is relevant to its independence and solvability. In general, the larger the subtask granularity is, the stronger the independence is, but the weaker the solvability is, vice verse.

2.2 Decomposition tree of design task

Product design is always a very complex process. Design task decomposition is a process from above to below and forms task tree finally, whose root is design task and whose leaves are solvable subtask. Finite set TR is composed of several nodes whose root is the task to be solved. Other nodes except root node can be divided into several nonintersecting finite sets TR_1, TR_2 ···TRn. Each set is a task tree that is subtask tree of the root. Each node of the task tree can be considered as a new design task independently, and the whole tree is oriented set composed of the independent task decomposition processes in proper order. The largest number of the layers equals to decomposition times Max (Disassemble Num (p, t)). It can be disassembled continuous. If P_{ir} is valid and resolvable, subtask of P_{ir} can be resolved. Suppose P_{ir} is disassembled to P_{irs}, then Disassemble Num (P_{irs}, T) = 3. This process shouldn't finish until it can't be disassembled. The result of the decomposition shows as Fig. 1. Tree construction.

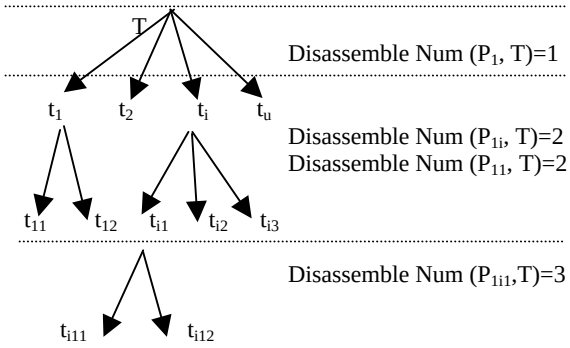


Fig. 1. Tree construction of task decomposition

Task tree is composed of several nodes in proper order whose basic unit is task node. Task node can be showed as Tree Node (Node ID, Parent Node ID, Child Node ID, and Attrib). Inserting and canceling tree nodes can complete task decomposition.

3. MODELING OF DESIGN TASK SCHEDULING

There are many relevant characterizes among the subtasks during the product design to CE, such as concurrent, intersect and reiteration. It shows as Fig. 2. It's the key to task scheduling that how to describe the interrelations

among the subtasks clearly [3].

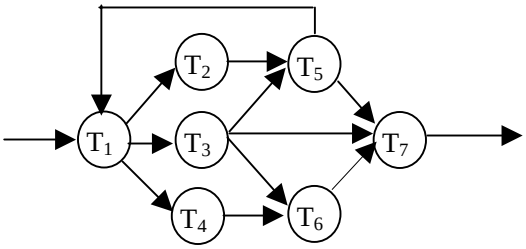


Fig. 2. Design flow chart and interrelations among subtasks

3.1 Task vector

Partial ordering may exist among the tasks, so the following introduces the concept of preorder task and posterior task. The concept of preorder task is the tasks that should be completed before implementing this task. Posterior task means the tasks that this task will be effect on. Preorder task of task T_3 is T_1, and its posterior tasks are T_5, T_6, T_7 as Fig. 2. So, T_3 is affected by T_1, and effects on T_5, T_6, and T7.

Each task can be represented by 7-tuple vectors T=(N, Flag, Tb, Tc, C, P, R). N is the series number of task. Flag shows task success or not and its execution status. Flag= (1, 0, -1, 2). 1 shows success, 0 means non-execution. –1 shows failure, and 2 means in-execution. Tb and Tc are preorder task and posterior task of T respectively. C= (t_i,b_i, c_i) is a time term of task. t_i is the basic time of the whole task. b_i is the beginning time. c_i is the terminal time. P is the priority of the task. R= (r_1,r_2,⋯,r_m) is resource demand vector of task T.

3.2 Resource description

Product design is composed of a series of tasks. Different task need different able designers and resources (including software and hardware). In order to describe the problem easily, designer and other resources are considered as general resources that have proper functions and can complete different tasks. Suppose the function of resource unit is the same in one kind resource, then resource is defined R= (RID, Total, Feature). RID is the tab of resource sort. Total is the total quantity of resource. Feature is characteristic description of resource.

3.3 Problem modeling

Modeling aims to complete task in the shortest time and should meet the constraint conditions as follows:

(1) Subtask shouldn't be interrupted from the beginning to the end if the subtask needn't be reiterated.
(2) The resource shouldn't be distributed to other subtasks when it is distributed to the subtask that is in-execution.
(3) Subtask can't be executed if the least resource can't be supplied.
(4) When a subtask is completed, other subtask can use the resources that belong to the subtask.
(5) A task can't be executed until its preorder tasks are all completed.

Suppose a task has n subtasks and m resources, then it can model as follows according to the aims and constraint conditions.

$$Min \quad c_n \qquad (1)$$

$$s.t. \quad c_i \le bj, \quad T_j \in Tc_i \qquad (2)$$

$$c_i = b_i + l_i \qquad (3)$$

$$\sum_{Flag=2} \mathrm{rh} \le \mathrm{Rh} \qquad h=1,2,\cdots,m \qquad (4)$$

C_n is the time when the n^{th} subtask is completed, that is, the total time of the project. Constraint condition (2) shows the constraint relation between preorder subtask and posterior subtask. Condition (3) describes that executing task need time. Condition (4) is resource constraint condition, that is, h^{th} resource belonging to the all executing tasks can't exceed the total of h^{th} resource.

4. DYNAMIC ALLOCATION ALGORITHM OF DSIGN TASK

4.1 Task priority

The priority of the task is confirmed by its importance in the project. The more important it is, the higher prior it is, and the task will have the priority in resource allocation. The importance of task can be confirmed as follows:

(1) Total time t_i. If the task is failure that need much time, it will waster much time. So, the task demanding more time is more important.
(2) Task order. Because the result of the task will effect on his posterior task, the task in more front is more important. Task chain can show task order. The longer task chain is, the more important it is. Task chain is the number of the longest task affected by the task from the beginning to the end. Task chain can show as $maxl_i$.
(3) Number of posterior task f_i. The more posterior task it has, the more important it is.
(4) Difficulty coefficient d_i. Difficult task should be executed before easy task. The task that is too difficult to be executed should be found as early as possible in order to adjust the planning.

In a word, task priority is defined as

$$P_i = \lambda_1 \cdot t_{io} + \lambda_2 \cdot maxl_i + \lambda_4 \cdot f_i + \lambda_4 \cdot d_i \qquad (5)$$

λ_1, λ_2, λ_3, λ_4 is weight coefficient of factors, and can affect the priority of the factor. The weight coefficient is confirmed by experience.

4.2 Resource allocation

Specifications Because of resource constraint, the task may be non-executed immediately after completing its preorder tasks. Resource should be distributed to the task with high priority. If the task can't meet the demand of resource, the resource can be distributed to the next important task, which is, taking full use of the resource. R_u= (RID, Total, Usage, remains) is resource usage vector, which can record change of resource usage. R_u can gain resource number and quantity from resource R. Usage and remains represent quantity of being using and remaining, which can be dynamically modified according to the task being selected or not and required resource quantity.

4.3 Task addition

The task added dynamically changes the posterior task set of its preorder task and preorder task set of its posterior task to make its posterior task being effected by addition task.

4.4 Algorithm description

(1) Find out the first task whose preorder task is 0 and save it in a datasheet.
(2) Order the tasks whose Flag equal to 0 according to priority and schedule the higher priority.
(3) If the least resource required by a unit can't be supplied, the resource should be added.
(4) Select schedulable task according to priority and resource quantity. If resource meets the scheduling conditions, the task can be executed. Then, modify task status Flag to 1, modify resource allocation vector R and update resource user state.
(5) Cancel the task, release its resource and update resource user state after task status is success.
(6) Check success task if it need adding posterior task or not. If the answer is no, then go to 7^{th} step directly. If the answer is yes, modify posterior task set and preorder task set of modified posterior task, and then go to 7^{th} step.
(7) Check if preorder task statuses of success task are posterior tasks are success or not. If the answer is yes, add it in the sheet and then go back to 2^{nd} step. If the answer is no, then go back to 2^{nd} step directly.

5. CONCLUSION

Product design to CE can be completed by teamwork from the members of multi-object group. It need be considered that how to order design task, alternate information among tasks and constrain resource. So, it is important that how to schedule design process efficiently, how to utilize resource rationally and shorten development cycle. This paper studies design task decomposition and dynamic scheduling task, and points out task decomposition principle and dynamic scheduling algorithm to add task and distribute resource dynamically, which arrange task and distribute resource according to the priority.

6. REFERENCE

[1] G.L. Xiong, B.H. Li, et al, "Concurrent engineering research and application", Tsinghua science and Technology, 4(1999), pp.1375-1385.
[2] D.S. Todd, P. Sen, "Tacking complex scheduling problems using operations based scheduling", Adaptive computing in design and manufacture 1 (1998), pp.45-58.
[3] B. Yang, K.Z. Huang. "Assignment and planning of Task supporting concurrent Engineering", Computer Integrated Manufacturing systems. 8(2002), pp.542-560.
[4] H. Zhu, "A Role Agent Model for Collaborative Systems", Proc. of int'l Conference on Information and Knowledge Engineering, USA, 2003, pp.438-444.

Application of Immune Genetic Algorithm to Multiobjective Optimization

Jun Xu, Li Liu, Wenbo Xu
School of Information Technology, Southern Yangtze University,
Wuxi, Jiangsu 214122, P.R.China
E-mail: xujun103737@yahoo.com.cn

ABSTRACT

In this paper, immune genetic algorithm is proposed for multiobjective optimization problems by introducing the capability of the immune system to the simple genetic algorithm. The optimizing ability of the proposed algorithm is identified by comparing the result of optimization with the simple genetic algorithm. Immune genetic algorithm owns the better convergence than the simple genetic algorithm.

Keywords: simple genetic algorithm, immune algorithm, immune genetic algorithm, convergence

1. INTRODUCTION

At preset, the simple genetic algorithm (SGA) unceasingly expands in its application, regardless of its solving the actual problem or modeling [1]. This mainly relies on that genetic algorithm itself gradually is mature. During these years, many genetic algorithm researches are different from the algorithm what Hollan has proposed [2], such as different gene expression way or different crossover and mutation operator. As we know, the SGA owns such advantages as the random initialization of a population and the parallel search. It has the ability to avoid becoming trapped in a "local optimum" solution, and is designed to locate the "global optimum" solution. The SGA has drawbacks too, of course. While the great advantage of the SGA is the fact that they find a solution through evolution, this is also the biggest disadvantage. Evolution is inductive; in nature life does not evolve towards a good solution - it evolves away from bad circumstances. This can cause a species to evolve into an evolutionary dead end.

Immune algorithm (IA) [3] is a kind of intelligent optimization algorithm that simulates the biology immunity system, and has potential to provide novel method for solving problem. It has been applied in Pattern recognition and Robot control and Network invasion examination. But at present, many developed algorithms based on the immunology principle are proposed in solving the concrete questions, and don't have the general form and pattern so widely like the genetic algorithm.

Based on the advantage of IA and SGA, an immune genetic algorithm (IGA) is proposed as a new multi-disciplinary intercross and seepage algorithm [4, 5].

2. THE INTRODUCTION OF THE ALGORITHM

The immune system in a vertebrate (an organism with a backbone) consists of all the cells and tissues that recognize and defend the body against foreign chemicals and organisms. IGA models the function of the auto-adapted recognition in biological immune system. It brings the diversity of the biological immune system into the SGA. During solving the concrete questions, object function and constraints will be input as antigens, then creating initial random population and computing a series of affinity of antibodies. In the condition of keeping the diversity of the antibody, find the best antibody.

2.1 Computing of the diversity and affinity

Antibody has the enough diversity in the immune system of the biological organism so that it can destroy different antigen. It keeps a subtle balance between the antigen and antibody, between the antibody and antibody in the whole immune system. Entropy is been introduced to judge both the diversity of the antibody and the change course of the allele probability in the immunity optimization.

An immune system is composed of *N* antibodies and each antibody owns *M* genes. The number j gene's entropy in the whole N antibodies is:

$$E_j(N) = \sum_{i=1}^{n_j} p_{ij} \log(\frac{1}{p_{ij}}) \qquad (1)$$

Where n_j is the allele total in position of the number *j* gene, p_{ij} is the probability of the number *j* gene to take the place of the number i allele.

The affinity formula of the random two antibodies is:

$$(A_{tt})_{ij} = \frac{1}{1+E(2)} \qquad (2)$$

Where $E(2)$ is the average entropy of two antibodies, *t* is the antibody.

$$E(2) = \frac{1}{M}\sum_{k=1}^{M} E_K \qquad (3)$$

The affinity of between antigen and antibody is defined as the fitness of an antibody.

$$(A_{yt})_i = g_i(x) \qquad (4)$$

Where $g_i(x)$ is the fitness function of the number *i* antibody, *y* is the antigen.

2.2 Adjusting the fitness by concentration

To keep the diversity of the antibody, there is a need to adjust the concentration. Adjusting the fitness by concentration is based on the original fitness, and concentration adjustment mechanism is introduced in it. As defined in the following:

$$g_{dj}(x) = \frac{g_i(x)}{1+\alpha \ln(1+D_i)} \qquad (5)$$

Where $g_{dj}(x)$ is the fitness function of the number *j* antibody, $0 \le D_j \le 1$, $\alpha > 0$ which is decided by the

unresolved problem and experiences.

2.3 The flow of the IGA

1) Create initial random population A_1.

2) Abstract vaccines according to the prior knowledge

3) If the current population contains the optimal individual, then the algorithm halts.

4) Perform crossover on the kth parent and obtain the results B_k.

5) Perform mutation on B_k to obtain C_k.

6) Perform vaccination on C_k to obtain D_k.

7) Perform immune selection on D_k and obtain the next parent A_{k+1}, and then go to step 3.

3. SIMULATION EXPERIMENT\

The test function is:

$$f(x,y) = 0.5 + \frac{\sin^2(x^2+y^2)^{1/2} - 0.5}{1.0 + 0.001(x^2+y^2)} \qquad (6)$$

Where $x, y \in [-1,1]$.

3.1 The comparing of reaching the maximum value

We optimize the function according to SGA and IGA. Each algorithm is operated for 30 times independently. Then we compute the average value of each algorithm. We find that IGA is faster than SGA in the speed of reaching the maximum value. The experimental results are shown in the Table 1.

Table 1. LBS mean the length of binary string. LIN means the largest iterative numbers. P1 means average iterative numbers when the function reaches the maximum value in SGA. P2 means average iterative numbers when the function reaches the maximum value in IGA.

Function	LBS	LIN	SGA	IGA
			P1	P2
f	22	100	22	18

3.2 The comparing of convergence rate

In the situation of keeping current optimal solution unceasingly, SGA is convergent to reach the global optimal solution. Its astringency has already been proven [6]. But the convergent time of global optimal solution is possibly very long. IGA has introduced the computation of antibodies and the adjustment mechanism in the SGA foundation. It can keep the diversity of the antibody effectively and avoid premature convergence. Its convergence also has already been proven [7]. We optimize the function according to SGA and IGA again. The iterative numbers is 100. According to the experimental results, it is obvious that convergence rate of IGA is faster than SGA.

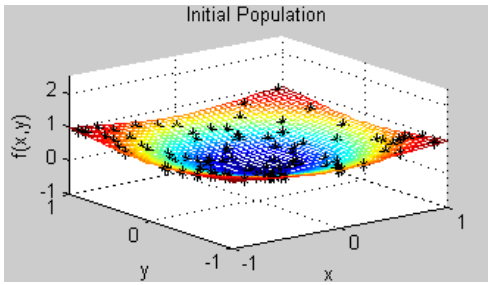


Fig. 1. The initial population of SGA

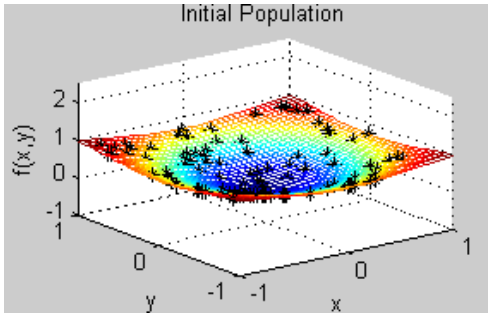


Fig. 2. The initial population of IGA

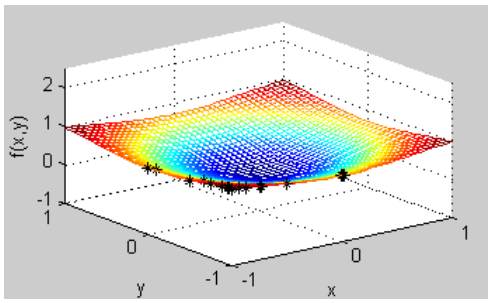


Fig. 3. GA experimental results after 100 iterative times

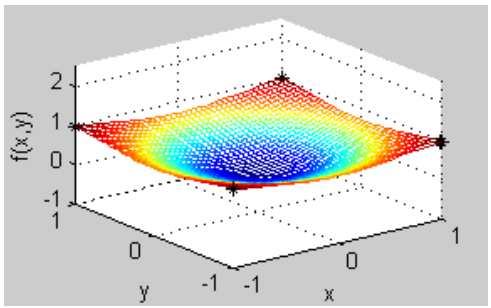


Fig. 4. IGA experimental results after 100 iterative times

4. CONCLUSION

In this paper, we introduce IGA and apply it into multiobjective optimization problem. According to analysis and comparison, this algorithm has the following characteristic:

1) It has the fast global searching performance. It overcomes the shortcoming that SGA is easy to fall into the partial optimal solution.

2) It has the ability of keeping multiple individuals. According to introducing the definition of affinity and concentration, it can keep the individual diversity in the evolution process. Thus it improves the ability of the global searching and avoids premature convergence.

5. REFERENCES

[1] C.A. COELLO, “An Updated Survey of GA-Based Multiobjective Optimization Techniques”, ACM Computing Surveys, 2000, 32(2): 109-143.

[2] J.H. Holland, Adaptation in natural and artificial systems, MIT Press, 1992.

[3] T. Li, Computer immunology, China Waterpower Press, 2004.

[4] L.A. Chen, P.M. Zhang, “Realization of immune algorithm in matlab”, Journal of Fuzhou University (Natural Science), 2004, 32(5): 821-827.

[5] W.Q. Zhang, Y.M. Xuan et al, “Application of Immune Genetic Algorithm to the Optimal Thermal Design of Louvered Panels”, Chinese Space Science and Technology, 2005, 25(5): 66-72.

[6] H.M. Li, J.S. Kou et al, The basic theory and application of genetic algorithm, Science Press, 2004.

[7] X.P. Luo, W. Wei, “General discussion on convergence of immune genetic algorithm”, Journal of Zhenjiang University (Engineering Science), 2005, 39(12): 2006-2011.

On the Optimization Methods & System of Trap Evaluation Schemes Based on Compatibility Degree and Difference Degree

Li Zhan, Changming Yang， Dongfu Yang
School of Economics, China University of Geosciences, Wuhan 430074, Hubei Province, China;
E-mail: zhanlijenny@126.com

Abstract

This paper researches on the improvement methods of trap evaluation schemes to help making more scientific petroleum-gas exploration decisions. It gave two improvement methods， optimazation model of compatibility degree maximization and optimazation model of compatibility degree maximization & difference degree minimal. The two methods were designed to evaluate petroleum traps based on Compatibility Degree and Difference Degree. And it also built an improvement system based both on four traditional trap evaluation methods and the two new ones. This system not only offered more than one evaluation schemes but also gave two improvement schemes. So it meets morden scientific decision-making principles better than many trational evaluation methods which rank traps by only one evaluation schemes. This system is successfully applied in some large oil field in China.

Keywords: compatibility degree; difference degree; trap evaluation; optimization system.

1. INTRODUCTION

Trap geology situation evaluation is a key step to petroleum-gas exploration decision analysis. Previous analysis techniqus almost gave a classification to all traps according to the evaluation results of their geological element index data by one certain rule or method, such as Probability of Geological Success[1], Artificial Neural Networks[2], Fuzzy Comprehensive Evaluation[3,4], and so on. That classification by one rule or method does privide exploration decision makers some frame of reference, but cannot meet modern scientific decision-making rules' request since modern scientific decision-making rules requair our evaluation results not only privide several evaluation outcomes but not only one, but also give the method to improve evaluation outcomes. Therefore, researching on traditional trap evaluation methods and creating improvement evaluation outcomes based on them has positive theorical and practical significance to petroleum-gas exploration decision makers.

R.L.Fu, S.K.Qin, and Z.B.Chen (1999) defined Compatibility Degree and Difference Degree to evaluate comprehensive evaluation schemes by different methods [5]. S.K.QIN (2002) researched on the optimazation system [6, 7]. This paper applies the definition of Compatibility Degree and Difference Degree in and exploation decision analysis and builds a trap evaluation optimazation system based both on four traditional evaluation metheds and on the two new ones that are created by two optimazation models. This system is successfully applied in some oil field in China. This application reveals that the new methods and system here are super to the conventional trap evaluation methods and models.

2. OPTIMIZATION METHODS & SYETEM OF TRAP EVALUATION SCHEMES

A trap evaluation scheme is the evaluation results according to a set of indicators combined with an evaluation rule or method. Several set of indicators combined with several evaluation rules or methods can provide several schemes. Then the key problem is how to create optimization methods from the several trap evaluation schemes for better decision.

2.1 Compatibility Degree of Trap Evaluation Schemes

Compatibility degree of a trap evaluation scheme is the weighted average of grading correlation coefficients of this trap evaluation scheme with other *h-1* trap evaluation schemes.

Grading Correlation Coefficient of each trap evaluation scheme can be calculated by the equation:

$$R_{ij}=1-\frac{6}{n(\eta^2-1)}\sum_{k=1}^{n}(a_k^{(1)}-a_k^{(2)})^2$$

$$(i,j=1,2,\cdots,h) \qquad (1)$$

Where $a_k^{(i)}$ represent the ranking number of trap#k in evaluation scheme *i*, $a_k^{(j)}$ represent the ranking number of trap#k in evaluation scheme *j*, *n* is the sum number of traps, *h* is the sum number of evaluation schemes.

Compatibility degree of this trap evaluation scheme $y=\{y_k\}$ with other *h-1* trap evaluation schemes can be calculated by the equation:

$$r_y=\sum_{j=1}^{h-1}\omega_j\ r_{yj} \qquad (2)$$

where $\sum_{j=1}^{h}\omega_j=1$ $(\omega_j>0)$ is the weight of trap evaluation

scheme*j*, ordinary is given $\frac{1}{h-1}$

when the decision making does not prefer to any trap evaluation scheme, y_k is the ranking number of trap*#k* in the trap evaluation scheme. If every trap evaluation method is independent, the trap evaluation scheme with higher value of compatibility degree is better than scheme with lower value of compatibility degree and its representativeness is better than the other one.

2.2 Difference Degree of Trap Evaluation Schemes

Difference degree of a trap evaluation scheme is the average of the sum number of traps whose ranking number not bigger than a certain number in this trap evaluation scheme is beyond the limition number in all other trap evaluation schemes when this trap evaluation scheme is considered as the standard ranking scheme. It can be calculated by the equation:

$$d_y = \frac{1}{h-1}\sum_{j=1}^{h} d_{yj} \tag{3}$$

where d_y is the difference degree of a trap evaluation scheme with other *h-1* trap evaluation schemes, d_{yj} is the sum number of traps whose ranking number not bigger than a certain number in this trap evaluation scheme is beyond the limition number in scheme*j*.
Clearly, if every trap evaluation method is independent, the trap evaluation scheme with lower value of compatibility degree is better than scheme with higher value of compatibility.

2.3 Optimization Model of Compatibility Degree Maximization

If there are *h* trap evaluation schemes, then a new trap evaluation scheme $y=\{y_k\}$ can be created by the rule of compatibility degree maximization**:**

$$\max_{y} r_y = \sum_{j=1}^{h} \omega_j \ r_{ij} \tag{4}$$

It can be proved that the chik with the serial of two incrementalby series of figures is not less than the chaotic chik of them by induction. Therefore, the new trap evaluation scheme with maxim compatibility degree with other *h* trap evaluation schemes can be created by the method of giving each trap a new ranking number incrementally *1，2，…，n* by its weighted average of ranking numbers in *h* evaluation schemes. And the new ranking number of each trap is just the ranking number of it in Optimazation Model of Compatibility Degree Maximization.

2.4 Optimization Model of Compatibility Degree Maximization & Difference Degree Minimal

If there are *h* trap evaluation schemes, then a new trap evaluation scheme $x=\{x_t\}$ can be created by the rule of compatibility degree maximization & difference degree minimal**:**

$$\max_{\{x_t | t\in D_{xk}\}} r_x^{(k)} \left[\min_{\{x_t | t\notin \bigcup_{i=0}^{k-1} D_I\}} d_{xk} \right]$$

$$(k=1,2,\cdots,L) \tag{5}$$

where d_{xk} is the difference degree of $\{x_t\}$ in *kth* classification, $r_x^{(k)}$ is the difference degree of $\{x_t\}$ in *kth* classification, D_{xk} is candidate groups of D_k，L is the number of classifications of trap evaluation schemes，D_k is the group of the ranking numbers of the traps in *kth* classification when $\{x_t\}$ is considered as standard trap evaluation scheme.

2.5 Selection Algorithms of Trap Evaluation

If d_{1j} represents the ranking number of trap evaluation scheme*#j* by compatibility degree and d_{2j} represents the ranking number of trap evaluation scheme*#j* by difference degree, then selection algorithms of trap evaluation are as follows:
1). Firstly，to calculate d_{1j} ($j=1,2,\ldots,h+2$) by compatibility degree;
2). Secondly, to calculate d_{2j} ($j=1,2,\ldots,h+2$) by difference degree;
3). Thirdly, to calculate the comprehensive ranking number d_j of trap evaluation scheme*#j*：

$$d_j = \sum_{i=1}^{2} \omega_i \ d_{ij}$$

$$(j=1,2,\cdots,h+2) \tag{6}$$

where ω_i is weight given by exploration decision makers

($0 \le \omega_i \le 1$， $\sum_{i=1}^{2} \omega_i = 1$)。

3. OPTIMIZATION SYSTEM OF TRAP EVALUATION SCHEMES

Fig. 1. shows the framework of trap evluation scheme optimazation system. And its operation steps are as just follows:
1). To input traps' geological element data.
2). To calculate the evaluation results and ranking number of each trap by *h* methods to provide *h* trap evaluation schemes；
3). To create two new trap evaluation schemes by optimazation model of compatibility degree maximization and optimazation model of compatibility degree maximization & difference degree minimal.
4). To calculate compatibility degree and difference degree of *h+2* trap evaluation schemes and to give them ranking numbers by selection algorithms of trap evaluation.

5). To select the best evaluation scheme as the last trap evluation results.

4. APPLICATION

The framework of trap evaluation scheme optimization system shown in Fig. 2. is designed for some oil field in West China. And table1 shows 18 traps' geological element index data of a constitution band of this oil field. Table 2. not only shows the ranking results of the trap evluation schemes by weighted average method, probability of geologicial succes, method of comprehensive index, and method of effectiveness rating, but also gives the ranking results of trap evaluation schemes by optimization model of compatibility degree maximization and optimization model of compatibility degree maximization & difference degree minimal which are created on the basis of the four ordinary evaluation methods. The compatibility degree and difference degree of all six trap evaluation schemes in Table 3. show that the trap evaluation schemes by optimazation model of compatibility degree maximization and optimazation model of compatibility degree maximization & difference degree minimal are the best trap evaluation methods, the secondary is the one by the method of probability of geologicial success, and the last one is by the method of effectiveness rating. Therefore, the trap evaluation schemes by optimazation model of compatibility degree maximization and optimazation model of compatibility degree maximization & difference degree minimal are the last trap evaluation results.

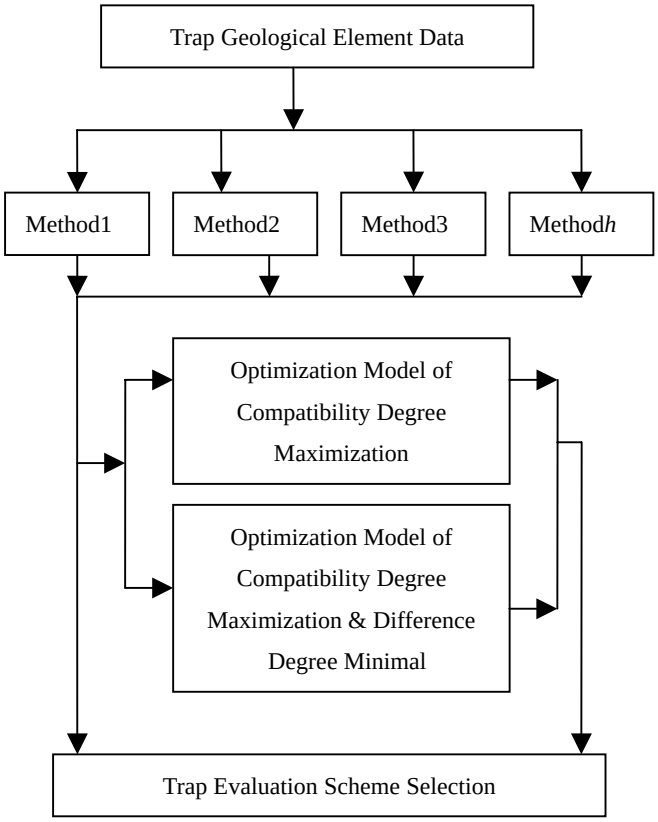


Fig. 1. The framework of trap evaluation scheme optimization system

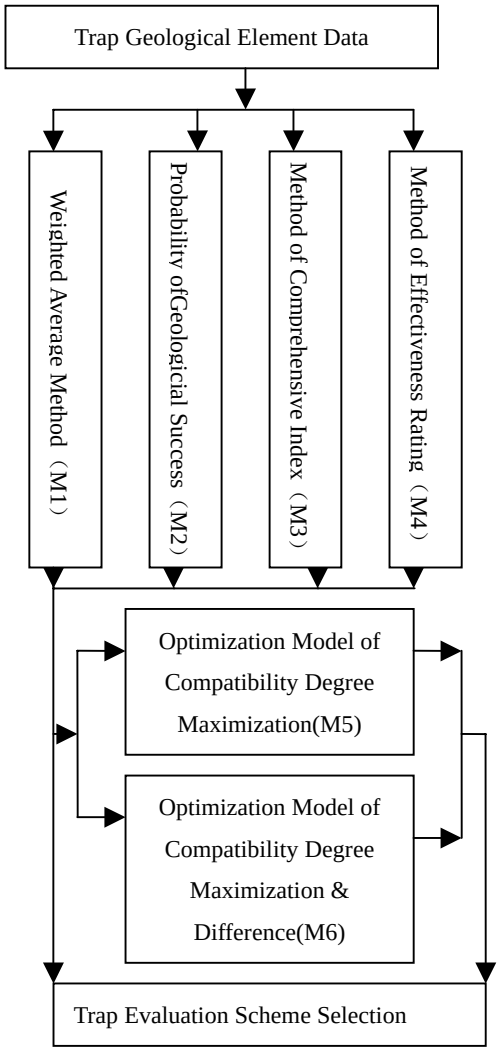


Fig. 2. The framework of trap evaluation scheme optimization system for some oil field in West China

Table 1. Trap geological element index data of a constitution band of some oil field in West China

Geological element / Trap	source rock	storing	preservation	traping	timing & migration pathways
#1	0.725	0.825	0.675	0.900	0.750
#2	0.725	0.825	0.825	0.975	0.900
#3	0.725	0.950	0.825	0.975	0.250
#4	0.725	0.950	0.825	0.750	0.250
#5	0.725	0.700	0.825	0.670	0.900
#6	0.725	0.825	0.825	0.875	0.750
#7	0.725	0.825	0.900	0.975	0.900
#8	0.725	0.825	0.825	0.975	0.250
#9	0.725	0.825	0.825	0.975	0.250
#10	0.725	0.825	0.825	0.900	0.250
#11	0.725	0.700	0.825	0.725	0.850
#12	0.725	0.700	0.825	0.725	0.850
#13	0.725	0.700	0.825	0.775	0.850
#14	0.725	0.650	0.825	0.725	0.850
#15	0.725	0.700	0.825	0.825	0.850
#16	0.725	0.700	0.825	0.850	0.850
#17	0.725	0.700	0.825	0.750	0.850
#18	0.725	0.700	0.825	0.675	0.850

Table 2. Ranking results of each evaluation scheme

Evaluation scheme / Trap	M1		M2		M3		M4		M5	M6
	Evaluation value	Ranking	Evaluation value	Ranking	Evaluation value	Ranking	Evaluation value	Ranking		
#1	0.451	11	0.854	13	107.4	11	0.537	13	10	10
#2	0.675	2	0.937	2	110.2	2	0.764	2	2	2
#3	0.206	14	0.789	14	93.4	14	0.633	5	11	11
#4	0.157	16	0.743	16	87.7	16	0.513	16	15	15
#5	0.455	10	0.874	9	102.3	9	0.559	12	9	9
#6	0.505	5	0.880	6	103.5	6	0.661	3	5	5
#7	0.717	1	0.952	1	111.5	1	0.830	1	1	1
#8	0.191	15	0.778	15	91.7	15	0.564	10	14	14
#9	0.191	15	0.778	15	91.7	15	0.564	10	14	14
#10	0.115	17	0.701	17	82.4	17	0.364	17	16	16
#11	0.479	7	0.880	7	103.0	7	0.581	8	7	7
#12	0.473	8	0.878	8	102.8	8	0.576	9	8	8
#13	0.494	6	0.885	5	103.6	5	0.594	7	6	6

Table2（Con.）										
#14	0.437	12	0.869	11	101.3	12	0.515	15	12	12
#15	0.523	4	0.894	4	104.7	4	0.617	6	4	4
#16	0.552	3	0.904	3	105.9	3	0.642	4	3	3
#17	0.455	9	0.873	10	102.1	10	0.561	11	9	9
#18	0.443	13	0.862	12	100.8	13	0.534	14	13	13

Table 3. Compatibility degree and difference degree of each evaluation scheme

Evaluation scheme	Compatibility Degree		Difference Degree		Comprehensive Ranking
	Evaluation value	Ranking	Evaluation value	Ranking	
M1	0.9569	2	0.2	1	2
M2	0.9482	4	0.2	1	4
M3	0.9548	3	0.2	1	3
M4	0.8454	5	1	2	5
M5	0.9643	1	0.2	1	1
M6	0.9643	1	0.2	1	1

5. CONCLUSIONS

Compatibility degree and difference degree of trap evaluation schemes not only provide rules to evaluate different trap evaluation schemes by different methods but also provide two new optimization methods which can create new trap evaluation schemes on the basis of ordinary trap evaluation methods.

Optimization system of trap evaluation schemes is a comprehensive system that combines innovation with seclection of trap evaluation schemes.

In optimization system of trap evaluation schemes, the former h kinds of trap evaluation methods are not limited in the four methods used in this paper. And other methods also can be applied, such as artifical neural networks, fuzzy comprehensive evaluation, and so on.

1. 6. REFERENCES

[1] Q.L. Cai, “Evaluation method and model of risks in oil-gas exploration”， Petrochemical technology economy, V.19 (6)，2003， pp 58-62.

[2]K.J.Zhu, J.X.Yang, et al, “The synthtical evaluation to the potential of petroleum prospecting based on neural network”, Systems engineering-theory & practice. N.(4), 2002, pp 131-135.

[3] C.M.Yang, M.L.Wang, et al, “The advantage of fuzzy mathematics comprehensive evaluation method”，Journal of southwest petroleum institute, V.21(2) 1999， pp 41-42.

[4]C.M. Yang, Research Report on petroleum-gas exploration & development assist decision-making system.CUG, Wuhan , 1997.

[5] R.L. Fu, S.K. Qin, et al, “Mathematical models of the compatibility with muti-valuation schemes and clasifications”, Journal of Systems Engineering, V. 14(2), 1999, pp 179-183.

[6] S.K. Qin, “Optimization system of evaluation schemes”, Journal of Systems Engineering, V. 17(2), 2002, pp 143-149.

[7] S.K. Qin, Comprehensive evaluation theory & application. Electric Industry Press, Beijing, 2003.

Design and Achievement of Integrated DTV TS Outputting and Getting Hardware Platform

Baoju Zhang
College of Physics and Electric Information, Tianjin Normal University
Tianjin, China
Email: wdxyzbj@mail.tjnu.edu.cn
And
Yuxin Cheng
School of Electronic, Peking University
Beijing, China
Email: cheng1@21cn.com
And
Guiling Li
School of Electronic Information Engineering, Tianjin University
Tianjin, China
Email: lgl@tju.edu.cn

ABSTRACT

Bit stream with different picture formats, scanning and coding parameter is widely used. It should be provided in studying digital television system, developing digital television production, supervising the running of digital television system and digital television equipment. The construction of integrated digital television (DTV) transport stream (TS) hardware platform is introduced. It was accomplished by the switched conversion of a circuit board, which was constructed on a foundation of S5933. S5933 is a special interface integration chip of Peripheral Component Interconnect (PCI) and its main functions are bidirectional. The data transmission bit-rate in outputting and getting is from 3Mbps to 30Mbps, which can satisfy digital television. The bit streams are continuous and have no failure bits.
.

Keywords: DTV, TS, Outputting, Getting, Integrate

1. APPLICATION OF S5933 IN THE SYSTEM

S5933 [1] is a 32-bit PCI controller and developed by Applied Micro Circuits Corporation (AMCC) of the United States, which has various, flexible PCI interface functions and is applicable in high-rate data transmission, such as network, digital video and multimedia.

The main functional modules of S5933 are shown in Figure 1[2]. S5933 has 3 real interfaces: the PCI bus interface, the Add-On bus interface, and the external non-volatile RAM interface through which data transmission can be realized. High-rate data transmission should be carried out between the PCI bus interface and the Add-On bus interface, and the external devices can be operated through the Add-On bus interface, which can be 8-bit, 16-bit, or 32-bit. It is also illustrated in Figure 1 that the internal FIFO (First In First Out) register, the mailbox register and the Pass-Thru data channel can all transmit data, and every channel has two groups of registers through which can realize bidirectional data transmission. The accesses to the internal registers of S5933 are convenient, and users can visit them by the PCI bus, or the Add-On bus [3].

In this bit stream outputting (getting) system, the external FIFO register is hung on the Add-On bus and cascaded with the internal FIFO register, so the bit streams are transported, passing through the PCI bus, into the internal FIFO register, then the external FIFO register and put out. (When getting, the bit streams flow in reversed direction.) The access to the internal FIFO in this system are carried out by the internal FIFO direct access pins (outputting: RDFIFO# AND RDEMPTY, getting: WRFIFO# and WRFULL) and the data-width of this direct access is controlled by the state-controlling pin, “MODE”. BPCLK provided by S5933 is the amortized PCI bus clock, which is used as a read (write) signal of the internal FIFO and a write (read) signal of the external FIFO in the outputting (getting) system after frequency-divided. SYSRST is used to control the reset of the external FIFO. All the software-controlling signals for control circuit of the external FIFO are provided by nvRAM interface pins.

2. THE OPERATING PRINCIPLE AND 16-BIT DATA TRANSMISSION OF FIFO

In the bit stream outputting system, an asynchronous FIFO is selected and used as data buffer, then the high-rate bit streams read out from the hard disk can be put out at

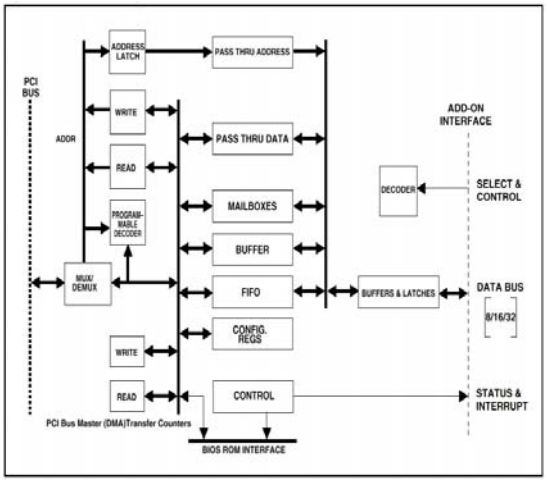


Fig. 1. The internal framework of S5933

constant bit-rate. In this project, the asynchronous FIFO-CY7C464-15PC (32Kbyte), made by CYPRESS company of the United States, is selected. It is similar to the FIFOs made by the other companies at the usage and the principle of operation.

The read/write processes of FIFO are controlled by read/write control signals respectively, and these two processes do not interfere with each other. When read control signal is in effect and the state of empty-flag-EF# is non-empty, the data of the internal memory cells which are pointed by the internal reading pointer will be read onto the outputting data bus Q0-Q8. When write control signal is in effect and the state of full-flag-FF# is non-full, the data of the inputting data bus D0-D8 will be loaded into the internal memory cells which are pointed by the internal writing pointers. When read and write control signals are both in effect and the state of empty-flag, EF#, and full-flag, FF#, are invalid, the operation of reading and writing will be executed by the internal read/write control logic units, and the instruction first in is executed first.

In consideration of the requirement of TS 8-bit parallel outputting, 8-bit reading to the internal FIFO is performed directly by the Add-On OP register access pins (RD#, BE#[3:0] and ADR [6:2]) of S5933.The maximum rate of data transmission which the outputting system requested is 3.75Mbyte/s (the corresponding rate of serial transmission is 30 Mbps), and the rate of writing in is higher than the rate of reading out in the external FIFO, so the read clock of the internal FIFO (the write clock of the external FIFO), WCLK, is set at 15MHz. When observing with the oscilloscope, the low 8 bits of every 32 bits displayed and the other 24bits disappeared. The data transmission was unstable and terminated automatically. It was found that both the full and the empty flags are true in the FIFO status indicating bits of the S5933 MCSR (Master Control Statue Register), that is to say, the internal FIFO is in an illegal status of both empty and full, which resulted from the high read clock frequency.

Because of mentioned above, the direct access pin, RDFIFO#, is used to read the internal FIFO, and the pin, MODE, of S5933 is set at high level. The data is put out from FIFO at 16-bit. Consequently, read pointer of the internal FIFO will be handled automatically, and read clock frequency of the internal FIFO will be reduced by two times at the same time. All these will eliminate the interference. In this project, there should be two external FIFOs out of S5933. The data is put in at 16-bit and put out at 8-bit under the control of the peripheral circuit, which make the circuits more complex but add leeway for raising the operating frequency. The plan of 16-bit data transmission resolves these two questions perfectly. Taking outputting as an example, Fig. 2. illustrates all the main signals time sequences in a transmission cycle of the system. A transmission cycle refers to the procedure from one time being read emptiness in the external FIFO to the next. In a transmission cycle, DMA transmission and hard disk transmission are executed once respectively, and an interrupt request is send out by S5933 to the system. DMA and HDD in the figure are not real signals, but refer to two procedures. CTL3 is software-controlling signal, which is used to control WR#. At the ending of a DMA transmission procedure, the interrupt service program is used to clear the counter status in control circuit through CTL3 (as shown in f), which makes the control circuit would not send out read write pulses, WR#, after DMA is enabled next time. R1# and R2# are the external FIFO read clocks in Fig. 2.

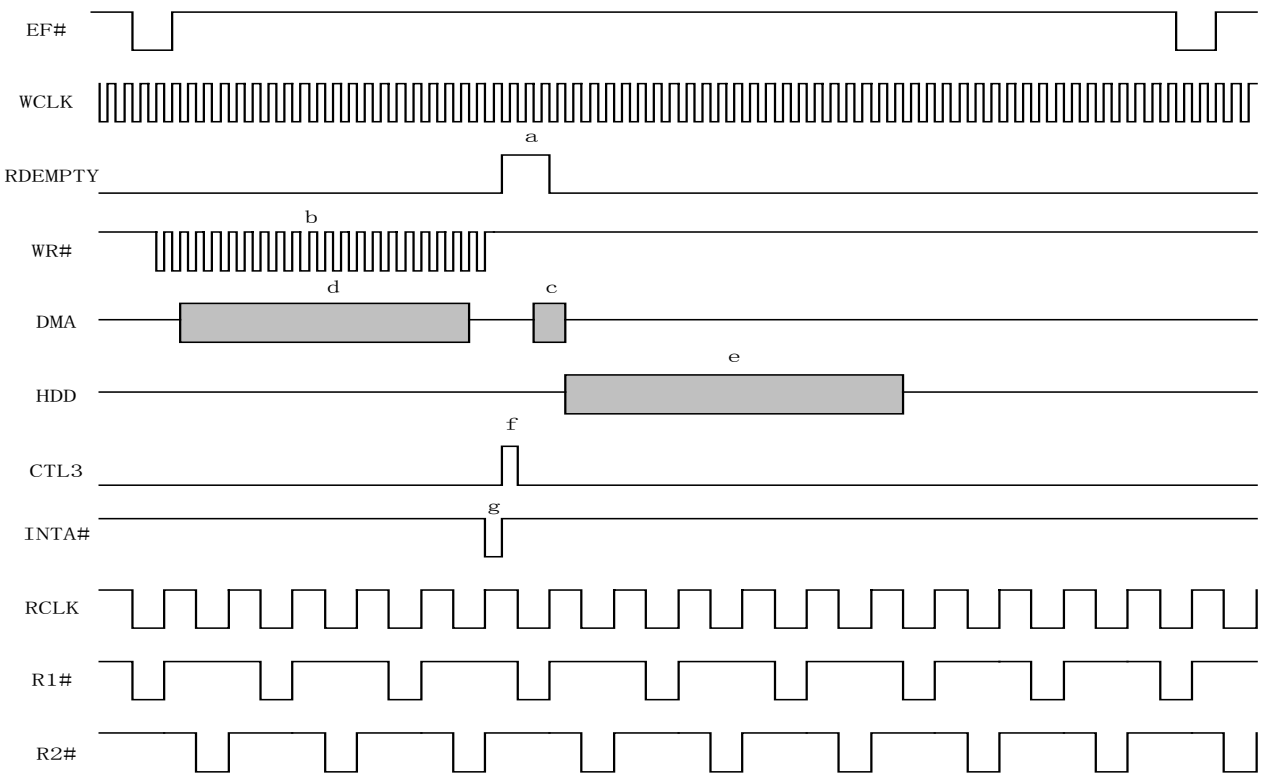


Fig. 2. TS outputting time sequence

3. THE PROOF OF AN INTERGRATING PLAN ABOUT TS OUTPUTTING AND GETTING

In the outputting system, MPEG-2 transport stream is stored in the hard disk as data files. Under the control of CPU, block data are transmitted to the memory and the FIFOs in the outputting card, passing through the PCI bus, and put out as parallel bit streams at last. In the getting system, the function is exactly contrary to the outputting system, as is shown in figure 3 and figure 4.

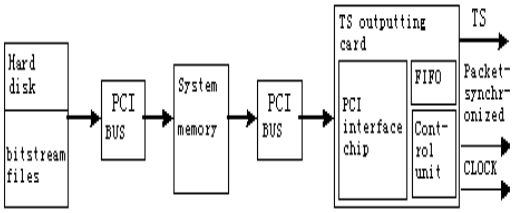


Fig. 3. The outputting system framework

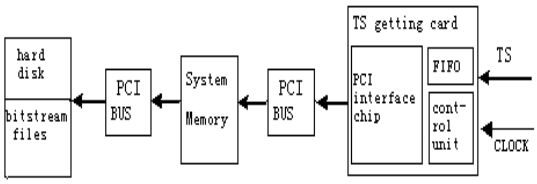


Fig. 4. The getting system framework

It can be seen from the discussion above that most functions of S5933 have the bidirectional character. Data transmission from the PCI to the ADD-ON or from the ADD-ON to the PCI can be finished separately by two sets of internal FIFOS of S5933 [4]. The read and write control circuits of the external FIFOs have no difference in the outputting circuit and the getting circuit, so these two circuits can be integrated. The outputting and getting functions can be realized with one set of external FIFO and its peripheral control circuit and a chip, S5933. The configuration of software is shown in Table 1.

An electronic switch can be used in the hardware circuit to perform the switching between TS outputting and getting modes. The clock circuit of writing (reading) the external FIFO in the outputting system can be used as read (write) clock circuit in the getting system [5]. The packet-synchronized extract circuit set in the outputting system is not needed at getting mode. The outputting system should also provide one set of parallel bit streams clock signals, which are put out through long line driving circuit [6,7,8]. In the getting circuit, the write clock signals of the external FIFO are extracted from the source. The integrated system framework is illustrated in Fig. 5.

4. TEST AND CONCLUSION

In the test, the outputting of the bit streams was accomplished first and the outputting card was realized. When debugging the system further, a set of testing codes were designed and the waveform displayed in the oscilloscope was observed to confirm the correctness of the system. To test the continuation property of bit streams, test codes were set at the beginning and ending positions of the memory buffer, and the signal, NULL, was used to trigger the oscilloscope, then the bit streams waveform could be displayed. After finishing this test, on-line communication was performed, and cyclical play was accomplished after optimizing the program at the foundation of single play. The digital satellite receiver and the DTV STB (set top box) were used as decoders separately, and the bit streams displayed in the monitor. Exceptions of video or audio did not appear in continuously playing.

Table 1. The Working Methods of TS Outputing and Getting

	Outputting	Getting
S5933 internal FIFO	PCI to Add-On	Add-On to PCI
Configure spatial registers	VID DID BADR INTLIN	VID DID BADR INTLIN
Set the address registers	MRAR	MWAR
Set the counter	MRTC	MWTC
Enable transport interrupt	Set INTCSR bit 15	Set INTCSR bit 14
Reset FIFO	Set MCSR bit 25	Set MCSR bit 26
Define FIFO management mode	Set MCSR bit 13	Set MCSR bit 9
Read/write transport priority	Set MCSR bit 12	Set MCSR bit 8
Enable DMA	Set MCSR bit 14	Set MCSR bit 10

At a foundation of these tests, the TS which were coded by software and stored in hard disk were put out and decoded by STB, then the output of the outputting card would be used as the source of getting if the decoded data was exact. First, the outputting card and the getting card were inserted separately in the PCI sockets of two PCs and connected by flat data line. Then, the test codes generated in one PC, which had the outputting card, or the standard bit streams stored in its hard disk, were put out and received by the getting card in the other PC. All these were useful to check the correctness of the getting data in the early stage of the test. At first, getting test codes. Test codes used in the system were: 1F E5 3C F2 58 76 B4 6F . These codes guaranteed that the level of every bit was contrary to its two adjacent levels and it was useful to observe the waveform. Every circle was made up of 8 Bytes. The getting codes in the early test were: E5 1F F2 3C 76 58 6F B4, and it owed to the irregular write time sequence of the two FIFOs. So, the write time sequence of the external FIFO was adjusted carefully and the compensation for time delay was added, and the getting codes become correct too. After this, getting the TS. The TS stored as hard disk files were put out through the outputting card and decoded by STB. The results displayed in the monitor were precise and had no error code.

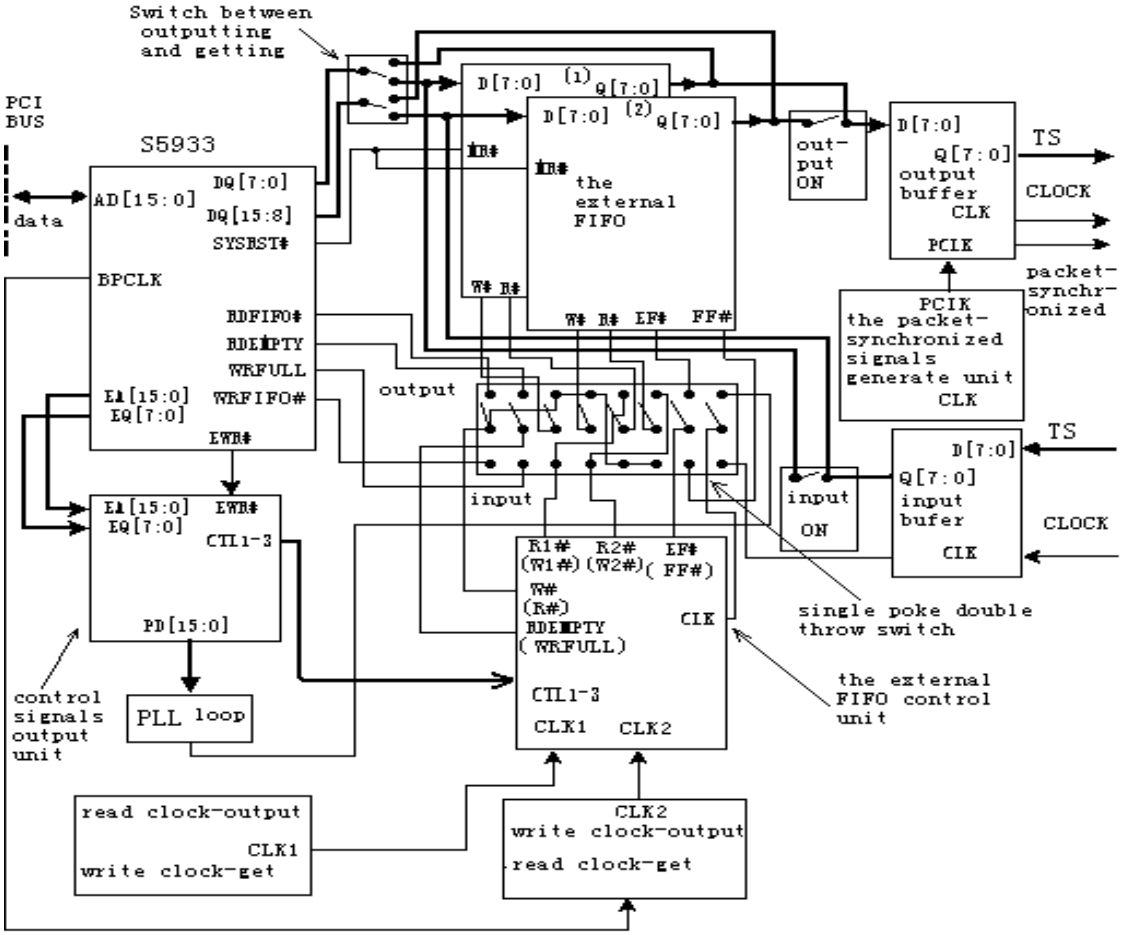


Fig. 5. The integrated system framework

As mentioned above in detail, the direction of bit streams in the outputting and getting system is reversible, so the read and write control circuits of the external FIFO in the outputting and getting circuits can be incorporated. Switched by a switch, the two functions of outputting and getting are performed by a set of external FIFO and its peripheral control circuit in one circuit board, and the integrating of outputting and getting is accomplished. Special attention should be paid in order to enhance the operating frequency when wiring the PCB. To avoid signal degradation, it must be less than 4 inches from the source to the destination for the copper clad wire, which transmits read, and write clock signals of the external FIFO.

The results of the test convey that the integrating plan is precise and feasible. The getting TS are correct and the decoded getting TS are exact and continuous, and have no error code.

5. REFERENCES

[1] S5933 PCI Controller Data Book. Applied Micro Circuits Corporation, 1997. 2-216.

[2] PCI Products Data Book S5933/S5920. Applied Micro Circuits Corporation, 199838-248.

[3] S5933 PCI Matchmaker Developer's Kit Technical Reference Manual. Applied Micro Circuits Corporation, 1997,18-25.

[4] Application Note: S5933 32 Bit FIFO Example, Revision 2. Applied Micro Circuits Corporation, Nov.1998, 2-11.

[5] Application Note: Bus Mastering With the S5933 PCI Match Maker. Applied Micro Circuits Corporation, 1997,2-8.

[6] ISO/IEC 13818-1, Generic Coding of Moving Picture and Associated Audio: Systems, 1994,46-84.

[7] ISO/IEC 13818-2, Generic Coding of Moving Picture and Associated Audio: Video, 1994,2369.

[8] ISO/IEC 13818-3, Generic Coding of Moving Picture and Associated Audio: Audio, 199415-6.

RAID5x: A Performance-optimizing Scheme against Double Disk Failures*

Ming Hu
Department of Computer Science, Wuhan University of Science and Engineering
Wuhan, HuBei Province 430073, China
Email: stereotype@263.net

And

Minghua Jiang
Department of Computer Science, Wuhan University of Science and Engineering
Wuhan, HuBei Province 430073, China
Email: mhjiang@126.com

ABSTRACT

A new scheme RAID5x was proposed to improve the performance of accessing storage systems with high reliability. It was a hybrid of an XOR-based code and mirroring for disk arrays to tolerate up to double disk failures. RAID5x is not only just higher in storage efficiency than other code-mixing solutions of the same type, but also it has the higher accessing performance over them and existing XOR-based codes. Moreover, its data placement is easier to find than that of XOR-based codes. Its mirroring placement makes it more adaptive and flexible to implement the dynamic load balance of the storage system than other schemes. Based on this, it can more efficiently improve the performance of accessing storage data, especially requests of small size and hotspot data.

Keywords: Disk Arrays, RAID5x, Data/Parity Layouts, Accessing Performance, and Storage Efficiency.

1. INTRODUCTION

Although RAID6 can protect against multi-disk failures, both Reed-Solomon coding and decoding [1] require complex finite field arithmetic, so it is very hard to attain a good performance. Thus, XOR-based codes such as EVENODD [2], RM2 [3], DH1, and DH2 [4] and multi-level RAIDs such as RAID5+5 and RAID5+1[5] were proposed. However, they belong to combinatorial problems, so such a solution is very hard to find. Even if such a solution is found, its data/parity layout may reduce the performance of accessing storage system to some extent. G.A.Gibson, L.Hellerstain, R.M.Karp, and so on [6] present standards for evaluating them: mean time to data loss, overhead of parity disks, parity-updating cost, and the parity group size. Although they have the optimal storage efficiency, they have more read/write synchronization problems than RAID5. Among multi-level RAID schemes, RAID5+1 is a code-mixing solution against double disk failures. It is easier to implement and it has the improved performance, but the very low storage efficiency.

A new code-mixing scheme against double disk failures was proposed for storing data in a disk array or other reliable storage system. It was called RAID5x because it had an XOR-based erasure code like RAID5 [7] and mirroring placement like RAID-x [8]. Its placement characteristics were not found in other code-mixing solutions. In its mapping layout, there are multiple groups of data and each group of data, their parity and mirrored copies are evenly distributed across distinct disks. A disk array with RAID5x placement has three separate areas for storing data, parity information, and mirrored copies respectively and so does each disk in the disk array. In what follows, these characteristics will be described in depth. Finally, a brief summary shows that RAID5x has the performance over other schemes against double disk failures.

* Supported by Natural Science Foundation under Grant No. 2004ABA015 and Key Technologies R&D Project of HuBei Province under Grant No. 2005AA104A15

2. PROBLEM FORMULATION

Without loss of generality, consider N disks each of which contains M stripe units denoted as $S(i,j)$, where $0 \leqq i \leqq M\text{-}1$ and $0 \leqq j \leqq N\text{-}1$. For example, the j-th disk has M stripe units denoted as $S(0,j),S(1,j),\ldots,S(M\text{-}1,j)$. There are two kinds of method to protect against disk failures: mirrored and parity redundancy. In a disk array, Stripe unit is a basic storage block over which check data for parity is computed or at which the parity, a mirrored copy or data is stored. To tolerate disk failures in a disk array, each m stripe units for logical data at least needs $m+f$ storage locations, where m physical stripe units are used for storing data and f physical stripe units for storing redundant information. The permutation of $m+f$ physical stripe units forms a layout period in a disk array, and is called a periodical stripe.

In a code-mixing scheme, each stripe unit for data contributes to both a stripe unit for parity and a stripe unit for mirrored copy. For m stripe units for data, mirroring needs m stripe units for its copy and XOR-based coding needs some stripe units for its parity, so $f>m$. For example, in RAID4+1 of 8 disks, 3 stripe units (m=3) for logical data need 8 physical stripe units, where 5 stripe units (f=5) for storing redundant information. In order to tolerate double disk failures and to guarantee even distribution of parity information across disks, a placement method of orthogonal mirroring and parity interleaving is defined as follows.

Definition 1. Orthogonal mirroring and parity interleaving (OMPI) placement: for N disks, N-2 stripe units for data on N-2 distinct disks contribute to computing their parity stored on one of two more disks, and copies of N-2 stripe units for data are sequentially stored one after another on another of them.

In order to guarantee even distribution of parity information across disks, a periodical stripe needs N stripe units for parity, one on each disk. For an OMPI placement, each N-2 stripe units for data generate its parity and N-2 copies, and N stripe units for parity need $m=(N\text{-}2)\times N$ and $f=N+(N\text{-}2)\times N=(N\text{-}1)\times N$, so these data, parities and copies form a periodical stripe with the size of $m+f=(2N\text{-}3)\times N$. A disk array with N disks at most has $L=[(M\times N)/((2N\text{-}3)\times N)]$

=[M/(2N-3)] periodical stripes and the storage efficiency u=L×(N-2)/M, where [x] is the greatest integer less than or equal to x. The size of a stripe unit is in sectors and its changing leads to a different value of M. If the value of M/(2N-3) is an integer, the storage efficiency of the disk array equals to the one of a periodical stripe, (N-2)/ (2N-3).

It is difficult to find a data/parity placement only for an XOR-based code against double disk failures. However, the following theorem shows that it is easy to find an OMPI placement against double disk failures.

Definition 2. Consistent data layout: the disk distribution of each stripe unit for data and each stripe unit for redundant information in a periodical stripe does not conflict with disk permutation requirements of redundancy-implementing scheme.

Theorem 1.Any consistent data layout obtained by an OMPI placement tolerates up to two disk failures.

Proof. We have to show that we can reconstruct lost data (, parity, or copy) when any one or two disks fail. Supposed a consistent data layout is obtained in accordance with an OMPI placement and Definition 2. A stripe unit for data just joins in a parity check group and mirrored copies in the same group are stored on the same disk, so we just consider a parity check group. When a disk fails, data on it can be reconstructed by reading its copy on another disk, parity on it can be recomputed by reading data in a parity check group on distinct disks, and copies on it can be reconstructed by reading data in a parity check group on distinct disks. When two disks fail, there are three cases for a parity check group: (a) no data on both of them; (b) data on one of them; or (c) data on both of them. For case (a), the parity on one of them can be recomputed by reading data in a parity group on other N-2 distinct disks and copies on another can be reconstructed by reading data in a parity check group on other N-2 distinct disks. For case (b), there are two cases: (d) its parity is on another of them or (e) correspondent copies are on another of them. For case (d), data on one of them can be reconstructed by reading its copy on other disk and its parity can be recomputed from the reconstructed data and remaining data on other N-3 disks. For case (e), data on one of them can be reconstructed from parity and remaining data on other N-3 disks and correspondent copies on another can be reconstructed from the reconstructed data and remaining data on other N-3 disks. For case(c), data on both of them can be reconstructed by reading their copies on other disk. Theorem is correct because each parity check group is independent.

D (k,l), P(k,l), and C(k,l) respectively denote the k-th stripe unit for data, parity, mirrored copy on the l-th disk, where 0≤k≤N-3 and 0≤l≤N-1. In accordance with Theorem 1 and Definition 1, only if a consistent data layout exists, it must protect against double disk failures. For the data layout, each client's logical data address A can be turned into D(k,l) in stripe unit, then leads to correspondent P(k,l) and C(k,l) according to redundant coding scheme, and finally turn into the physical address S(i,j) in the disk array, where 0≤i≤M-1 and 0≤j≤N-1. For a disk array with N=4, a data layout of a periodical stripe is obtained in accordance with Definition 1 as follows.

D_0	D_1	$P_{0,1}$	D_0	$P_{0,1}= D_0 \oplus D_1$
D_2	D_3	D_2	D_1	$P_{2,3}= D_2 \oplus D_3$
$P_{4,5}$	D_4	D_3	$P_{2,3}$	$P_{4,5}= D_4 \oplus D_5$
D_6	D_5	D_4	D_5	$P_{6,7}= D_6 \oplus D_7$
D_7	$P_{6,7}$	D_6	D_7	

Here is another such that N=4.

D_0	D_1	D_2	D_3	$P_{5,6}= D_5 \oplus D_6$
D_4	D_5	D_6	D_7	$P_{2,3}= D_2 \oplus D_3$
$P_{5,6}$	$P_{2,3}$	$P_{4,7}$	$P_{0,1}$	$P_{4,7}= D_4 \oplus D_7$
D_2	D_4	D_0	D_5	$P_{0,1}= D_0 \oplus D_1$
D_3	D_7	D_1	D_6	

3. THE PERFORMANCE-OPTIMIZING LAYOUT RAID5X

Theorem 1 shows that, if the layout meets the consistent condition, it is easy to find a data layout against double disk failures, and coding and decoding for a parity check group do not care in what order data in the parity check group are stored in the disk array. So another, turned into from one mapping structure, is feasible if it does not destroy the consistency of data layout. However, a different mapping structure leads to the different performance of accessing storage system to some extent, so a performance-optimizing data layout RAID5x is chosen for a disk array and described in Table 1. It requires three sequential storage areas in each disk, where L× (N-2) stripe units are used for storing data, L stripe units for storing parity, and L×(N-2) stripe units for storing mirrored copies. Supposed each disk contains M stripe units. Each disk still has M-L× (2N-3) stripe units for storing system information or others. A disk has at most L× (N-2) stripe units for storing data and a disk array with N disks has L×(N-2)×N stripe units for storing data, so the total capacity of a disk array available for data is L×(N-2)×N in stripe units.

The physical mapping of RAID5x is described as follows:
D(i,j)→S(i,j) where 0≤j≤N-1 and 0≤i≤L(N-2)-1.
P(i,j)→S(i+L(N-2),j) where 0≤j≤N-1 and 0≤i≤L-1.
C(i,j)→S(i+L(N-1),j) where0≤j≤N-1 and 0≤i≤L(N-2)-1.

Table1. The Storage layout of RAID5x

	Disk0	Disk1	…	Disk(N-1)
Data area	D(0,0)	D(0,1)	…	D(0,N-1)
	D(1,0)	D(1,1)	…	D(1,N-1)
	…	…	…	…
	D(L(N-2)-1, 0)	D(L(N-2)-1, 1)		D(L(N-2)-1, N-1)
Parity area	P(0,0)	P(0,1)	…	P(0,N-1)
	P(1,0)	P(1,1)	…	P(1,N-1)
	…	…	…	…
	P(L-1,0)	P(L-1,1)	…	P(L-1,N-1)
Mirrored area	C(0,0)	C(0,1)	…	C(0,N-1)
	C(1,0)	C(1,1)	…	C(1,N-1)
	…	…	…	…
	C(L(N-2)-1, 0)	C(L(N-2)-1, 1)	…	C(L(N-2)-1, N-1)

The mapping scheme of data area is simple. The client's address A can be calculated in form: A=iN+j, where 0≤i≤L (N-2)-1 and 0≤j≤N-1. Conversely, the physical address S (i,j) can be calculated in form: i=[A/N] and j=Amod$N$, where 0≤A≤L×(N-2)×N-1. From the beginning of data area, each N-2 stripe units for data are associated with one data group by which check data for parity can be computed and mirrored copies can be constructed. Consider data group S(I_0,J_0), S(I_1, J_1), …, S(I_{N-3}, J_{N-3}), stripe unit S(I_p,J_p) for parity and S(I_{C0},J_C),S(I_{C1},J_C),…,S(I_{CN-3},J_C) for mirrored copies such that J_k≠J_p≠J_C, S(I_0, J_0) ⊕ S(I_1, J_1) ⊕ … ⊕ S(I_{N-3}, J_{N-3})=S(I_p, J_p), and S(I_0, J_0)=S(I_{C0}, J_C),S(I_1, J_1)=S(I_{C1}, J_C), …, S(I_{N-3}, J_{N-3})=S(I_{CN-3}, J_C), where 0≤k≤N-3. In this way, the mapping from data area to parity and mirrored area is obtained one data group by one data group.

For any data stripe unit $S(i,j)$ such that $0 \le j \le N-1$ and $0 \le i \le L(N-2)-1$, the address $S(I_p,J_p)$ of the stripe unit for its parity is computed as follows:

$$I_p = L \times (N-2) + [(iN+j)/(N(N-2))]$$

$$J_p = \begin{cases} \{([iN+j)/(N-2)] \bmod N + 1) \times (N-2)\} \bmod N \\ \{([iN+j)/(N-2)] \bmod N + 1) \times (N-2)\} \bmod N + 1, \\ \text{if } N \text{ is even and } [(iN+j)/(N-2)] \bmod N \ge N/2 \end{cases}$$

The address $S(I_C,J_C)$ of the stripe unit for its mirrored copy are computed by the way the address $S(I_p,J_p)$ is computed.

$$I_C = L \times (N-1) + [(iN+j)/(N(N-2))] + (iN+j) \bmod (N-2)$$

$$J_C = \begin{cases} \{([iN+j)/(N-2)] \bmod N + 1) \times (N-2)\} \bmod N + 1 \\ \{([iN+j)/(N-2)] \bmod N + 1) \times (N-2)\} \bmod N, \\ \text{if } N \text{ is even and } [(iN+j)/(N-2)] \bmod N \ge N/2 \end{cases}$$

Theorem2. The data layout of RAID5x is consistent.

Proof In accordance with the distribution mechanism of parity and mirrored copies, I_p and I_C are computed in a periodic stripe such that $j \ne J_p \ne J_C$, so, for the same data group, the data layout of RAID5x meets the demands of Definition 1. Supposed the data layout of RAID5x is not consistent. There must exist $S(i_1, j_1)$ and $S(i_2, j_2)$ which belong to the different data groups and whose parities are on the same disk. Let $s = [(i_1N+j_1)/(N-2)]$ and $t=[(i_2N+j_2)/(N-2)]$ such that $\{(s \bmod N+1)\times(N-2)\}=\{(t \bmod N+1)\times(N-2)\} \bmod N$. In the same periodical stripe, there are at most $N\times(N-2)$ stripe units, so $|s-t|<N$ and $s=t \bmod N$ is impossible. If N and $N-2$ are relatively prime, $\{(s \bmod N+1)\times(N-2)\} = \{(t \bmod N+1)\times(N-2)\} \bmod N$ is impossible because $s \ne t \bmod N$. If N is even, N and $N-2$ are not relatively prime. The product of any integer and $N-2$ are even and $\{-2(s \bmod N+1)\} = \{-2(t \bmod N+1)\} \bmod N$ if $t = s+N/2$. However, when $t=s+N/2$, J_p of $S(i_1, j_1) \ne J_p$ of $S(i_2, j_2)$, and similarly, J_C of $S(i_1, j_1) \ne J_C$ of $S(i_2, j_2)$. This is contrary with the hypothesis, so the data layout of RAID5x is consistent. The theorem is complete.

Example1. Here is the data layout of the first periodical stripe in RAID5x with N=4. The other L-1 periodical stripes begin with the second one which follows the first one closely and are located in the area with dots as follows.

D_0	D_1	D_2	D_3	$P_{0,1}= D_0 \oplus D_1$
D_4	D_5	D_6	D_7	$P_{2,3}= D_2 \oplus D_3$
⋮	⋮	⋮	⋮	$P_{4,5}= D_4 \oplus D_5$
⋮	⋮	⋮	⋮	$P_{6,7}= D_6 \oplus D_7$
$P_{2,3}$	$P_{6,7}$	$P_{0,1}$	$P_{4,5}$	
⋮	⋮	⋮	⋮	
D_6	D_2	D_4	D_0	
D_7	D_3	D_5	D_1	
⋮	⋮	⋮	⋮	

Example2. Similarly, the next is the data layout of the first periodical stripe in RAID5x with N=5. The way for data layout is like the case with N=4.

D_0	D_1	D_2	D_3	D_4	$P_{0,2} = D_0 \oplus D_1 \oplus D_2$
D_5	D_6	D_7	D_8	D_9	$P_{3,5} = D_3 \oplus D_4 \oplus D_5$
D_{10}	D_{11}	D_{12}	D_{13}	D_{14}	$P_{6,8} = D_6 \oplus D_7 \oplus D_8$
⋮	⋮	⋮	⋮	⋮	$P_{9,11} = D_9 \oplus D_{10} \oplus D_{11}$
⋮	⋮	⋮	⋮	⋮	$P_{12,14} = D_{12} \oplus D_{13} \oplus D_{14}$
$P_{12,14}$	$P_{3,5}$	$P_{9,11}$	$P_{0,2}$	$P_{6,8}$	
⋮	⋮	⋮	⋮	⋮	

D_6	D_{12}	D_3	D_9	D_0
D_7	D_{13}	D_4	D_{10}	D_1
D_8	D_{14}	D_5	D_{11}	D_2
⋮	⋮	⋮	⋮	⋮

4. CONCLUSIONS

In a periodical stripe, XOR-based codes against double disk failures have the lower redundancy and the higher storage efficiency up to $(N-2)/N$; RAID5x has the storage efficiency of $(N-2)/(2N-3)$; RAID4+1 and RAID5+1 have the storage efficiency of $(N-2)/(2N)$. Thus, the storage efficiency of RAID5x is higher than that of code-mixing solutions of the same type and is lower than that of XOR-based codes against double disk failures.

If RAID5x's parity and mirrored area in Table 1 are disregarded, its data placement is the same as RAID0. For a read request, RAID5x has the parallel degree N of storage accessing and multiple requests with continuously sequential stripe units can be coalesced into a single request at the same disk to reduce the overhead of disk accessing, so it has the similar performance like RAID0. The coalesced requests make RAID5x have the performance over RAID4+1 and RAID5+1.

In accordance with its mapping mechanism, RAID5x's data placement is evenly distributed across multiple disks and each disk has three areas with the same structure: data, parity, and mirrored area as is illustrated in Example1 and Example 2. For a write request, RAID5x has the parallel degree N and the higher performance of storage accessing over other scheme against double disk failures. Its layout makes data, parity, and mirrored area have the higher request-coalescing capability to reduce the total overhead of disk accessing and does not leads to performance bottleneck on individual disks.

Many XOR-based codes against double disk failures require more read/write synchronization operations than RAID5. In case of a single disk failure, RAID5x can recover the lost data by mirrored copies, so it has the better recovery performance over RAID5. In case of two disk failures, RAID5x can recover the lost data by mirrored copies or its parity depending on the disk location of parity or mirrored copies, so it has the better performance over other erasure codes against double disk failures. When failing disks are displaced by the normal disks, RAID5x has the better reconstruction performance over other erasure codes against double disk failures.

If there are three request queues respectively associated with data, parity, and mirrored area in a disk, the storage system with RAID5x can apply the strategies for load-balancing, attain the higher dynamic load-balancing capability, and reduce the total overhead of disk accessing. It is difficult for other schemes against double disk failures to attain. In most cases, a read request of small size can directly access its mirrored copy on a single disk in a disk array with RAID5x while it may be divided into two smaller requests on the distinct disks in a disk array with RAID0. The better performance of accessing read requests of small size and load-balancing capability makes RAID5x process the hotspot data better.

5. REFERENCES

[1] I.S. Reed, G. Solomon, "Polynomial Codes over

Certa_
in Finite Fields", Journal of The Society for Industrial and Applied Marthmatics, 1996, 8.300-304.

[2] M. Blaum, J. Brady, et al, The EVENODD code and its generalization, In J.Jin, T.Cortest, and R.Buyya, editors, High Performance Mass Storage and Parallel I/O: Technologies and Applications, New York:IEEE and Wiley Press, 2001.

[3] Y.J. Nam, D.W. Kim, et al, "Enhancing Write I/O Performance of Disk Array RM2 Tolerating Double Disk Failures", In Proceedings of International Conference on Parallel Processing, 2002, 211-218.

[4] N. Lee, S.B. Yang, et al, "Efficient Parity Placement Schemes for Tolerating up to Two Failures in Disk Arrays", Journal of Systems Architecture, 2000,46.1383-1402.

[5] Q. Xin, E.L. Miller,et al, "Reliability Mechanisms for Very Large Storage Systems", In Proceedings of 20th IEEE/11th NASA Goddard Conference on Mass Storage Systems and Technologies, 2003, 146-156.

[6] G.A. Gibson, L. Hellerstain, et al, "Coding Techniques foe Handling Failures in Large Disk Arrays", Computer Science Technical Report, CSD88-477, 1988, University of California, Berkeley, 1-14.

[7] S. Wu, H. Jin, et al, "Parity-distribution: a Short-cut to Reliable Cluster Computing System", In Proceedings of The Fourth International Conference/ Exhibition on High Performance Computing in the Asia-Pacific Region, 2000, 1.417-422.

[8] K. Hwang, H. Jin, et al, "RAID-x: a New Distributed Disk Array for I/O-centric Cluster Computing", In Proceedings of The Ninth International Symposium on High-Performance Distributed Computing, 2000, 279-286.

A New Dynamic Scheduling Algorithm for Real-Time Homogenous Multiprocessor Systems *

Yuhai Yang [12] Xuelian Bin [2] Shengsheng Yu [3] Jingli Zhou [4]
[1,3,4] National Storage System Laboratory, School of Computer Science &Technology, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China;
[1,2] Airforce Army Radar Academy, Wuhan, Hubei 430019, China
Email:yhyang@wtwh.com.cn

ABSTRACT

Many time-critical applications require predictable performance and tasks in these applications have deadlines to be met. In this paper, we develop a new algorithm for non-preemptive scheduling of dynamically arriving real-time tasks (aperiodic task) in multiprocessor systems. The new algorithm includes two policies: task selection policy and processor selection policy, named grouping policy and properly choosing policy (PCP), respectively. Grouping policy and PCP aim to maximize the effective utilization of resources and processors, respectively. Simulation results show this algorithm outperforms the myopic algorithm and the thrift algorithm.

Keywords: Multiprocessor Systems, Dynamic Scheduling, Non-preemptive Scheduling, And Aperiodic Task.

1. INTRODUCTION

Multiprocessor has emerged as a powerful computing means for real-time applications such as avionics control and nuclear plant control, because of their capability for high performance and reliability. The problem of multiprocessor scheduling is to determine when and on which processor a given task executes. This can be done either statically or dynamically. In static algorithms, the assignment of tasks to processors and the time at which tasks start execution are determined a priori. Static algorithms are often used to schedule periodic tasks with hard deadlines. The main advantage of static algorithms is that, if a given task set is schedulable, then one can be sure that all deadlines will be guaranteed. However, this approach is not applicable to aperiodic tasks whose characteristics are not known a priori. Scheduling such tasks in a real-time multiprocessor system requires dynamic scheduling algorithms. In dynamic scheduling, when a new task arrives, the scheduler dynamically determines the feasibility of scheduling these new tasks without jeopardizing the guarantees that have been provided for the previously scheduled tasks. In many applications, if a new task cannot be feasibly scheduled on a site, alternatives may exist. These include sending the task to another site or invoking a less resource-intensive (and typically, lower quality) task instead. Dynamic scheduling exploits these possibilities. A dynamic scheduling can be either centralized or distributed. In a centralized scheme, all the tasks arrive at a central processor from which they are distributed to other processors in the system for execution. In this paper, we will consider the centralized dynamic scheduling scheme.

* The National Science Foundation of China (NSFC) under the grant No 60073003 supported this work.

It was shown in the paper [1] that there does not exist an algorithm for optimally scheduling dynamically arriving tasks with or without mutual exclusion constraints on a multiprocessor system. This negative result motivated the need for heuristic approaches for solving the scheduling problem. [2] and [3] proposed the heuristic search algorithms for multiprocessor systems with tasks having resource constrained, called myopic and thrift, respectively. The authors of [2,3] have shown that an integrated heuristic of the deadline and the earliest start time of a task perform better than simple heuristic, such as EDF, least laxity first, and minimum processing time first. A variant of the myopic was proposed for the parallizable tasks scheduling presented in the paper [4]. The paper [5] modified myopic so that it can be adapted to systems with hard real-time and soft real-time tasks. [6] and [7] extended the thrift so that it can be used to heterogeneous systems and the one with hard real-time and soft real-time tasks, respectively.

The myopic and the thrift include two policies. One is the task selection policy and the other is the processor selection policy. The main difference between the myopic and the thrift is the policy to select a processor. The myopic assigns a task to a processor with the minimum earliest available time. The thrift firstly obtains the set of processors that can guarantee the deadline of a task. And then the task will be assigned to the processor with the latest earliest available time in this processor set. Their task selection policies are same, which is to select the task with the minimum H to extend.

In this paper, a new dynamic algorithm, named GP (Group-Proper), is developed. GP consists of two policies. A task selection policy and a processor selection policy, named grouping policy and properly choosing policy (PCP), respectively. Grouping policy aims to increase the effective utilization of resources, while the aim of PCP is to enlarge the effective utilization of processors. The objective of GP is to increase the performance of the system.

The rest of the paper is organized as follows. In section 2, the models of scheduler, task, definitions and related works are presented. In section 3, the new algorithm GP is developed. In section 4, simulation results are presented to show the performance of theses three algorithms. Finally, main conclusions are drawn in section 5.

2. MANUSCRIPTS SCHEDULER AND TASK MODELS

Scheduler Model

In this paper, we assume a centralized scheduling scheme. In a centralized scheme, all the tasks arrive at a central processor called the scheduler, from which they are distributed to other processors in the system for execution.

The communication between the scheduler and the other processors is through dispatch queues. Each processor has its own dispatch queue. This organization ensures that the processors will always find some tasks in the dispatch queues when they finish the execution of their current tasks. The scheduler will be running in parallel with processors, scheduling the newly arriving tasks, and periodically updating the dispatch queues. For this parallel operation the scheduler has to ensure that the dispatch queues always find some tasks in the dispatch queues when they finish the execution of their current tasks. The scheduler will be running in parallel with processors, scheduling the newly arriving tasks, and periodically updating the dispatch queues. For this parallel operation the scheduler has to ensure that the dispatch queues always have a minimum of tasks (if there are tasks left with it).

Task Model

We assume that the system only consists of aperiodic task with m processors and s resources, where $m>1$. Each task in this real-time system is a dispatcher entity.

- Each task T is characterized by the arrival time (a_T), the release time (r_T), the worst-case processing time (c_T), and the deadline time (d_T).
- A task cannot be parallelized.
- Tasks are independent, aperiodic and non-preemptive.

Tasks may use other resources except processors. A task uses a resource either in shared mode or in exclusive mode and holds the requested resource as long as it executes.

Definitions

Definition 1. A Task is feasible if its timing and resource requirements are met in the schedule. A schedule is said to be a feasible schedule if all tasks are feasible in the schedule.

Definition 2. A partial schedule is a feasible schedule for a subset of tasks. A partial schedule is said to be strongly feasible if all the schedules obtained by extending the current schedule by any one if the remaining tasks are also feasible.

Definition 3. EAT_k^s (EAT_k^e) is the earliest time when a resource R_k becomes available for shared (or exclusive usage).

Definition 4. Let P be the set of processors, and R_i be the set of resources requested by task T. Earliest start time of a task T, denoted as $EST(T)$, is the earliest time when its execution can be started, defined as : $EST(T) = MAX(r_T, \min_{j\in P}(avail\ time(j)), \max_{R_k\in R_T} EAT_k^u)$, where $avail\ time(j)$ denotes the earliest time at which the processor P_j becomes available for executing a task, and the third term denotes the maximum among the earliest available times of the resources requested by T, in which $u=s$ for shared mode and $u=e$ for exclusive mode.

3. GP ALGORITHM

In the paper, we will present a new algorithm called GP to obtain a higher performance. GP consists of two policies. One is the grouping policy and the other is the properly choosing policy. Grouping policy and properly choosing policy aims at maximizing the effective utilization of resources and processors, respectively.

Properly Choosing Policy

PCP is described by the following.

- If task T only needs a processor. T is assigned to the processor with the minimum earliest available time.
- If T needs other resources except processors.
  If the earliest available time of resources required by T is less than the minimum earliest available time of processors, T is assigned to the processor with the minimum earliest available time.
  Otherwise, T is assigned to the processor with the earliest available time closest to that of resources required by T.

Grouping Policy

The main idea of grouping policy is to select the task with the minimum H value accessing resources in shared mode provided that the task with the smaller H value meets its deadlines. Each resource R_r has one corresponding shared group GR_r. Tasks using a same resource in shared mode are grouped into the corresponding shared group. We know that tasks in a shared group can be executed parallel. In order to maximize the effective utilization of resources, we introduce a new term overlap ratio to represent the overlapped time that a task uses resource and the other tasks access it.

Firstly, we will describe the term overlap ratio.

The earliest available time of a resource R_r accessed by T can be computed using EAT_r^s and EAT_r^e. The earliest start time of a task T is $EST(T)$. If the task T accesses resource R_r in shared mode, EAT_r^s will not be changed and EAT_r^e=max (EAT_r^e, $EST(T)+C_T$). If the task T accesses resource R_r in exclusive mode, $EAT_r^s=EAT_r^e$=max (EAT_r^e, $EST(T)+C_T$). Supposing a task T_a using R_r in shared mode is to be scheduled, it is obvious that only if $EAT_r^s \neq EAT_r^e$ and $EAT_r^s \leq EST(T)<EAT_r^e$, the time that the task T_a uses the resource R_r in shared mode overlaps the time that other tasks uses the resource R_r in shared mode, which have been scheduled.

Let $overlap_{R_r}(T)$ be $\dfrac{\text{The overlapped time that } T \text{ uses } R_r \text{ in shared mode and } R_r \text{ is accessed by other tasks in shared mode}}{\text{The time that } R_r \text{ is accessed by other tasks in shared mode}}$

The method to compute $overlap_{R_r}(T)$ is the following.

When $EAT_r^s=EAT_r^e$, $overlap_{R_r}(T)=0$.

When $EAT_r^s \neq EAT_r^e$ and $EAT_r^s \leq EST(T) <EAT_r^e$,

$$overlap_{R_r}(T) = \begin{cases} \dfrac{EAT_r^e - EST(T)}{EAT_r^e - EAT_r^s} & EST(T)+C_T \geq EAT_r^e \\ \dfrac{C_T}{EAT_r^e - EAT_r^s} & otherwise \end{cases}$$

Below, grouping policy is described in detail.

1. Initially, select the task with the minimum H to extend and determine whether the partial schedule

is strongly feasible. If the partial schedule is strongly feasible, then moves this task in the task queue, adds it into the corresponding dispatch queue, and modifies the earliest available times of processors and resources.

2. Select one of a task in the feasibility check window. Let T be the task having the minimum H in the feasibility check window.
   - If none in the feasibility check window is in the shared groups, then T is selected to extend.
   - T is selected to extend, if there isn't any task that has the following properties. It uses resources (a resource) in shared mode, and its overlap ratio is larger than the overlap ratio factor q, where q is an input parameter. If there is at least one task having the above properties, then selects the shared group having the largest overlap ratio within all the shared groups. Thereby, a task T' with the minimum H among the tasks having overlap ratios larger than q in this shared group is determined. If T' is the task T, then T is selected to extend. Otherwise, pre-choosing is needed. Pre-choosing determines whether T is able to meet its deadline if T' is extended. If T is unable to meet its deadline, then T is selected to extend. Otherwise, T' is selected to be extend.

To reduce the error numbers of pre-choosing, grouping policy requires $overlap_{R_r}(T) \geq q$. Otherwise; the performance of the group policy will be low.

To implement the grouping policy does not need to construct the shared groups for resources. Assume a system has s resources. Every resource R_r has two variables, named $max\_overlap_R(r)$ and $overlap_R(r).task$. $max\_overlap_R(r)$ records the maximum overlap ratio of tasks in the shared group GR_r. $overlap_R(r).task$ records the task with the minimum H among the tasks having overlap ratios larger than or equal to q in the shared group GR_r. At the beginning computing H of tasks in the feasibility check windows, $max\_overlap_R(r)$ and $max\_overlap_R(r).task$ of the resource R_r $(1 \leq r \leq s)$ are 0 and null, respectively. While computing H of the tasks in the feasibility check window, $overlap_{R_r}(T)$ of the task T in it can be checked whether larger than or equal to q. If $overlap_{R_r}(T) \geq q$ and $max\_overlap_R(r)=0$, then $max\_overlap_R(r)= overlap_{R_r}(T)$ and $overlap_R(r).task=T$. If $q \leq overlap_{R_r}(T)$ and $overlap_{R_r}(T) > max\_overlap_R(r)$, then $max\_overlap_R(r)= overlap_{R_r}(T)$.

By this way, $max\_overlap_R(r)(1 \leq r \leq s)$ of each shared group is computed. Hence, the shared group GR_i with the maximum overlap ratio among all shared groups can be known. $overlap_R(i).task$ represents the task which has the minimum H among the tasks in GR_i with the overlap ratio larger than or equal to q. If $overlap_R(i).task$ is null, then chooses the task T with the minimum H in the feasibility check window to extend. Otherwise, checks whether extending $overlap_R(i).task$ results in T missing its deadline. If T misses its deadline, then T is selected to extend. Otherwise, $overlap_R(i).task$ is selected to extend.

GP Algorithm

Starting at the root of the search tree that is an empty schedule, GP tries to extend the schedule (with one more task) by moving to one of the vertices at the next level in the search tree until a full feasible schedule is derived. GP uses grouping policy to select a task to extend the current partial schedule. PCP is used to select a processor to assign a selected task on. If the current vertex is not strongly feasible, the algorithm backtracks to the previous search point. If the previous schedule selects task T with the minimum H in the feasibility check window, then this schedule selects the task with the next minimum heuristic value to extend. If the previous schedule selects the task in a shared group that is not T, this schedule selects T to extend. The termination condition is either (a) a complete feasible schedule has been found, or (b) reached the maximum number of backtracks, or (c) no more backtracking is possible.

Assume k, H, n, m and s represent the feasibility check window, the heuristic function value, the number of tasks, the number of processors, and the number of resources, respectively. GP has a total of n steps. Each step uses the grouping policy, the properly choosing policy, computing the heuristic function H and determining whether a partial schedule is strongly feasible. The complexity of computing H in the feasibility check window is O (k). Choosing a task to extend includes selecting the task with the minimum H in the feasibility check window and selecting the task with the minimum H among the tasks with overlap ratio larger than or equal to q in the feasibility check window. If no need of pre-choosing or the result of pre-choosing is the task in the shared group, the complexity of choosing a task is O (k+s). Otherwise, the complexity of selecting a task is O (k+s) +O (1). Whether a partial schedule is strongly feasible can be determined at the same time of choosing a task. Hence, the complexity of the grouping policy and determining whether a partial schedule is strongly feasible is O (k+s) +O (1). The properly choosing policy selects a processor among m processors. Thereby, its complexity is O (m)

For $O(k)+O(k+s)+O(1)+O(m) \leq O(2k+s+1+m)$, the complexity of GP is $O((2k+s+1+m)n)=O(K'n)$). Moreover, m and s are constants and they are often much less than the number n of tasks. Hence, the scheduling cost is linearly proportional to the number of tasks that remains to be scheduled. Thereby, the complexity of GP is O (n^2), which is equal to that of the myopic and the thrift.

4. SIMULATION STUDIES

In dynamic scheduling, when a new task arrives, the scheduler dynamic determines the feasibility of scheduling these new tasks without jeopardizing the guarantees that have been provided for the previously admitted tasks. If the task is unfeasible, the task will immediately be rejected. Otherwise, the task will be scheduled to meet its deadline. Therefore, the performance metric used to evaluate the effectiveness of these three scheduling algorithms is the guarantee ratio defined as the ratio of the number of tasks in a given set that are admitted by the scheduler compared to the number of tasks in this set.

We have conducted extensive simulation results to show the performance of algorithms.

The method to generate task sets is the same as described in [2]. Tasks (of a task set) are generated till schedule length, which is an input parameter, with no idle

time in the processors. The computation time of a task *T* is randomly chosen between *Max_exec* and *Min_exec*, where *Max_exec* and *Min_exec* represent the maximum processing time of tasks taken as 60 and the minimum processing time of tasks taken as 30, respectively. The arrival time of a task *T* is the minimum earliest available time of processors. The deadline of a task *T* is randomly chosen in the range [(1+*D*)**f*,(1+*D*)**SC*], where *f*, *SC* and *D* are the shortest completion time, the largest completion of the task set generated in the previous step, and laxity of task (which denotes the tightness of the deadline), respectively. The resource requirements of a task are generated based on the input parameters *Use_P* and *Share_P*, where *Use_P* and *Share_P* are the probability that a task uses a resource and the probability that a task uses a resource in shared mode, respectively. In the experiments, the release time of a task is equal to its arrival time.

Each point in the performance curves in the figure 6~11 is the average of ten simulation runs, each with 200 task sets. Each task set contains approximately 200 to 300 tasks by fixing the scheduling length to 2000. For all the simulation runs, the number of processors is 5 and the number of resource is taken as 2, each having 1 instance.

The effect of overlap ratio for GP

Firstly, simulations run to get the value of *q* so that the performance is GP is best.

For the case that *Use_P* and *Share_P* are low (<0.5), the value of *q* does not have much effect on the performance of GP. Therefore, this paper only plot the simulation results taken high *Use_P* and *Share_P*. Figure 1, 2, 3 and 4 show the simulation results when *Use_P* and *Share_P* are high. We can know from these figures that GP always achieves high guarantee ratio, as *q*=0.5. Hence, the following simulation will take *q* as 0.5.

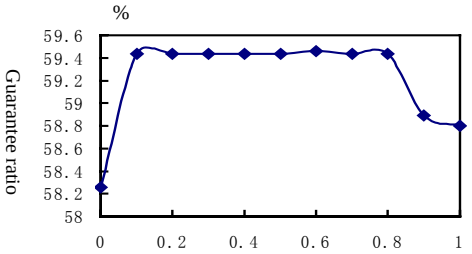


Fig. 1 The effect of overlap ratio of GP *Use_P*=1, *Share_P*=0.6,*K*=3,*W*=1,*Backnum*=∞

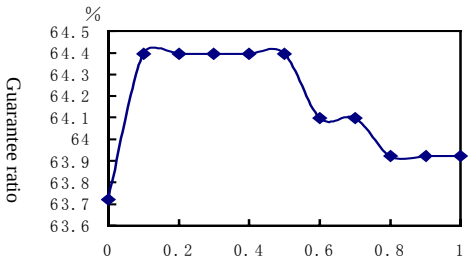


Fig. 2 The effect of overlap ratio of GP *Use_P*=1, *Share_P*=0.7,*K*=3,*W*=1, *Backnum*=∞

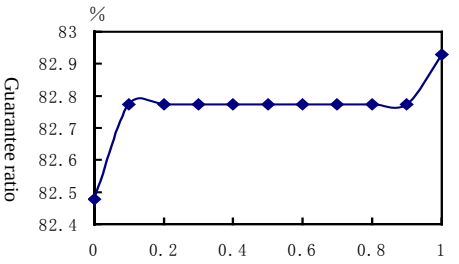


Fig. 3 The effect of overlap ratio *q* of GP *Use_P*=0.6, *Share_P*=0.7,*K*=3,*W*=1, *Backnum*=∞

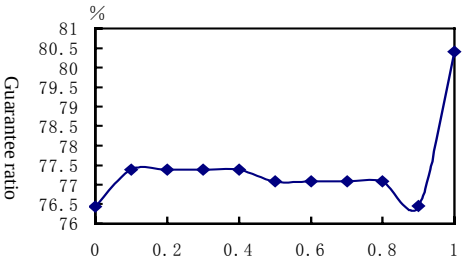


Fig. 4 The effect of overlap ratio *q* of GP *Use_P*=0.6, *Share_P*=0.6,*K*=3,*W*=1, *Backnum*=∞

Number of Backtracks

To compare the impact of the number of backtracks on their performance, we assume the number of backtracks of these three algorithms are unlimited. Moreover, in the following simulations, the number of backtracks will be added to get the total numbers of backtracks needed by an algorithm.

The mean total numbers of backtracks needed by these three algorithms are listed in table 1 for various *Use_P*.

The mean total numbers of backtracks needed by these three algorithms are listed in table 2 for various *Use_P* and *Share_P*.

From the table 1 and the table 2, we can observe that the mean total numbers of backtracks of the myopic and GP are approximately similar and the mean total number of backtracks of the thrift is much higher than that of myopic and GP.

Because of the large number of the backtracks needed by the thrift, the number of the backtracks is assumed to be unlimited to compare the performance of these three algorithms in the following simulations.

Effect of W, k, Use_P and D

In fig. 5, 6, 7 and 8, the impact of *W*, *k*, *Use_P* and *D* are plotted, respectively. From fig. 5, 6, 7 and 8, it is clear that the guarantee ratio of GP is always higher than that of the myopic and the thrift.

5. CONCLUSIONS

In this paper we have proposed a new algorithm for non-preemptive scheduling of aperiodic tasks that possesses good properties such as high guarantee ratio, in a multiprocessor real-time system. This algorithm includes a task selection policy and processor selection policy, named grouping policy and properly policy, respectively. Simulation results show that this algorithm outperforms the myopic and the thrift.

Table 1. Mean total numbers of backtracks for *Share_P*=0.5, *K*=4, *W*=1 and *D*=0.1

Algorithm	*Use_P*										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Myopic	0	0	0.5	0.9	1.6	1.8	2.6	2.6	2.9	2.6	3.3
Thrift	24.9	33.4	47.4	53.9	66.8	79.7	90.8	111.4	122.2	139.7	154
GP	0	0	0.5	0.9	1.6	1.8	2.6	2.6	2.9	2.6	3.2

Table 2. Mean total numbers of backtracks for *K*=4, *W*=1, *D*=0.1

Algorithm	*Use_P* and *Share_P*										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Myopic	0	0.02	0.39	1.07	1.52	1.97	2.3	2.54	2.75	2.95	2.94
Thrift	27.39	33.01	50.25	61.57	73.48	87.68	101.61	101.61	134.95	152.32	169.61
GP	0	0.02	0.39	1.07	1.52	1.96	2.3	2.54	2.75	2.94	2.94

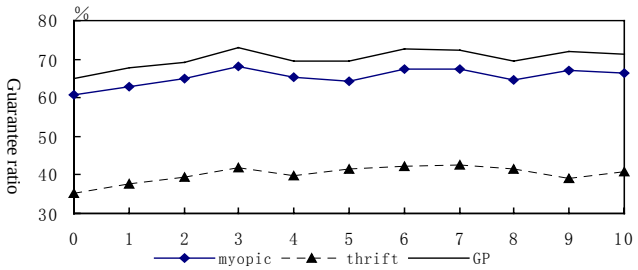


Fig. 5 Effect of *W* parameter
Use_P=0.7, *Share_P*=0.5, *K*=4,*D*=0.1, *Backnum*=∞

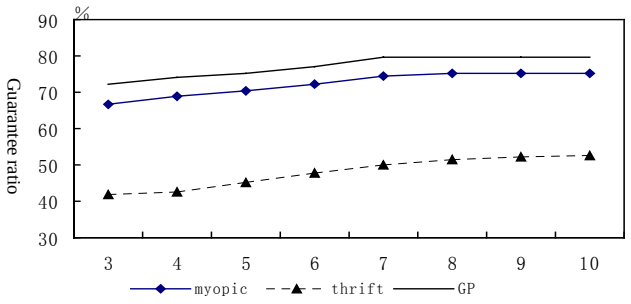


Fig. 6 Effect of size *k* of feasible window
Use_P=0.7, *Share_P*=0.5, *W*=1,*D*=0.1, *Backnum*=∞

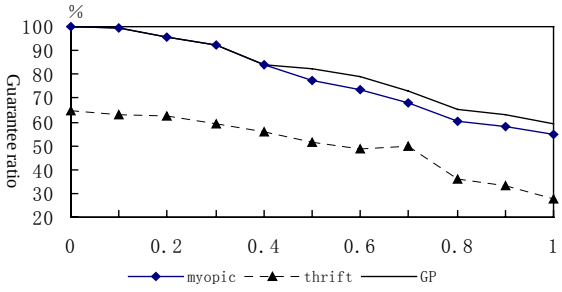


Fig. 7 Effect of resource usage *Use_P*
Share_P=0.5, *K*=4,*W*=1,*D*=0.1,*Backnum*=∞

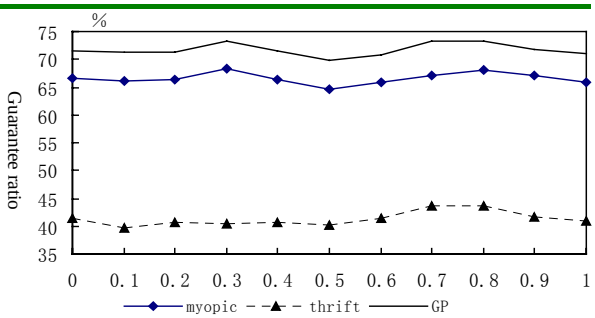


Fig. 8 Effect of laxity parameter *D*
Use_P=0.7, *Share_P*=0.5, *K*=3,*W*=1,,*Backnum*=∞

6. REFERENCES

[1] M. L. Dertouzos, A Kmok, “Multiprocessor on-line scheduling of hared real-time tasks”, IEEE transaction on software engineering, 1989, 41(12): 1497-1506.

[2] K. Ramamritham, J. Stankovic, “Efficient scheduling algorithms for real-time multiprocessor systems”, IEEE transactions on Parallel and distributed systems, 1989, 1(2): 184-194.

[3] Y. Qiao, H.A. Wang, et al, “Developing a new dynamic scheduling algorithm for real-time multiprocessor systems”, Journal of Software (in Chinese), 2002, 3(1): 51-58.

[4] T. Manimaran, C.S.R Murthy, “An efficient dynamic scheduling algorithm for multiprocessor real-time systems”, IEEE Transactions on Parallel and Distributed Systems, 1998, 9(3): 312-319.

[5] A. Mittal, G. Manimaranet, et al, “Integrated dynamic scheduling of hard and QOS degradable real-time tasks in multiprocessor systems. In: Mok, A.K., Tokuda, H... eds”, Proceedings of the 5th international conference on real-time computing systems and applications. Los Alamitos, CA; IEEE press, 1998, 162-172.

[6] K. Wang, Y. Qiao, et al, “Study of a dynamic scheduling algorithm for real-time heterogeneous systems (in Chinese)”, 2002, 39(6): 725-731.

[7] Y. Qiao, B. Zou, et al, “Design and evaluation of an algorithm for integrated dynamic scheduling in real-time heterogeneous systems”, Journal of Software (in Chinese), 3(12): 2251-2258.

ForTER – An Open Programmable Router Based on Forwarding and Control Element Separation*

Weiming Wang, Ligang Dong, Bin Zhuge
Institute of Networks and Communications Engineering, Zhejiang Gongshang University
149 Jiaogong Rood, Hangzhou, P.R.China, 310035
Email: {wmwang, donglg, zhugebin}@mail.zjgsu.edu.cn

ABSTRACT

ForCES (Forwarding and Control Element Separation) shows great advantages for routers to be flexible, programmable, and cost-effective. An implementation of a ForCES-based router (ForTER) was illustrated. The implementation architecture was first discussed. A software model for the ForTER implementation was presented. Based on the model and related IETF ForCES specifications, Control Element (CE) and Forwarding Element (FE) in the ForTER were designed and implemented. An algorithm to prevent ForCES redirect data DoS attacks coming from a UDP channel was introduced. The algorithm tried to bind UDP stream to TCP stream for congestion control when transmitting ForCES protocol messages from FE to CE. Experiment results shown the feasibility of the system architecture and the software model of ForTER, and shown the effectiveness of the DoS attack prevention algorithm. Experiments also shown the feasibility of ForTER as a router. As a result, the implementation of ForTER testifies the feasibility of ForCES and related specifications, which are being standardized by IETF ForCES working group.

Keywords: ForCES, Open programmable, router, protocol.

1. INTRODUCTION

Several demands from the Internet service providers and network end-users have greatly impacted upon the designs of next generation network equipments (formally called as Network Elements, NEs) [1]. The demands for next generation NEs can be described as: 1)being flexible enough for the deployments of ever growing new services; 2)being open and modular enough so that the equipment market is hard to be monopolized, so as to low the equipment prices; 3)being capable of providing QoS services in order to support real-time applications such as multimedia transmissions.

Open programmable networks [2, 3] is considered as the most prospective architectural approach to meet above demands. In open programmable networks, a NE (e.g., a router/switch) is systematically separated into a control plane and a forwarding plane. Forwarding plane receives packets from outer networks, processes the packets according to functional requirements of the NE, and then outputs the packets back to outer networks. Forwarding plane usually needs the ability to process packets at line speed. Control plane controls forwarding plane for the whole forwarding process, and provides adequate parameters for the process. More importantly, the interface between the control plane and the forwarding plane is standardized. Moreover, resources at the forwarding plane, which are used to process packets, are also described in a standardized way. As a result, control plane can access and control the forwarding plane resources in a standard way. This makes it feasible for control plane and forwarding plane to be separated at their product level, i.e., control plane and forwarding plane as separate products from different vendors can work together to form one NE with full interoperability.

Accordingly, researches on open programmable networks are focused on the way to standardize the interface between control plane and forwarding plane and to setup a model to standardize the resources in forwarding plane.

Early ideas in Opensig [8], Active Networks [2], and Virtual Networks [6] helped forming the basic idea of open programmable networks. Opensig emphasizes the standardization of interfaces, active networks more on the programmability of network functions, and virtual networks on providing QoS by allocation or reservation of network resources.

Works in the past years related to the standardization of interfaces and resources in open programmable networks include IEEE P1520 [16], General Switch Management Protocol (GSMP) [13], Multi-service Switching Forum (MSF) [7], and Forwarding and Control Element Separation (ForCES) [10]. Amongst the works, the ForCES, which is a working group under Routing Area of IETF, has made the most prominent achievements for open programmable networks. It has been widely accepted as currently the most typical solution for open programmable networks. ForCES is also supported by works from Network Processing Forum (NPF) [9], ITU-T NGN Focus Group [5, 12], and Intel IXA [11].

ForCES proposes that an Network Element (NE) (typically a router) consists of several Control Elements (CEs), among which one acts as an active controller and others as backups, and multiple Forwarding Elements (FEs). Multiple FEs in forwarding plane are interconnected together, forming a distributed forwarding plane FE topology to fulfill complex packet forwarding tasks. A standardized interface called ForCES protocol [15] is defined for the communication between CEs and FEs. Up to the present, the ForCES group has completed the work of ForCES requirements (RFC 3654) and ForCES Framework (RFC 3746). The works for ForCES Protocol Specification (i.e., ForCES protocol) [15] and ForCES FE Model [4] are also near its completion.

Presented in the paper is an implementation of an open programmable router that is based on ForCES architecture and ForCES protocol. The implementation is called ForTER – a ForCES-based rouTER. Key design issues for ForTER are discussed and implementations of the key elements in it are presented in details in this paper.

* Project sponsored by National 863 Project (2005AA121310), NSF China (60573116), and Zhejiang Key Sci & Tech Project (2005C21013)

Background of this paper is highly related to IETF ForCES work. As key participants in the ForCES work and authors of the ForCES protocol, we set the motivation of the ForTER implementation as for the evaluation of the ForCES protocol and the associated ForCES architecture. As a proposed standard for Internet, ForCES is currently in urgent need of such an implementation to evaluate its protocol designs.

Section 2 described the architecture of ForTER. Section 3 presented the implementation details of key elements in ForTER. Section 4 introduced some experiments and tests results.

2. ARCHITECTURE

2.1. Forwarding and Control Element Separation

Forwarding and Control Element Separation is the key for the architecture of ForTER. We show the architecture as in Fig. 1. , which complies with the ForCES Framework defined in RFC 3746. In the architecture, there exists one primary Control Element (CE). There may also be some redundant CEs for system High Availability (HA) purpose. There are multiple Forwarding Elements (FEs) (may be up to hundreds of FEs for core routers), which are separated from CEs. The interface between CE and FE is to be standardized and the communication protocol upon it is called the "ForCES protocol". The CE is responsible for the management of FEs by use of the ForCES protocol. The Fr is the interface connecting CEs, and the Fi connecting FEs. Outer networks connect to ForTER via Fi/f interfaces. After ForCES protocol has been standardized, it will make possible CE and FE being manufactured by different vendors.

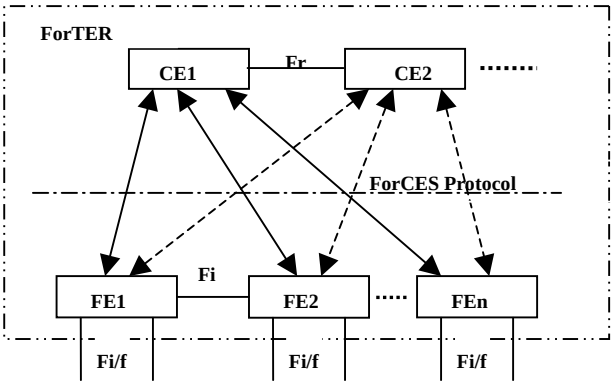


Fig. 1. Architecture of ForTER

FE is further standardized by ForCES FE Model [4]. The resources in an FE are represented by various kinds of standardized Logical Function Blocks (LFBs), each of which has specific functions for IP packet forwarding. Typical LFBs are that like classifier LFB, scheduler LFB, and IPv4 or IPv6 forwarder LFB. Multiple LFBs are interconnected via datapaths to form an LFB topology in FE, so that the FE can carry out a complex process on packets forwarding purpose. CE is responsible for the management of LFBs in FE. The management of LFBs includes the configuration and inquiry of LFB attributes, capabilities, or events. ForCES protocol makes it possible for CE to dynamically manage the LFBs and the associated LFB topology, such as to add/remove/modify some LFBs, the attributes, and the associated LFB topology. In this way, CE is able to provide various IP forwarding services in a "on the fly" way and by configuring the very bottom resources of FE. Note that this management ability of CE to FE resources is far beyond the management ability that general router console or IOS can provide. This makes ISPs able to configure new services in a very handy, economic, and flexible way.

Mostly, FE forwards packets from some Fi/f to some other Fi/f, but in some cases, FE also have to redirect some packets to CE. Packets of these are usually packets for routing protocols or network management protocols. These packets will be encapsulated as ForCES protocol messages and redirected to CE then. A special security issue called DoS attack from redirect messages is induced by the redirect packets. Section 3.3 describes it and provides a solution for it.

2.2. Layered Model

IETF ForCES proposed the architecture and the protocol for forwarding and control element separation. However, a detailed software model based on the architecture is essential for any ForCES implementations and has not been presented yet. We present here a layered software model which is applied to ForTER and is also fit for other ForCES implementations.

The layered software model of ForTER is shown is Fig. 2. As in the figure, the model is coarsely layered as CE and FE, while FE is divided into FE SlowPath and FE FastPath. FE FastPath is responsible for the packet processing that needs running at line speed. The behavior of FE FastPath is flexibly controlled by CE via FE SlowPath.

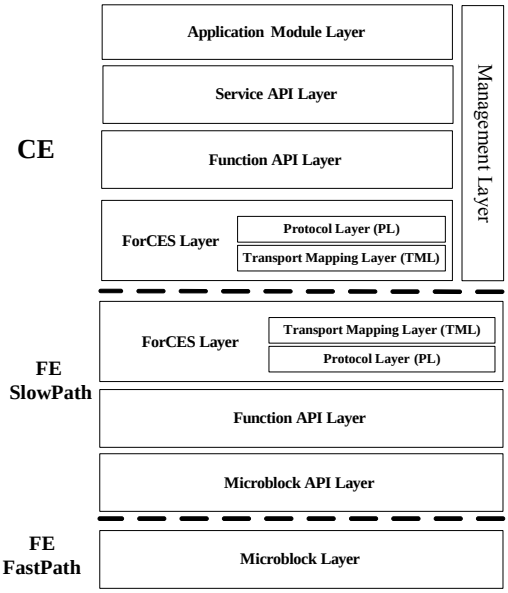


Fig. 2. Layered Model of ForTER

The Application Module Layer inside the CE contains independent software modules for ForCES applications. Typical application modules include routing protocol modules, network management module, modules for QoS, and modules for other services. Routing protocol module includes routing protocol stack software like RIP, OSPF, etc. Network management module provides core software for network management protocols. The modular design makes it able to widely use third-party software modules as the ForCES application modules. In the ForTER

implementation, we adopt Zebra routing module as the routing protocol module, and AgentX++ and NetSNMP as the SNMP module.

Service API Layer in CE is a bridge between upper softwares and ForCES-specific layers. On the one hand, Service APIs in CE convert a service (e.g., DiffServ) from upper application softwares to a sequence of operation on LFBs. On the other hand, Service APIs in CE provides call back functions, which are interfaces to third-party network software, for underlying function APIs in CE. In this way, Service API Layer has hidden the feature of ForCES for upper application and third-party network software. Service API Layer provides sets of function calls and data structures for various upper softwares, e.g., sets of function calls for SNMP, routing, DiffServ, etc. Service API Layer is implemented based on APIs in underlying Function API Layer.

Function API Layer in CE contains function calls and data structures for direct operation (i.e., configuration, inquiry, event report) on every attribute and capability of all the LFBs. On the one hand, Function APIs in CE will call functions in ForCES Layer to realize actual operation on LFBs in FE, on the other hand, functions in Function API Layer of CE will be called by ForCES Layer when received protocol messages from FE are executed.

The CE ForCES Layer includes PL（Protocol Layer）and TML（Transport Mapping Layer）. PL provides generation and execution of ForCES protocol messages for certain operations on LFBs in FEs. TML supports the transportation of ForCES messages over various transport medias, like Ethernet, IP, ATM, etc. In most cases, The TML is an IP based media.

A CE Management Layer is set for the purpose of the whole CE management. The layer traverses all CE functional layers， providing a management interface for all the layers. In ForTER, we set a GUI interface for CE management purpose, which provides friendly interface for management of the whole ForCES Network element.

In FE, the FE ForCES Layer is identical to the ForCES layer in CE, which also consists of PL and TML sub layers for ForCES protocol management.

According to ForCES FE Model[4], LFBs in FE are logical functional blocks that are abstracted and synthesized from FE physical resources. In the layered model of ForTER as in Fig. 2. , the physical resources in FE are described by microblocks. A microblock is a physical component with a single function, which usually runs at line speed to process packets, hence belongs to FE fastpath layer. To summarize, microblocks are described more from the implementation and physical point of view and not directly related to the ForCES structure, while LFBs in ForCES are modeled more from the logically abstract point of view. LFBs in FE are constructed based on the microblocks.

Microblock API layer provides management APIs to underlying microblocks. The layer will exploit as much as possible programmability from the microblocks. The programmability of microblocks is an essence for programmability of whole FE.

A Function API Layer is necessary for connecting ForCES Layer and Microblock API Layer in FE. The layer implements the converting task between LFB operations and micriblock operations.

The layered model in Fig. 2. is featured as being layered and modular, which helps exploiting the generic features of ForCES as being open, programmable, modular, and standard.

3. DESIGN AND IMPLEMENTATION

ForTER is implemented based on the layered model shown in Fig. 2. We further present the design and implementation of elements of ForTER. We also present a security consideration regarding DoS attacks and propose an algorithm to protect the attacks.

3.1. Control Element (CE)

Fig. 3. shows the implementation structure of CE in ForTER. The main platform of CE is based on Windows OS, while third-party network software (including routing protocol software Zebra and SNMP software NetSNMP) run on a separate Linux OS platform, which is called the Routing & SNMP server.

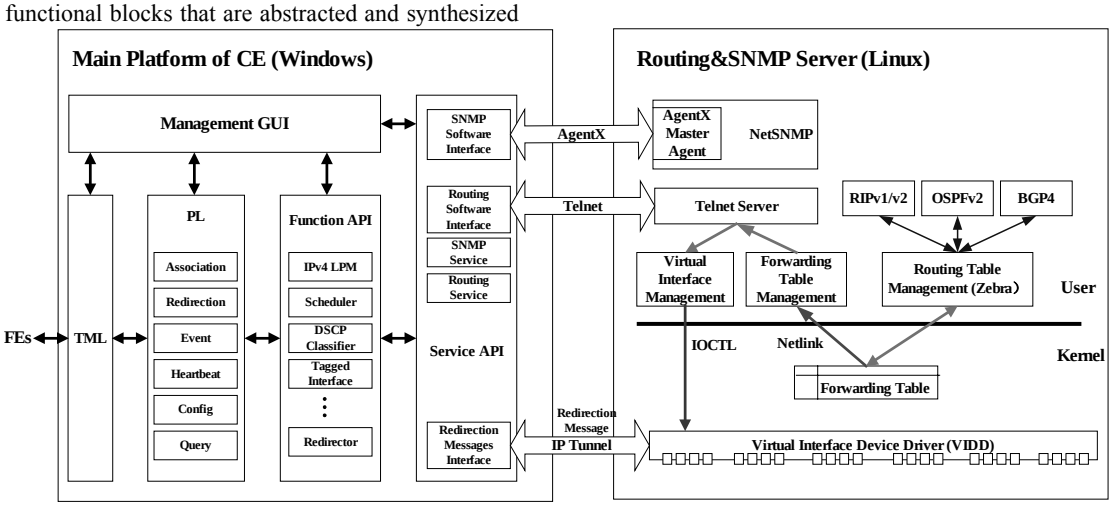


Fig. 3. Structure of Control Element

As described in the layered model, CE PL supports the generation, encapsulation, decapsulation, and processing of ForCES protocol messages. The messages and associated modules include: 1) Messages and for association, e.g., AssociationSetup message, AssociationSetupResponse message, AssociationTeardown message; 2) Messages for configure and query, e.g., Config message, ConfigResponse message, Query and QueryResponse message; 3) Event notification message; 4) Heartbeat message, 5) Redirection message. Redirection message is the message that contains

data packets that are to or from outer networks via FEs but that need to be processed by the CE, including packets for routing protocols and SNMP protocol. ForCES just encapsulates these packets directly in the redirection message and transport them between CE and FEs. Among the messages above, 1) to 4) messages are called ForCES control messages. Different sub-modules in PL are used to process above different ForCES messages, like the sub-module for association, the sub-module for redirection, as shown in Fig. 3.

We use the TCP+UDP based TML, which is based on a proposed ForCES TML specification [14] that we have submitted to IETF, as the TML scheme for ForTER. In this scheme, TCP is applied to the transmission of ForCES control messages and UDP to redirection message. According to ForCES requirements in RFC 3654, the transmission of control messages should be reliable and congestion-controllable, while redirection messages do not need to be reliably transmitted. We will discuss the TML security problem in Section 3.3.

CE Management Layer as described in Fig. 2 is represented by Management GUI (Graphical User Interface), as shown in Fig. 3. The GUI is provided in ForTER for management of every part in CE. In the GUI, we designed a FE description tree and a LFB topology diagram for very handy operation of FEs and the LFBs inside. An FE consists of multiple LFBs, and an LFB occupies its individual attributes and capabilities. We use XML to describe LFBs and use a tree structure to show them. Fig. 4 shows an instance for such FE and LFB description tree in the GUI, where the FE attributes like its ID, IP address, and associated LFBs like FE Object LFB and FE Protocol LFB[15] are shown. LFBs in an FE are interconnected and form an LFB topology. We represent the LFB topology in the CE GUI with a LFB topology diagram. This diagram looks like a circuit diagram, where an LFB is alike an electronic element. Fig. 5 shows an example of the LFB topology diagram, where every node represents a LFB in the FE and the lines represent the datapaths connecting them. By the diagram, we can clearly see the LFBs and their interconnection relationship inside the FE, hence provide the ability to see the functions and services of a router, which was very hard to do for traditional routers. Moreover, via ForCES configuration messages, CE is also able to dynamically add or remove some LFBs in the diagram, in order to modify the functions or services the FE provides. This makes the FE quite open at its resources level and very flexibly programmable at the functional level. For example, by configuring and modifying the LFBs and the topology, we can change the ForTER from IPv4 router to IPv4/IPv6 double stack router in a very handy way. Management GUI also provides the operation on Service API Layer, e.g., start/stop/restart of routing, SNMP, and QoS service, parameter configuration of interfaces to routing software, SNMP software, etc. We also provide dialog boxes for TML management like TML parameter (e.g., transport protocols, service ports) configurations. This actually provides the functions as a CE manager does.

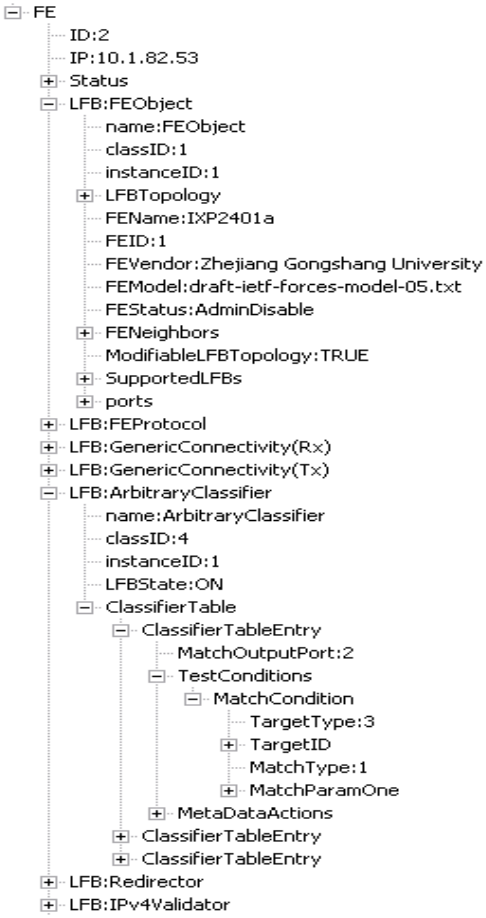


Fig. 4. Example of an FE and LFB Description tree in CE GUI

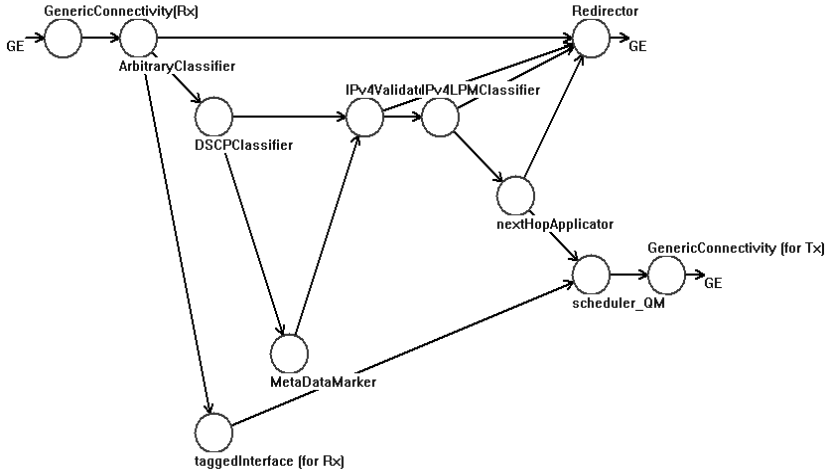


Fig. 5. Example of an LFB Topology Diagram in CE GUI

As described before, ForTER is flexible to include third-party software modules as the CE application module. In the implementation, Zebra is taken as the routing protocol software module, and NetSNMP as the SNMP module. As mentioned before, routing protocol packets and SNMP packets are encapsulated in ForCES redirection messages so as to be transported between FE and CE. The redirection messages are further transported between CE main platform and Routing & SNMP server by use of IP Tunnels, as shown in Fig. 3.

A VIDD (Virtual Interface Device Driver) is used in the Routing & SNMP server to mirror FE real network interfaces (described in Fig. 1 as Fi/f) to this server. It makes all FE network interfaces just like network interfaces locally on this server. Redirect messages that contain routing protocol or SNMP protocol messages are transported over IP tunnel to the VIDD, and the VIDD recovers them back to routing or SNMP protocol messages and outputs them to routing protocol or SNMP software modules for use. In this way, the third-party software modules can just treat VIDD as network interfaces that they interact.

3.2. Forwarding Element

ForTER implements the FE based on Intel IXP network processors. In this case, FE SlowPath and FE FastPath in the Fig. 4. layered model are exactly referred to Intel XScale Layer and MicroEngine (ME) Layer. Structure of FE in ForTER is shown as in Fig. 6.

The FE Management module in the figure is a manual interface to provide some necessary manual management for modules in the FE. The FE Management is based on CLI (Command Line Interface).

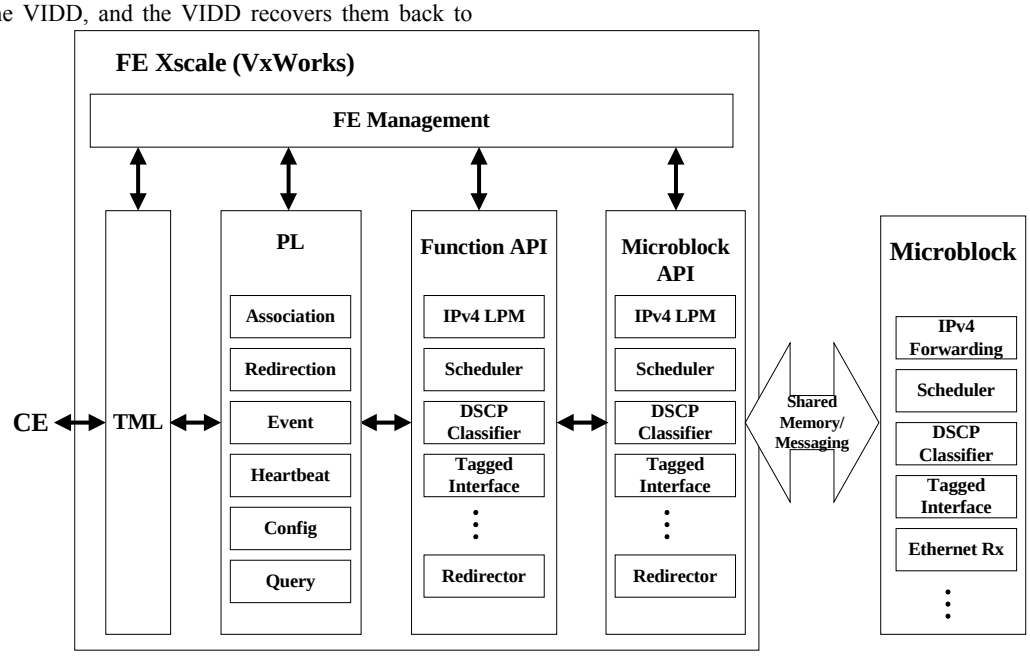


Fig. 6. Structure of Forwarding Element

Upon receiving control messages from CE, the FE may be triggered to call LFB operations in the Functions API Layer. The operations on LFBs are further converted to the operations on microblocks, which are carried out by the Microblock API Layer.

We have built Microblock API layer and Microblock layer in ForTER based on Intel IXP environment, where Microblock API Layer matches to Core Component (CC) Layer, Core Component Infrastructure Layer, and Resource Manager Layer, while Microblock layer to IXP MicroEngine(ME) layer.

A set of typical LFBs have been accordingly constructed in ForTER. The LFBs include IPv4 LPM forwarding LFB, DSCP classifier LFB, Tagged Interface LFB, scheduler LFB, etc.

3.3. Security Consideration

Because of the separation of control and forward elements, we must consider a special security problem for a ForCES router. As described before, FE may receive some packets from outer network that have to be delivered to CE for processing. In this case, FE just forwards the packets to the CE over the ForCES channel and via ForCES redirect messages. This leaves a hole for malicious attackers. Attackers may try to send a huge amount of such packets to FE and make FE forwards them to CE. As a result, and specifically because of CE and FE separation, bandwidth of the ForCES channel that connects FE and CE is easy to be exhausted, making the CE loses control of the FE, and subsequently making the ForCES router totally down. We call such attack as DoS attack from redirect data.

In Sect. 3.1, we proposed to adopt the TCP + UDP TML for ForTER. In the TML, TCP is for PL control messages and UDP for redirect messages. As is known, UDP is more aggressive than TCP, which leads to unfair bandwidth allocation and TCP congestion collapse when TCP and UDP are in the same link. In the ForCES case, this will leave space for DoS attack from redirect data.

We propose an algorithm here to avoid the DoS Attack from redirect data. In the algorithm, we try to suppress UDP stream in some way that the UDP will not completely block TCP control message transmission. Note that in the algorithm, we don't try and don't need to setup a very strict and full congestion control for the UDP stream, for the main goal for the algorithm is only to guarantee a minimum bandwidth for control message transmissions between CE and FE.

The key idea of the algorithm is to try to control UDP stream with reference to TCP stream regarding its congestion status, i.e., to roughly bind UDP to TCP for the congestion control. We call TCP and UDP binding each

other in this way as a TCP-UDP pair. In this way, we expect the UDP stream not going into a status that the UDP stream completely blocks the TCP stream.

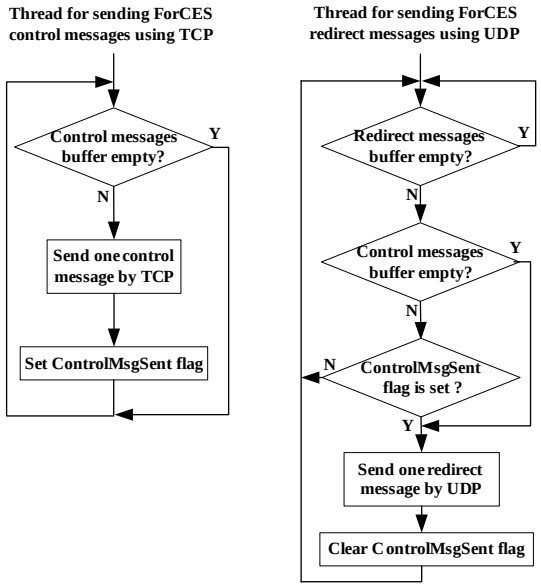


Fig. 7. Diagram of Algorithm for Protecting Redirect Data DoS Attack

Fig. 7. shows the diagram for the algorithm. Two parallel threads are established in the ForCES TML to send ForCES messages from FE to CE, one for control message sending and another for redirect message sending. Control message sending thread uses TCP as the transport protocol, and redirect message sending thread uses UDP. A sending buffer is respectively set in the threads to temporarily store sending messages. A flag called ControlMsgSent flag is used as a global variable for the two threads, indicating the state for control message sending. In the control message sending thread, whenever the sending buffer is checked none empty, a control message is sent using TCP and the ControlMsgSent flag is set. In the redirect message sending thread, the local redirect message buffer is first checked to see if there is any redirect message to send. When is, the sending buffer in the control message thread is further checked to see if, at the moment, there are control messages waiting for sending. If no, the redirect message sending thread just picks one redirect message and sends it by UDP. If yes, the thread must further check the ControlMsgSent flag state to decide if it is allowed to send a redirect message right now. When the flag is set, it actually means at least one control message has been sent before last time the redirect message was sent. In this case, the thread is allowed to send one redirect message over UDP at the moment. Otherwise, when the ControlMsgSent flag is not set, it actually means no control message has been sent since the thread sent one redirect message last time, and moreover, it indicates there is control message waiting for sending, for the control message buffer is not empty. In this case, the thread is not allowed to send any redirect message over the UDP channel and has to go back to wait for the control message sending thread to send at least one control message. Whenever a redirect message is sent, the ControlMsgSent flag is cleared so as to block the redirect message sending, till the flag has been set again by the control message sending thread.

4. EXPERIMENTS

4.1. Purposes

The nature of the ForTER implementation is for the evaluation of the IETF ForCES architecture and the associated protocols, we specifically list the purposes of the experiments we have made based on ForTER as below:

1)To see the feasibility of the ForCES architecture as defined in RFC 3654 and RFC 3746.

2)To see the feasibility of ForCES protocol as defined in ForCES Protocol Specification [15].

We go with our experiments for above purposes by means of constructing several CE basic applications. The applications include:

1) Routing function and associated routing protocols, which are fundamental for routers.

2) SNMP network management

Besides the purposes above, we also made an experiment on the algorithm presented in this paper for the protection of DoS attacks from redirect data.

The ForTER used for the experiments consists of one CE and three FEs. CE includes a CE platform on Windows and a routing & SNMP server on Linux, structured as Fig. 3. FE is structured as Fig. 6. Among the three FEs, two FEs run on Intel IXMB2401, and one FE on IXMB2851, which contain IXP2400 and IXP2850 Intel NPs respectively. CE and FEs are interconnected by ForCES protocols which run over TCP+UDP TML. An ZX5000 ATCA based switch board is used as the switch fabric to connect the three FEs..

A SmartBits600 and the TeraRouting Tester software are used as the test flow generator and the measure tool.

Followed sections explain the individual experiments and show the results.

4.2. Routing and Running Routing Protocols

As described before, routing protocols are run on the routing Server. Via ForCES protocol, the routing protocol software cooperatively works with resources on FE to make ForTER as one router to run the routing protocols.

We have implemented RIP and OSPF routing protocols on the ForTER.

A TeraRouting Tester based on SmartBits600 with a LAN3321A port module is used to test ForTER for its running of RIP and OSPF. TeraRouting Tester establishes a topology of simulation network, which contains many simulation routers to send the messages of routing protocol via SmartBits600 ports. ForTER exchanges the routing protocol messages with the SmartBits600. The CE of ForTER learns the topology of the simulated network and gains the routing table about it. Then the routing table is distributed to all FEs of ForTER. In the test scenario, another interface of SmartBits600 sends IP traffic, whose destination IP address is in the range of the simulation network to ForTER. According to the generated routing table, FEs of ForTER forward the IP traffic to its output port which is connected back to another interface of SmartBits600. If ForTER correctly runs the routing protocols tested, SmartBits600 then can receive all IP traffic which is sent by itself. Moreover, through TeraRouting Tester, lost IP packets along the cycled paths can be shown. If routing protocols are in such ForCES based architecture, it is expected the TeraRouting Tester should show the returned traffic with no lost packets. Fig. 8 shows the final results for testing the routing protocol and the routing ability of the ForTER. Area 1 of Fig. 8 shows information of all IP traffic

streams that are generated. Area 2 selects the tested routing protocol type (OSPF as in the experiment), and shows the information associated with the routing protocol. At last, area 3 shows the state of IP traffic streams that are making round trips between Smartbit and ForTER, in which the lost frames and the % loss show the loss state for the streams. Our experiment has shown no packet loss here. This can only be achieved when the ForTER has run the routing protocol well, and has correctly produced routing tables and forwarded packets according to the tables in the FEs. ForCES protocol is the basis for the CE and FEs to interact to run the whole routing protocol, as a result, this routing ability and routing protocol experiment has also well evaluated the feasibility of ForCES protocol and ForCES architecture for routers. ForCES protocol messages like configure message, query message, association message, redirect message are frequently used in the process.

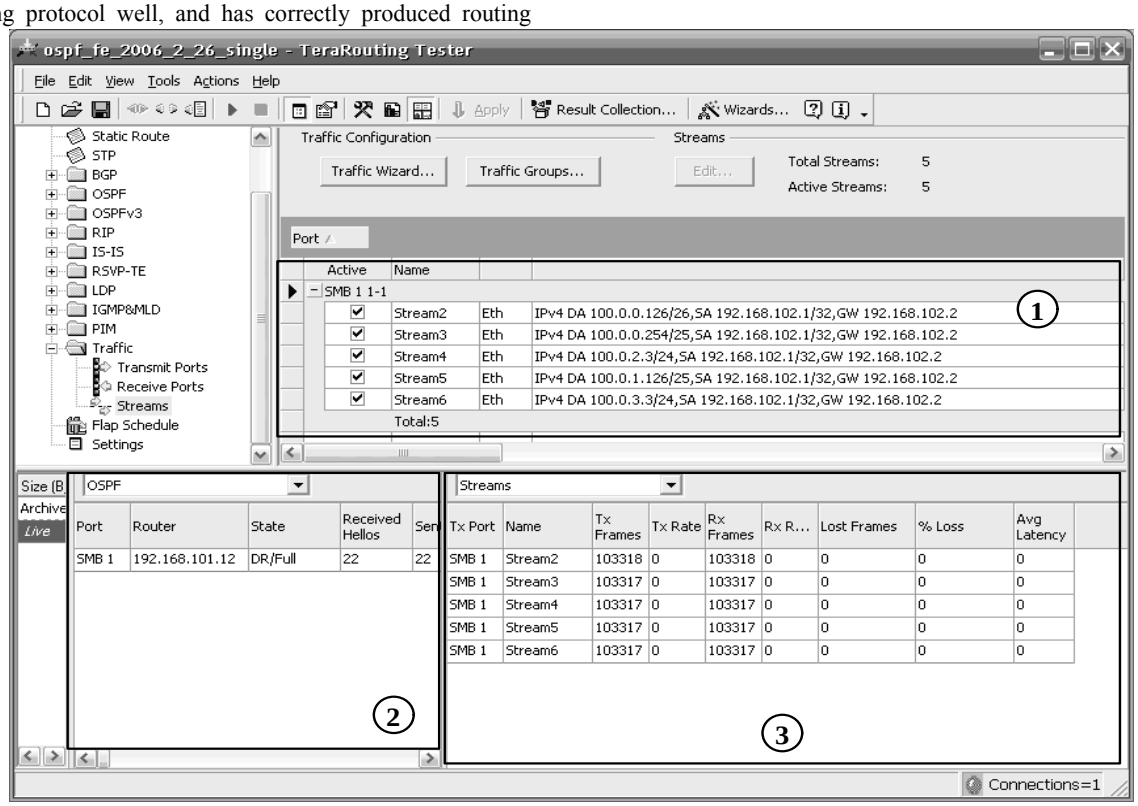


Fig. 8. Result of Routing and Routing Protocol Experiment

4.3. SNMP-based Network Management

As Fig. 3. explains, SNMP agent is run on the SNMP server. The server collects information on CE and FEs to form MIB information. Information on FEs is actually collected by means of ForCES protocol.

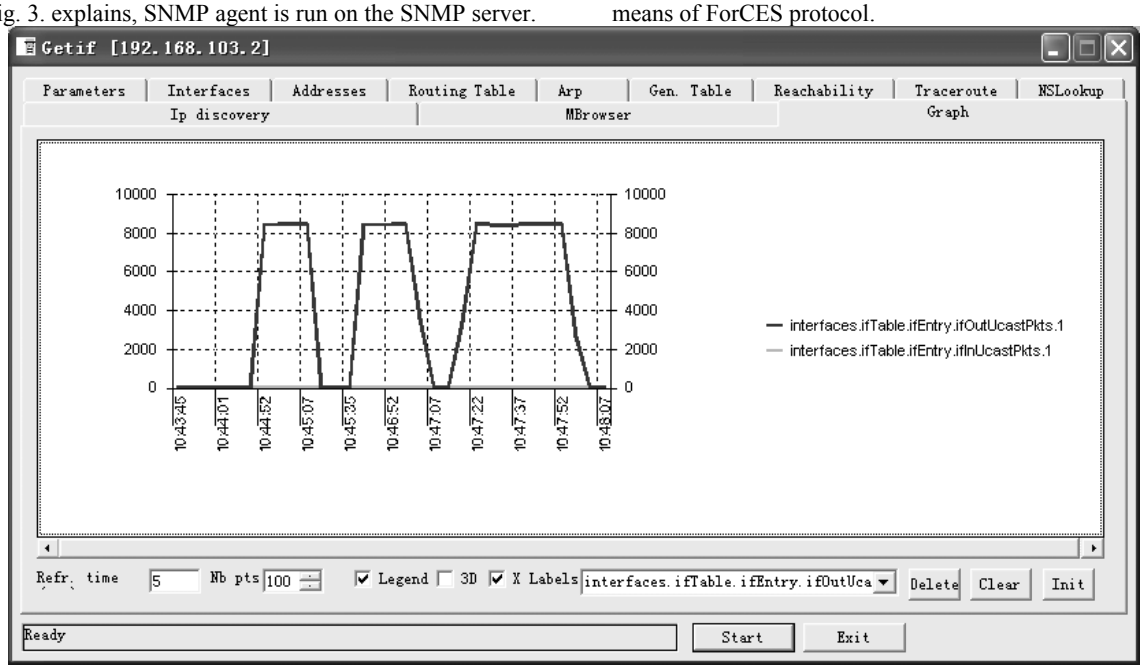


Fig. 9. Result of Testing SNMP

The experiment is to show the ability of ForTER to support SNMP. It has been tested that the SNMP agent in ForTER can configure and query the information on the FEs to correctly form various MIB information, and it can also form SNMP trap messages to report events happened on FEs. We use Getif software module as a SNMP management browser, which is located in the outer network of ForTER. Fig. 9 is the experiment result shown on the Getif that shows the IP traffic of a specific port in an FE of the ForTER. This is done by means of SNMP to GET the ifOutUcastPkts and ifinUcastPkts MIB on the FE. We still use SmartBits600 as the traffic generator of the port, and use SmartBits600 start/stop button to intently shift the traffic. Then, we monitor the traffic got by the Getif to see the coherence. Fig. 9. shows an snapshot of the traffic diagram, where ifOutUcastPkts and ifinUcastPkts are recorded.

The experiment shows that SNMP is well supported in the implementation of ForTER, hence, shows the feasibility of ForCES architecture and ForCES protocol from another point of view. The ForCES architecture and the ForCES protocol is also the essence for ForTER correctly running the SNMP protocol.

4.4. DoS Attack Prevention Algorithm

The algorithm to prevent the redirect data DoS attack, which is to bind UDP to TCP regarding its congestion control, as proposed in Fig. 7. , is tested its feasibility using a platform as shown in Fig.10. There is only one CE and one FE simulated. Connecting CE and FE is a 10Mbps Ethernet link. SmartBits600 are used to generate rate adjustable TCP and UDP flows that are input to FE. The TCP flow simulates ForCES control message flows that have to be sent from FE to CE via a TCP+UDP TML, and the UDP flow simulates ForCES redirect messages that also have to be delivered to CE from FE via the TML. By adjusting the SmartBits600 rate, we can then adjust the control and redirect messages rates respectively. At the CE side, it is designed to recover the TCP and UDP flows again when CE has received the flows, and then returns the flows back to SmartBits600. In this way, in the SmartBits600, we then can monitor how the TCP flow or the UDP flow has been transported from FE to CE, or in other words, how the throughput is over the CE-FE link for the two flows, to see if some of them are blocked or not.

We test the feasibility of the DoS attack algorithm by means of comparing the throughput status of the two flows when the algorithm is applied and when not applied.

Fig. 11. shows the throughput status when no DoS attack prevention means is applied. In this case, the UDP flow is independent of the TCP flow regarding the congestion control method when been sent from FE to CE. TCP sending traffic keeps 5Mbps, and UDP sending traffic changes from 1Mbps to 20Mbps. Note that heavy UDP traffic really simulates the DoS attack from redirect data. Fig. 11 shows the throughput changes of the two flows following the UDP traffic changes. The throughput is represented by a percentage, indicating the relative throughput. A little more than 1.0 of the relative throughput is from the system rounded error. From the figure, we can clearly see the TCP throughput decreases along with the UDP traffic increases. This means, in the ForCES case, the ForCES control messages become harder and harder to be delivered to CE when ForCES redirect message traffic overwhelms.

Fig. 12. shows the different result when the redirect DoS attack prevention algorithm applies. As described, the algorithm binds UDP traffic to TCP traffic. Fig. 12. shows that when the UDP traffic increases from 1M to 20M, the TCP relative throughput can roughly keep at 100% status, meaning the ForCES control messages are not blocked. This also means attackers are unable to destroy the communications between CE and FE.

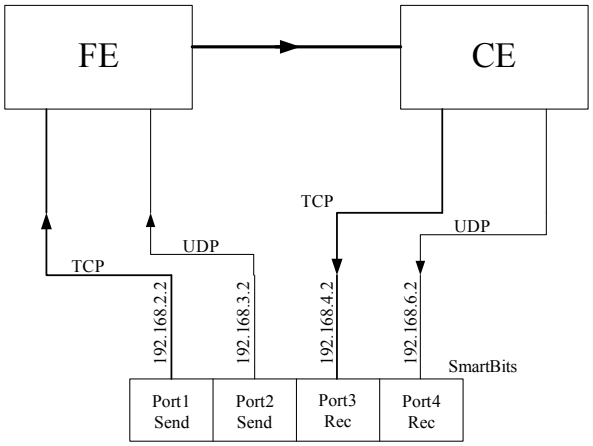


Fig. 10. Dos Attack Testing Platform

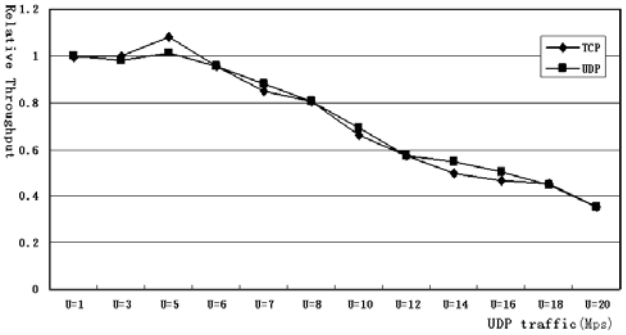


Fig. 11. UDP Independent of TCP

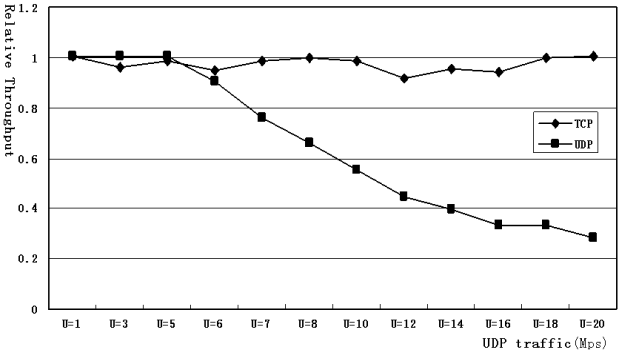


Fig. 12. UDP Binding to TCP

5. CONCLUSION

In this paper, we introduced the design and implementation of the ForTER, an IETF ForCES-based router. The architecture and the layered software model for ForTER are described. The detailed design schemes for the Control Element and the Forwarding Element are presented, and an algorithm to prevent DoS attacks from ForCES redirect data is proposed.

Experiments have shown the feasibility of ForTER with its designs regarding the ability for routing and running routing protocol, and the ability to support SNMP network

management. More importantly, the experiments have, as a result, actually illustrated the feasibility of IETF ForCES architecture as defined by RFC 3654 and RFC 3746, and the IETF ForCES protocol, which is near to be standardized as an RFC protocol.

IETF ForCES working group is now working hard and is towards its completion. Implementations based on ForCES are highly expected from vendors as well as the ForCES working group so as to make one important step forward towards completion of ForCES specifications, and for ForCES applications. As the only one full ForCES implementation ever presented till now, it is expected ForTER implementation introduced in this paper can contribute to IETF ForCES work for its progress.

6. REFERENCES

[1] B. Struck, M. Weingarten, “Next Gen Switch/Router Design Issues”, Business Communication Review, 2004, 44-50.

[2] K.L. Calvert, A.T. Campbell, et al, “Guest Editorial. Active and Programmable Networks", IEEE Journal on Selected Areas in Communications, 2001, 19(3).

[3] W.M. Wang, L.G. Dong, “Design and Performance Evaluation of General Router Management Protocol (GRMP) ", Proc. IEEE TENCON (Analog and Digital Techniques in Electrical Engineering), 2004.

[4] L.Yang, J. Halpern, et al, “ForCES Forwarding Element Model", work in process, 2006, http://www.ietf.org/internet-drafts/draft-ietf-forces-model-06.txt.

[5] SoftRouter, http://www.bell-labs.com/org/11347A/projects.htm#SoftRouter.

[6] M.El-Darieby, J.Rolia et al, “Performance Modeling for Virtual Network-based Service Provisioning", Proc. Integrated Network Management, 2001.

[7] Multiservice Switching Forum, http://www.msforum.org/.

[8] OPENSIG, http://comet.ctr.columbia.edu/opensig/.

[9] Network Processing Forum (NPF), http://www.npforum.org/.

[10] Forwarding and Control Element Separation (forces), http://www.ietf.org/html.charters/forces-charter.html

[11] U. Naik, R. Yavatkar, et al, “IXA Portability Framework: Preerving Software Investment in Network Processor Applications", Intel Technology Journal, 2002, 6(3).

[12] T. Woo, I. Faynberg, “The Softrouter Concept and the Proposal for the Study of Separation of Forwarding and Routing in the Next Generation Network", http://www.itu.int/ITU-T/ngn/fgngn/docs/FGNGN-004-June04.doc, 2004.

[13] General Switch Management Protocol, http://www.network sorcery.com/enp/protocol/gsmp.htm.

[14] W.M. Wang, L.G. Dong, et al, TCP and UDP based ForCES Protocol TML over IP Networks, Internet-Draft, 2006.

[15] A. Doria, W.M. Wang, et al, “ForCES Protocol Specification”, Internet-Draft (IETF ForCES WG Document), 2006. http://www.ietf.org/internet-drafts/draft-ietf-forces-protocol-08.txt

[16] J. Biswas, A. Lazar, et al., “The IEEE P1520 Standards Initiative for Programmable Network Interfaces", IEEE Communications Magazine, 1998, 36:64-72.

Customer Intelligence Based on Web Services

Ke Wang
College of Computer and Information Engineering, Zhejiang Gongshang University,
Hangzhou, Zhejiang, China
Email:rebecca1027@163.com
And
Guangming Wang
College of Computer and Information Engineering, Zhejiang Gongshang University,
Hangzhou, Zhejiang, China
Email: g.m.wang@hotmail.com

ABSTRACT

By means of processing and analyzing the data related to customers, CI is an effective method to help an enterprise make sound decisions to improve customer relationship. But for some medium and small sized enterprises, the difficulty and cost of data integration required by CI are very high, and the expense of establishing CI system is farily enormous. It is urgent to seek a feasible solution for this situation. Web service is an important technology for distributed computing and resource sharing, so it can support the implementation of distributed CI services. This article presents a Distributed Customer Intelligence System (DCIS), which was developed with Java Web Services Developer Pack (JWSDP), one web service developing technology. DCIS can effectively provide distributed services for CI application.

Keywords Customer Intelligence, Web services, data mining, SOAP, Java API.

1. INTRODUCTION

Nowadays, the market ecnomy tends to become strongly competitive, almost all enterprises are up against the pressure of competetion. Various kinds of enterprises transform their focus from product to customer.Thus Customer Relationship Management (CRM) turns into the urgent affairs of enterprises. Customer Intelligence (CI) focusing on CRM is a subsegment of business intelligence. It is the application of advanced analytical methods to identify, attract and retain the best and most profitable customers. Analytical results are used for anticipating, measuring and optimizing customer relationships.

Enterprises have piled up a mass of data related to customer during operation including manufacture, vendition and services. Databases are overflowing with details about customers and contacts. The challenge is how to turn this raw data into useful information. Whether CRM could meet the need of enterprises is determined by how to dig useful information from historical data and making decision according to this information.

The typical solution is to build a customer-database integrating all data related to customers, and analyze the data with data mining tools in order to discover valuable patterns. This solution is not effectual in some situation. For some medium and small sized enterprises (MSE), the difficulty and cost of data integration are very high, and the expense of establishing CI system is farily enormous.

The application of Internet has gone deep into sorts of fields, such as business, education, research and life. It has become international research hotspot how to offer intellectualized services and help users achieve high-quality services. This article presents a Distributed Customer Intelligence System (DCIS) based on Web services to resolve the problems discussed above.

2. TECHNOLOGY FACTORS

2.1 Data mining

Data mining discovers the meaningful patterns and relationships in data, and provides decision-making information. Data mining techniques include the following [1].

Association: looking for patterns that describe how one event is connected to another.

Sequence or path analysis: looking for patterns that describe how one event leads to a later event.

Classification: looking for patterns that describe how one class is distinct from others.

Clustering: finding distinct groups that are not previously known.

Forecasting: discovering data patterns leading to reasonable predictions.

The technology of data mining has been developed quite mature. Hence we can achieve the function of CI using data mining technique. Data mining provides a clear picture of what is going to happen, such as which customers might be most valuable, which customers are likely to losing, or, which kind of product is popular with customers.

2.2 Web services

Web services provide a standard means of interoperating among different software applications, running on a variety of platforms and/or frameworks [2]. A Web service is a software module designed to support interoperable machine-to-machine interaction over a network. It has an interface described in a machine-processable format, specifically Web service description language (WSDL). Other systems interact with the Web service using SOAP messages. The mechanics of the message exchange are documented in a Web service description (WSD). The WSD is a machine-processable specification of the Web service's interface, written in WSDL. It defines the message formats, data types, transport protocols, and transport serialization formats that should be used between a requester and a provider. It also specifies one or more network locations at which a provider can be invoked. In essence, the service description represents an agreement governing the mechanics of interacting with that service.

During past years, Web services techniques, such as SOAP, WSDL and UDDI, have progressed rapidly and

gained the favour of many large software vendors. Web services can be implemented with kinds of programme languages, thus it has a broad application foreground. Web services turns into an indispensable infrastructure in distributed computing [3], which can be adopted in DCIS.

3. THE FRAMEWORK of DCIS

The framework of DCIS is showed in Fig. 1. It regards distributed customer intelligence program modules as Web services providers, and each distributed client system as a Web services requester. Customer intelligence program modules release their services in Web services agency. Each client system has its own self-governed data source and application program. In order to get customer intelligence services, a client system only needs submit request to Web services agency using SOAP messages. Web services agency can then find and bind the right service for client system according to its request. This kind of framework is appropriate to provide customer intelligence services for some medium and small sized enterprises.

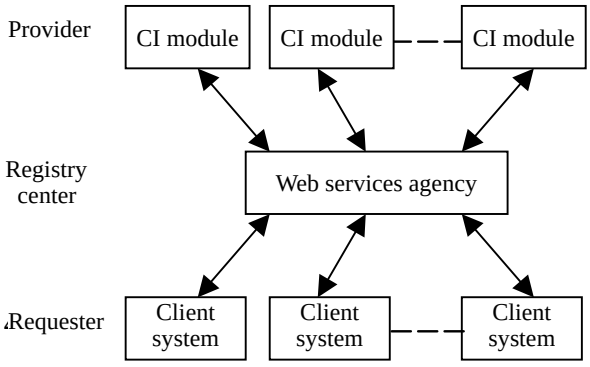


Fig. 1 Framework of Distributed Customer Intelligence System

4. FUNCTIONS of DCIS

4.1 Functions of customer intelligence program modules

1) Customer Subdividing

Customer subdividing is to classify customers according to their characters, which divides large members of customers into several different classes. Customers in a same class have similar characters, whereas customers in different classes have dissimilar characters. This function can be implemented with the technique of classification and clustering.

2) Cross-selling Analysis

Cross-selling analysis is to better sale new products and services to primary customers. It analyzes customers' behavior to understand what types of products are often purchased together, and how much is the probability of certain class customers buying certain product. This function can be implemented with the technique of association.

3) Consuming Trend Forecasting

Consuming trend forecasting is to forecast customers' demand in the future, which needs to consider the tendency of long-term change, circular change, seasonal change and random change. This function can be implemented with the technique of sequence analysis and forecasting.

4) Customer Losing Analysis

Customer losing analysis is to find the characters of lost customers and discover the factors influencing customer losing. It is necessary to analyze lost customers' data to gain the rules of customer losing and forecast the losing probability of current customers. This function can be implemented with the technique of factor analysis and correlation analysis.

4.2 Functions of Web services agency

DCIS is developed by means of Java Web Services Developer Pack (JWSDP) [4], which can implement the releasing, finding and binding of Web services, and the communicating between requestesr and providers.

1) Web services agency is the registry center of DCIS. Its elementary tasks include connecting, querying and managing the data of registry center. Both providers and requesters can communicate with registry center using the Java API. The basic steps to perform queries and to update the registry are described as the following [5].

Step1 Managing Registry Data

A provider can submit registry information to the registry, modify it, and remove it. It involves getting authorization from the registry, creating an organization, adding services and service bindings to an organization, publishing an organization, publishing a specification concept, and removing data from the registry.

Step2.Establishing a Connection

A requester must establish a connection to the registry, which involves getting access to the registry, obtaining a connection factory, creating a connection, setting connection properties, obtaining and using a service object.

Step3 Querying the Registry

The simplest way to use the registry is to query the information about the service providers that have submitted data to it. Requesters may search for services information according to their demand using the friendly interface of the registry.

2) The program of requester and provider can be implemented with Java API for XML-based Remote Procedure Call (JAX-RPC), which offers a sort of mechanism that makes it possible to invoke object across network using SOAP messages. JAX-RPC allows a Java-based client to call Web service methods in a distributed environment. JAX-RPC relies on WSDL for describing Web services. The basic steps to build and run Web services application are depicted as the following [6].

Step1 Provide the Service Endpoint Interface

To make CI service available to clients through JAX-RPC, developers need to provide a JAX-RPC service endpoint definition, which describes the remote methods that can be called by the client and the method signatures.

Step2 Register the Service

Each provider needs to publish information about its CI service in the registry. It includes business name, contact information, service name, and service binding.

Step3 Generate the Artifacts

In order to handle communication between a client and a service endpoint, JAX-RPC needs various classes, interfaces, and other files on both the client and server side of the communication. These files are collectively called artifacts. The wscompile tool provided by JWSDP can generate these artifacts.

Step4 Deploy the Service

The CI service providers need to deploy their services in a container, J2EE or Tomcat. To make its WAR files

deployable, each service provider can use the wsdeploy tool provided by JWSDP.

Step5 Create the Requester

A requester can invoke the remote method on a service endpoint through a local object called a stub that can be generated from the service endpoint definition or from a WSDL document with the wscompile tool, or dynamically invoking the method using the JAX-RPC Dynamic Invocation Interface (DII).

3) As the agency of the communication between providers and requesters, the registry center requires that every provider must upload relevant requester program of invoking service. When a client system request a kind of CI service, the registry asks it to input necessary arguments, then creates a stub object to invoke the remote method on the service or dynamically invokes it, and gives the result to the requester. In this process, the client system doesn't need to know how to invoke the remote service, which can be dealt with by the registry.

5. CONCLUSION

This article discusses customer intelligence based on Web services, mainly presents the framework and functions of DCIS using the technology of data mining, SOAP, XML and Java API. This solution resolves the problem of distributed customer intelligence application for MSE, and extends the application of Web services. With the improvement of Web services and customer intelligence, distributed customer intelligence will be applied broadly.

6. REFERENCES

[1] M.H. Dunham, Data Mining Tutorial［M］, Publishing House of Tsinghua University, 2003.

[2] W3C Working Group, Web Services Architecture ［EB/OL］, http://www.w3.org/TR/ws-arch/, 2004.

[3] D.A. Chappell, T. Jewell, Java Web services［M］, China Electric Power Press, 2003.

[4] J. Foster, M. Porter, Developing Web services with Java APIs for XML Using WSDP［M］, China Water and Electricity Publishing Company, 2003.

[5] Java Web Services Developer Pack Tutorial 2.0 FCS ［EB/OL］, http://java.sun.com/webservices/downloads /webservicespack.html, 2006.

[6] Web Services Developer Pack Part 2: RPC Calls, Messaging, and the JAX-RPC and JAXM API［EB/OL］, http://java.sun.com/developer/technicalArticles/WebSer vices/WSPack2/jaxrpc.html, 2006.

Analysis of Ant Colony Algorithms in Network Routing Problems

HongBin Xia
School of Information Technology, Southern Yangtze University, Wuxi 214036, China

ABSTRACT

Ant Colony (ACO) algorithm simulates the phenomena that a swarm of ants in the search for food and shows the remarkable capability of finding shortest paths between a found food source and the anthill. It is introduced as a population-based evolutionary optimization technique. ACO is known as a problem solver for Combinatorial Optimization (CO) problems. In this paper, we discuss the applicability of ACO to a well-known CO problem, Network Routing problem. We compare the ACO-based Routing algorithm with traditional methods and provide the performance comparison between these algorithms. Some new ideas and possible research direction in network routing problem are provided in the last section of the paper.

Keywords: Ant Colony, Swarm Intelligence, Routing, and Combinatorial Optimization

1. INTRODUCTION

Although an ant is a simple creature, collectively a colony of ants performs useful tasks such as finding the shortest path to a food source and sharing this information with other ants by depositing pheromone. In recent years, computer scientists were able to transform the models of collective intelligence of ants into useful optimization and control algorithms [1]. In the interesting and emerging field of ant colony optimization (ACO) [1], a colony of (biological) ants is typically modeled as a society of mobile agents (or artificial ants). ACO has been applied in many combinatorial optimization problems such as the asymmetric traveling salesman problem [2], and vehicle routing problem [1]. This paper focuses on surveying ACO approaches in network routing.

2. ANT COLONY OPTIMIZATION (ACO)

In the following we will present a definition of the problem representation on which ACO algorithms work. We will describe how solutions are constructed in the representation and present the ACO algorithm as pseudo-code.

The central properties of ACO are based upon the self-organized collective foraging behavior of ants. The key property collective in the foraging behavior of ants is their ability to find shortest paths between the location of their anthill and the location of food sources. When the ants move on a path between their anthill and the location of a food source they lay a pheromone trail. Other ants can then follow these generated paths of pheromone trails. This means that ants tend to converge on the same path.

In ACO, solutions are constructed repetitively by adding solution components to partial solutions stochastically. Solutions are constructed by taking into account (i) heuristic information when adding solution components (if available), and (ii) (artificial) pheromone trails which change dynamically based on the experience of the ants. Stigmergy handles the propagation of experience between ants.

The algorithm constructs solutions by constructing paths in a graph. This process can be stated in pseudo-code, which is the intention of the following section.

```
Procedure run_ACO() {
        While (time++ < tmax) {
              ants_generation_and_activity();
              pheromone_trail_update();
             daemon_actions();
             }
  }
Procedure ants_generation_and_activity() {
       While ( no_of_ants++ < m ) {
             new_active_ant();
           }
}
Procedure new_active_ant() {
     ant = create_new_ant();
     Position = get_starting_vertex();
     initialize_ant( position );
    while ( not( in_target_state( ant ) or
        could_not_reach_target_state( ant ) ) ){
     A = read_
possible_routes_from_current_vertex( ant );
     P = compute_transition_probabilities( A,
ant, . );
     Next = stochastically_choose_path( A, P );
     ant_move_to_next( next );
        If
(online_step_by_step_pheromone_update ) {
           deposit_trail_on_the_last_[edge/vertex]
( ant );
         }
     }
   If ( online_delayed_phero_update ){
     deposit_trail_on_all_[edges/vertices]( ant );
      }
   If ( could_not_reach_target_state( ant ) ) {
       ant_deaths++;
     }
  Else {
     // Update the current shortest tour
     If ( length_of_tour < shortest_tour ) {
         shortest_tour = length_of_tour;
      }
    }
  }
```

3. ACO VERSUS TRADITIONAL ROUTING

In OSPF, routing is achieved by having each node transmit a link-state-packet (LSP) to every other node in a network through a flooding processing. Although an LSP (which carries information about the costs to all the neighbors of a node) is generally smaller than a routing table, the flooding process ensures that every node in receives a copy of the LSP. Since an LSP from a node can be disseminated via different paths to other nodes, multiple identical copies of the same LSP may be transmitted to the same node.

Routing in RIP involves the transmission of routing tables of each node to every one of its neighbors. For a large network, the routing table of (which consists of a list of cost vectors to all other nodes in) is large. Since each needs to transmit its routing table to every of its neighbors, the routing overhead can be very large.

Routing in ACO is achieved by transmitting ants rather than routing tables or by flooding LSPs. Even though it is noted that the size of an ant may vary in different systems/implementations, depending on their functions and applications, in general, the size of ants is relatively small, in the order of 6 bytes [3]. This is because ants are generally very simple agents.

4. ROUTING USING ACO ALGORITHM

Caro and Dorigo's AntNet [4] was originally designed for routing in packet-switched networks. Unlike traditional routing algorithms (such as OSPF and RIP), which focused on minimal or shortest path routing, routing in AntNet, was carried out with the aim of optimizing the performance of the entire network.

Tony White [5] designed ant system with genetic algorithm (ASGA) to solve problems of point-to-point, point to multipoint and cycle (multipath) routing in circuit-switched networks. Routing in ASGA is achieved by using explorer ants to update pheromone tables. Although similar to AntNet, explorer's travel in a round trip, but unlike backward ants in AntNet, explorers deposit the same amount of pheromones in their return trips [5]. In addition, evaporation agents and pheromone heuristic control were used to mitigate stagnation. The addition of GA was to increase the adaptivity of ants. For instance, if the best path is congested, it increases the likelihood of ants to find an alternative path. However, unlike the ABC system, ASGA was not designed to solve the load-balancing problem in circuit-switched networks.

5. ACO IN ROUTING

Dynamic routing algorithms that try to alleviate congestionor guarantee quality of service based on network traffic measurements are not new. The "new" ARPANET routing protocol computed link weights from link delay measurements, and recomputed routing tables when link delays changed significantly. However, this protocol caused wide oscillations in link utilization under heavy load. Khanna and Zinky [6] revised the delay metric to improve the stability of this protocol. Since then, many other adaptive algorithms have been proposed in the literature.

AntNet is a routing algorithm that employs two-way mobile agents to update link probability tables, which is similar to TB. The two-way agents are sent from the source to the destination, and they update the link probability tables on their return trip. The main difference between TB and AntNet is that TB does not require a lengthy learning period to discover the network topology, since TB assumes that the underlying link-state protocol provides this information. Moreover, the AntNet paper did not investigate the interaction between ant-based routing and congestion control protocols, such as TCP. TB is also simpler than the AntNet algorithm.

Another ant-like algorithm was developed by Subramanian et al. in [7]. This algorithm sends one-way messages from the destination to the source. Like AntNet, this paper did not investigate the interaction with congestion control protocols. This algorithm also requires a lengthy learning period to discover the network topology.

In addition to ant-based algorithms, several other types of adaptive algorithms have been developed. Basu et al. [8] propose a class of routing algorithms based on potentials. They also develop a specific traffic-aware routing protocol using potentials. Traffic awareness is achieved by defining the potential of a link for a given source-destination pair as a function of the queue length of the link and the distance to the destination. Another approach to adaptive routing is presented byWang and Crowcroft [9]. In this paper, a protocol that uses emergency exits to alleviate congestion on shortest paths is proposed. These emergency exits are precomputed and are only used when queue lengths are detected to be above a certain threshold. In a somewhat different approach, Shaikh et al. [10] present a protocol that distinguishes between long-lived and short-lived IP flows. Adaptive routing is performed only on long-lived flows, which are detected at the routers by keeping track of the number of bytes for the flow. Once a flow is determined to be long-lived, another mechanism such as MPLS is is used to create a dynamic route. Chen et al. [11] also present a dynamic routing protocol that only applies adaptive techniques to some flows in the network. Their approach is to focus on destination networks that receive a lot of traffic. These "hot" destinations initiate paths to themselves using dynamic link metrics. Packets for all other destinations use the normal link state protocol. The protocol of [11] is called Scout, and it uses scout packets sent by the popular destinations to periodically update the routing tables in other nodes. The Scout protocol and the behavior of their scout packets are very different from ours. This protocol is based on the periodic flooding of scout packets to discover link metrics and to perform the path updates. The seminal work here is the minimum delay routing algorithm by Gallager [12]. This work developed an iterative algorithmthat converges to the optimal solution for stationary input traffic statistics. Vutukury and Garcia-Luna-Aceves present an approximation to minimum delay routing is presented. If the traffic matrix that gives the point-to-point volume is available, Fortz and Thorup [13] propose an algorithm that precomputes OSPF weights to alleviate congestion and to achieve good utilization. TB differs from these algorithms in that it

does not depend on any prior knowledge of traffic patterns.

6. DISCUSS AND NEW DIRECTIONS

Although the ant metaphor has been successfully applied to routing of data packets both in wireless and fixed networks, little is known yet about its appropriateness for search in peer-to-peer environments. [14] presents SemAnt, a distributed content-based routing algorithm based on the Ant Colony Optimization meta-heuristic and adapted for deployment in peer-to-peer networks. Under the assumption that content is annotated according to taxonomy, it is possible to determine the hierarchical relationships between queries, and to exploit this information to improve the routing process. their results show that using taxonomies enhances search performance in peer-to-peer networks. The degree of enhancement is highly dependent on the content distribution in the network.

Generally speaking, ant algorithms can be said to use heuristics to delimit the hypothesis space, and through the Stochastic State Transition mechanism explore the hypothesis space. The ant algorithms can suggest hypotheses that can be tested with the Q-function. Ant algorithms learn to produce (suggest) viable hypotheses. Inductive machine learning algorithms (IMLA) mainly operate through receiving already suggested hypotheses s and the resulting quality Q(s). The IMLA consequentially results in approximating the quality function Q. This is a good idea when Q is not known, and the overall task is to be able to classify hypotheses according to Q. IMLA learn to classify hypotheses. Inductive machine learning algorithms combined with ant algorithms.

REFERENCES

[1] J M. Dorigo, G. D. Caro, et al, Gambardella, “Ant algorithms for discrete optimization,” Artif. Life, 1999, 5(2) .137–172.

[2] M. Dorigo, L. M. Gambardella, “Ant colony system: A cooperative learning approach to the travelling salesman problem,” IEEE Trans. Evol. Comput, 1997, 1. 53–66.

[3] D. Subramanian, P. Druschel, et al, “Ants and reinforcement learning: A case study in routing in dynamic networks,” in Proc. Int. Joint Conf. Artificial Intelligence, Palo Alto, CA, 1997, IJCAI-97, pp.832–838.

[4] G.D. Caro, M. Dorigo, “AntNet: Distributed stigmergetic control for communications networks,” *J. Artif. Intell. Res.*, 1998, 9. 317–365.

[5] T. White, F. Oppacher, et al, “Connection management using adaptive agents,” in Proc. Int. Conf. Parallel and Distributed Processing Techniques Applications, July 12–16, 1998, pp. 802–809.

[6] A. Khanna, J. Zinky, The revised ARPANET routing metric, In Proceedings of ACM SIGCOMM ’89, Austin, TX, September 1989, pp.45-46.

[7] D. Subramanian, P. Druschel, et al, Ants and reinforcement learning: A case study in routing in dynamic networks, In Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI-97), August 23-29 1997, pp. 832–839.

[8] A. Basu, A. Lin, et al, Routing using potentials: A dynamic traffic-aware routing algorithm. In Proceedings of ACM SIGCOMM ’03, Karlsruhe, Germany, August 2003.

[9] Z. Wang, J. Crowcroft, Shortest path first with emergency exits. In Proceedings of ACM SIGCOMM ’90, Philadelphia, PA, August 1990, pp.166-176.

[10] A. Shaikh, J. Rexford, et al, Load-sensitive routing of long-lived ip flows. In Proceedings of ACM SIGCOMM ’99, Cambridge, MA, August 1999, pp.215-226.

[11] J. Chen, P. Druschel, et al, A new approach to routing with dynamic metrics. In Proceedings of IEEE INFOCOM ’99, New York, NY, March 1999, pp.661-670.

[12] R.G. Gallager, A minimum delay routing algorithm using distributed computation. IEEE Transactions on Communications, January 1977, 25(1): 73-85.

[13] B. Fortz, M. Thorup, Internet traffic engieneering by optimizing OSPF weights. In Proceedings of IEEE INFOCOM ’00, pages 519–528, Tel Aviv, Israel, March 2000, pp.519-528.

[14] Gabber, S.T. Blazer, A Routing Algorithm Inspired by Ants. Proceedings of the 12th IEEE International Conference on Network Protocols (ICNP’04) Berlin, Germany, October 2004.

FS-SVM Based Intrusion Detection Feature Selection and classification

Xueqin Zhang, Chunhua Gu
College of information science and Engineering, East China University of science and technology
Shanghai, 200237, China
Email: zxq@ecust.edu.cn

ABSTRACT

Feature selection is an important issue in machine learning. Support Vector Machine (SVM) is a popular topic in Intrusion Detection System (IDS) for its abilities to perform classification and regression in recent years. Solving a support vector machine problem is a typical quadratic optimization problem, which is influenced by its time complexity. A FS-SVM algorithm based on Fisher score and SVM was presented in this paper. This algorithm was applied on KDDCUP'99 standard Intrusion Detection dataset. The experiment results show, using FS-SVM algorithm to select the important features from the original feature space is an effective method to reduce the dimension of the example feature vector. The classification accuracy has not decrease comparing to the original feature space in general.

Keywords: FS-SVM, Feature Selection, Intrusion detection.

1. INTRODUCTION

Support Vector Machine (SVM) is a new machine learning method building on statistic theory and Structural Risk Minimization Principle for classification and regression. In the last few years, it has seen a rise of support vector machines as powerful tools for solving classification and regression problems [1, 2, 3].

Solving a Support Vector Machine is typically a quadratic optimization problem. The time complexity is related with the dimension of feature vector and the number of examples. The training time and memory grow dramatically with the increase of the size of training dataset. But, for the situation of distributed high-speed network detection, in order to obtain satisfied detection accuracy, people usually select enough features to describe the characteristic of intrusion activity. It produces a heavy pressure to SVM classifier.

Feature selection is important issue in machine learning. In recent years, Fisher score is increasingly used as a feature selection strategy for classification problems. Especially when it was applied with high performance classifiers such as SVM, it often shows superb results.

This article is organized as following. In section 2, we introduce support vector classification and Fisher score technologies. In section 3, we introduce the algorithm of FS-SVM .In section 4, we describe the detail of experiment on KDDCUP'99 standard Intrusion Detection dataset, and give the experiment results. Finally, we make the conclusions in section 5.

2. RELATED BACKGROUND

2.1 Fisher Score

Given training example: (x_1, y_1) ,…, (x_l, y_l) , $x \in R^d$, $y_l \in \{+1,-1\}$, l is the number of training example, d is the dimension of the example feature, y_l is the class label. N is the number of training example. N_1 is the number of positive example and N_2 is the number of negative example.

Fisher score F is defined as the followed [4, 5]:

$$F = \frac{(\bar{m}_1 - \bar{m})^2 + (\bar{m}_2 - \bar{m})^2}{\frac{1}{N_1}\sum_{x\in+1}(X-\bar{m}_1)^2 + \frac{1}{N_2}\sum_{x\in-1}(X-\bar{m}_2)^2} \quad (1)$$

Define S_b as:

$$S_b = (\bar{m}_1 - \bar{m})^2 + (\bar{m}_2 - \bar{m})^2 \quad (2)$$

S_b is called between-class scatter matrix, it describes the distance of two classes. $\bar{m}_1$, $\bar{m}_2$, $\bar{m}$ is the mean of the positive class, negative class and all example respectively:

$$\bar{m}_1 = \frac{1}{N_1}\sum_{x\in+1} x,\ \bar{m}_2 = \frac{1}{N_2}\sum_{x\in-1} x\ ,$$

$$\bar{m} = \frac{1}{N}\sum x$$

Define S_W as:

$$S_W = \frac{1}{N_1}\sum_{x\in+1_i}(X-\bar{m}_1)^2 + \frac{1}{N_2}\sum_{x\in-1_i}(X-\bar{m}_2)^2 \quad (3)$$

S_W is called within-class scatter matrix, it describes the distance between one class. The Fisher score is the weight of each feature. These values show the contribution of each feature to classification. The feature with high Fisher score can be selected to construct a reduced feature subset. Fisher score is an efficient method to seek features for classification.

The feature selection problem can be evaluated by: fixed m, find the m features that obtain the smallest expected generalization errorγ; or fixed maximum allowable generalization errorγ, find the smallest m. In both of these problems γ must be estimated. In our algorithm, we used SVM to evaluate the performance of the feature selection strategy.

2.2 Support Vector Machine

Given training example described as section 2.1. The support vector technique solves the following optimization problem for linear problem [6,7]:

$$\min \frac{\|\omega\|^2}{2} + C\sum_{i=1}^{l}\xi_i \quad (4)$$

s.t. $y_i(\omega \cdot \phi(x) + b) - 1 \geq 0$

Where, C is a penalty coefficient, ξ is a relax variable, $\xi_i \geq 0$. The solution to Eq. (4) equal to the solution to the constrained quadratic optimization problem using the Wolfe dual theory as:

$$\max_{\alpha} \; W(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{i=1}^{l} \alpha_i \alpha_j y_i y_j k(x_i \cdot x_j) \quad (5)$$

s.t. $\sum_{i=1}^{l} y_i \alpha_i = 0$, $0 \leq \alpha_i \leq C, i = 1, 2, ..., l$

Here α_i are the Lagrange multipliers. For linear non-separable problem, the training examples in input space X can be mapped into a high dimension space *H* by kernel function *K*. The common kernels are linear kernel, polynomial kernel, and radial basis function (RBF) kernels.

According to the Karush-Kuhn-Tucker (KKT) necessary and sufficient optimality conditions, there are:

$$\alpha_i [y_i(\omega \cdot \phi(x) + b) - 1] = 0 \quad (6)$$

Here, $y_i(\omega \cdot \phi(x) + b) - 1 = 0$ if and only if $\alpha_i > 0$. These examples corresponding to $\alpha_i > 0$ decide the interval boundary. They are called support vectors (SVs).

3. FS-SVM ALGORITHM

The FS-SVM algorithm focuses on combining SVM with Fisher Score to do feature selection and classification. It is described as following:

1. For given dataset, calculate Fisher score of each feature.
2. Choose some thresholds, for each threshold, do the following：

a) Drop features with Fisher score below this threshold.
b) Split the training data into two parts randomly: D_train and D_test. Let D_train be the training dataset and D_test dataset be the testing dataset.
c) Apply the SVM classification method to train D_train and test on D_test to get the validation error.
d) Repeat the steps above some times, and then evaluate the average validation error.

3. Choose the threshold with the lowest average validation error.
4. Abandon features with Fisher score below the chosen threshold. Then apply the SVM algorithm in Section 2 to do classification.

4. EXPERIMENTS AND RESULTS

4.1 Data Description

KDDCUP'99 intrusion detection standard dataset, which is issued by MIT lab, was used in our experiment [8]. In this dataset, there are about 5 million network connection records that including normal network connection records and abnormal attack records, such as land attack, neptune attack, password guess, port scan, etc. Every connection is a sequence of TCP packets starting and ending at some well defined times, between which data flows to and from a source IP address to a target IP address under some well defined protocol. Each connection is labeled as either normal, or as an attack, with exactly one specific attack type. These attacks fall into four main categories:

DOS: denial-of-service, e.g. syn flood;
R2L: unauthorized access from a remote machine, e.g. guessing password;
U2R: unauthorized access to local super user (root) privileges, e.g., various ``buffer overflow'' attacks;
PROB: surveillance and other probing, e.g., port scanning.

All of the data have 41 features, including 34 numerical features and 7 character features. The name, serial number and description and other detail information of the 41 features can be found on the website: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

In our experiment, near 50 thousands was selected randomly from ‘10percent’ dataset of KDDCUP'99. The percent of normal record and attack records in this dataset are according to the original percent in ‘10percent’ dataset. Those character features of the examples were encoded with decimal number by picking up all probability dispersant value of the feature in the dataset. After this disposing, all the data were scaled into a [0,1] range by $x_i = (x_i - x_{\min}) / (x_{\max} - x_{\min})$.

4.2 Experiment Result

To evaluate our FS-SVM algorithm, we computed three statistics variables in our experiments: the Accuracy (Acu), the detection rate (DR) and the false positive rate (FR). The Accuracy is the percentage of records that are classified correctly. The detection rate is the percentage of records generated by the malicious programs, which are labeled correctly as anomalous by the classifier. The false positive rate is the percentage of normal records, which are mislabeled anomalous.

For the four categories of attacks, in order to seek the important features for each of them, we abstracted four dataset from 1percent dataset, named Normal_DOS, Normal_PROB, Normal_U2R, and Normal_R2L respectively. By applying the FS-SVM algorithm mentioned in the section 3 to the four dataset, we got the following selection result for each attack:

DOS: 12, 23, 32, 2, 24, 37, 36, 31, 6

PROB: 35, 27, 4, 28, 40, 41, 34, 12, 29

R2L: 10, 22, 33, 1, 34, 36, 38, 37,

U2R: 13, 17, 18, 14, 16, 11, 37, 3

Here, the number is the serial number of each feature given by KDDCUP’99. We selected 9 features for DOS attack, 8 features for PROB attack, 9 features for R2L attack and 8 features for U2R attack. These selected features are the important features with high Fisher score. The prediction accuracy is 99.58%, 99.66%, 99.75% and 99.98% respectively when doing cross validate testing. The rest features that we do not list here are those features with low Fisher score.

In order to compare the classification performance of standard SVM and FS-SVM, the ROC curves of DOS, PROB, R2L and U2R were given in Fig.1. [9]. In this figure, the X-axes represents the false positive rate, the Y-axes represents detection accuracy. Fig.1. (a) is the curve of DOS dataset, Fig.1. (b) is the curve of PROB dataset, Fig.1. (C) is the curve of R2L dataset and Fig. 1. (d) is the curve of U2R dataset. The experiment results show the classification accuracy of FS-SVM is equivalent to that of standard SVM

in general.

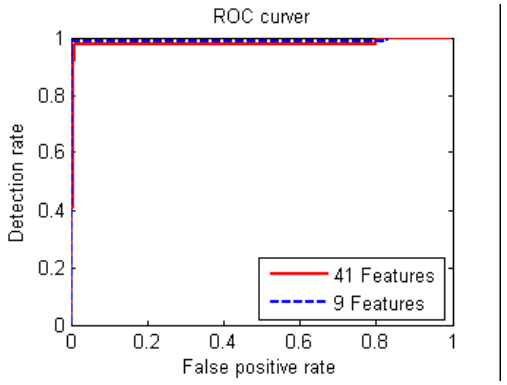


Fig. 1. (a)

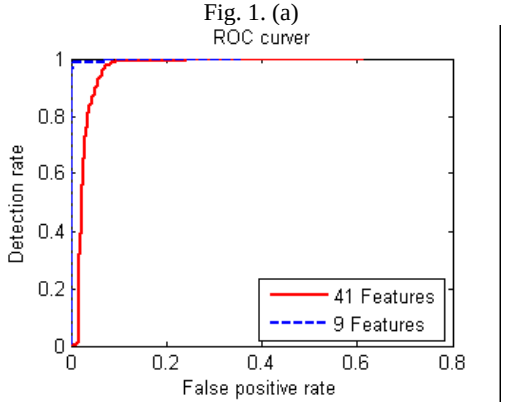


Fig. 1. (b)

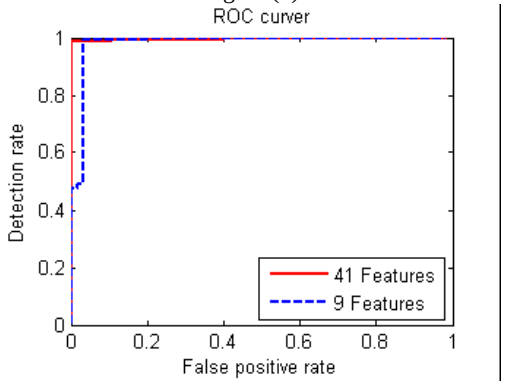


Fig. 1. (c)

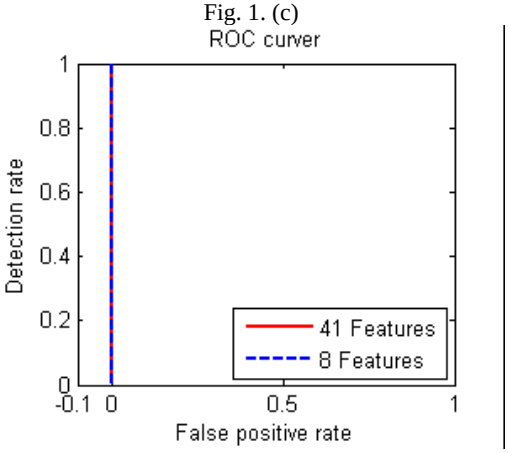


Fig. 1. (d)

Fig.1 Detection ability Comparison of FS-SVM and SVM

We also compared the classification executing time of FS-SVM and SVM algorithm. In Table 1, we list the training and testing time on the four databases when applying the two different algorithms. By do comparing, it is exactly that the total running time of FS-SVM decreases dramatically.

Table 1. Running time comparisons of FS-SVM and SVM

	SVM		**FS-SVM**	
	Train (s)	**Test (s)**	**Train (s)**	**Test (s)**
DOS	0.978	2.757	0.836	0.769
PROB	0.2293	0.4456	0.0425	0.1545
R2L	0.1786	0.4367	0.1295	0.1894
U2R	0.1383	0.3354	0.04166	0.1367

5. CONCLUSIONS

Fisher score is a simple and effective feature selection method. In this paper, FS-SVM algorithm was presented based on Fisher score and support vector machine. This method was applied on KDDCUP'99 intrusion detection standard dataset. By checking the accuracy, detection rate, false positive rate and training/testing time, we can see that using FS-SVM algorithm on classification can reduce the dimension of original dataset effectively without decreasing of classification precise. This FS-SVM method will be a useful way to reduce the data quantity of large-scale dataset. Our future work will focus on doing some comparison with other feature selection algorithms and using some technologies to reduce the number of training examples.

6. REFERENCES

[1] Z.R. Yang, K.C. Chou, "Bio-support vector machines for computational proteomics", Bioinformnatics, 2004, 20(3): 735-741.

[2] G.. Guo, S.Z. Li, "Content-based audio classification and retrieval by support vector machines", IEEE Transactions on Speech and Audio Processing, 2003, 14(1):209-215.

[3] S. Mukkamala, G. Janoski, et al, "Intrusion Detection Using Neural Networks and Support Vector Machines", In Proceedings of IEEE International Joint Conference on Neural Networks, 2002, 1702.

[4] K. Tsuda, M. Kawanabe, et al, "Clustering with the Fisher Score", In Proceedings of Advances in Neural Information, 2003, 729-736.

[5] S. Theodoridis, K. Koutroumbas, "Pattern Recognition (Second Edition)", Elsevier Science, 1999.

[6] C. Cortes, V. Vapnik, "Support Vector Networks", Machine learning, 1995, 20(3):273-297.

[7] C.J.C. Burges, "A tutorial on support vector machines for pattern recognition", Data Mining and Knowledge Discovery, 1998, 2(2):121-167.

[8] KDDCUP '99 dataset, available at: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[9] Y.Y. Gao, X. Fu, et al, "IDS evaluation approach based on ROC curves", Journa1 on Communications, 2005, 26(1):113–115.

A TCP-Friendly Multicast Congestion Control Scheme Based on Active Network *

Yunqing Fu
College of Distance Education, Chongqing University
Chongqing 400044/023, P.R. China
Email: yqfu@cqu.edu.cn
And
Zhu Liang
Department of Computer Science, Chongqing Industrial Polytechnic University
Chongqing 400030/023,P.R. China
Email: liangzhulz@163.com

ABSTRACT

Using multicast rather than sending data to each receiver individually(referred to as unicast) has the potential of saving on sender and network resources, and reducing the time to completion by overlapping transmission to multiple receivers. The differences between multicast and unicast, however, bring great challenges to congestion control for large-scale multicast applications. Effective multicast congestion control mechanism is urgently needed with the deployment of multimedia multicast applications in Internet. In this paper, a novel multicast congestion control scheme is proposed through introducing active network technology, hierarchical congestion detection, report and adaptive multi-rate control based on TCP flux model. Simulations and results show that the proposed scheme responds more quickly to congestion anywhere in the multicast tree and is more stable and TCP-friendly, and maximizes the utilization of network resource due to the fact that it effectively avoids and releases congestion within limited scope of network.

Keywords: Active Network, Multi-rate, Congestion Control, Multicast, TCP-Friendly.

1. INTRODUCTION

Multicast is an effective and efficient transport mechanism for simultaneously transferring bulk data to multiple receivers. The differences between unicast and multicast, however, bring great challenges to flow control and congestion control in multicast. Firstly, different receivers within a same multicast group always have diverse characteristics (e.g., available link bandwidth, data processing capability and Qos requirement), which thus makes congestion control for multicast more complicated. Secondly, since TCP strongly relies on other network flows to use congestion control schemes similar to its own, TCP-incompatible multicast traffic can completely lock out competing TCP flows and monopolize the available bandwidth. Hence, it is highly unlikely that congestion control for large-scale multicast will become widely accepted without TCP-friendly mechanism. Furthermore, multicast flow insensitive to existing congestion is likely to cause simultaneous congestion collapses in many parts of network, even the whole network.

Due to the fact that traditional network only forwards packet passively, congestion control research generally focuses on end-to-end control, even in the router-based congestion control scheme, its main control functions are still completed by end systems. But in fact, congestion always occurs inside network. Thus, two problems will be resulted from, i.e., congestion cannot be detected in time, and the corresponding control actions thus fail to keep up the congestion, thus causing vibration in network.

In this paper, we introduce a TCP-friendly and multi-rate congestion control scheme, which incorporates several novel features, including:

1) Active network-based congestion detection and control that makes it possible to place policies inside every part of network [1], instead of only on end systems.
2) Hierarchical congestion status detection and report distribute processing load across multicast group in a balanced way.
3) Hop-by-hop congestion detection and control and multi-rate multicasting limit the congestion control within parts of network, and maximize the utilization of network.

2. MULTI-RATE MULTICASTING PRINCIPLE

Fig. 1. shows a multicast tree of a certain multicast group. It is constructed corresponding with the shortest path tree. The root S of the tree denotes the multicast traffic source. The forking nodes R_1, R_2 and R_3 are junction nodes. Leaves R_{m1}, R_{m2}, R_{m3} and R_{m4} are multicast receivers. Following the definition of graph theory, R_1 is the parent of R_2 and R_3, R_2 is the parent of R_{m1} and R_{m2}, and S is the parent of R_1. Conversely, R_1 is a child of S, R_2 and R_3 are children of R_1, and so on. All nodes in the multicast tree, including the sender and receivers are active nodes [2,3].

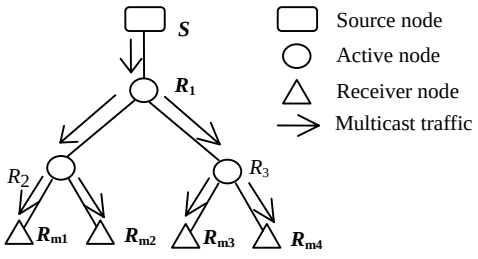


Fig. 1. Simplified multicasting network

The architecture of active node is illustrated in Fig. 2, where

* The research is supported by CSCT (2005BB2212), China.

AA denotes active applications, and *EE* is the environment of execution for active application.

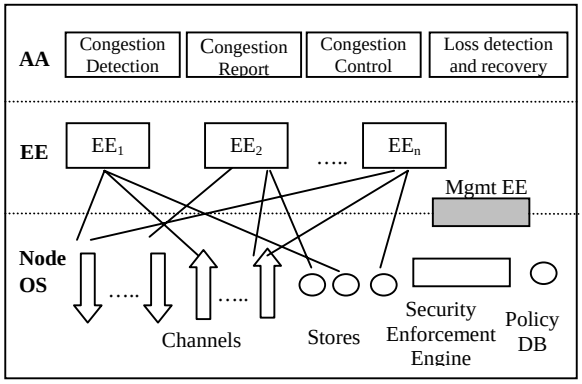


Fig. 2. Architecture of active node

The sender transmits data packets to its children. Each node in the multicast tree copies the packet received from the sender into its active buffer, then forwards the buffered packet to each of its children, and so on [5]. Active node (or the sender) sets a timer, called retransmission timer, for each transmission of a packet. Each receiver may send a positive acknowledgement (referred to as *ACK*) or negative acknowledgement (referred to as *NACK*) to its parent in the tree. Received packets are reported in *ACK*s and missing packets are requested in *NACK*s. An active node (or the sender) discards a buffered packet after receiving an *ACK* from all of its children. On the other hand, an active node (or the sender) retransmits a packet (1) upon receiving a *NACK* reporting that the packet is missing, or (2) if it does not receive an *ACK* for the packet from all its children in the multicast tree before the timer associated with the packet expires.

Due to the heterogeneity of multicast receivers, the rate control in our scheme is multi-rate [6]. Each active node decides its sending rate based on its utility function, congestion status of its children and the congestion signals received from its children.

3. CONGESTION CONTROL SCHEME

Each active node independently monitors the congestion status of each child, and conducts corresponding control when there may exists a potential congestion at its children. Once an active node finds that it is going to be involved in congestion, it sends a congestion indication to its parent, the parent then adjusts its sending rate to avoid making the congestion worse, and reports upward if potential congestion maybe occurs. As a result, congestion is effectively avoided and released by this hop-by-hop congestion detection, reports and control.

3.1 Congestion Indication

Not only link state but also working state of network node may contribute to congestion. *RTT* (Round Trip Time) can reflect the state of communication links. For an active node, if its recent RTT between itself and a child exceeds the average *RTT* value for a long period, it is highly possible that there exists potential congestion on the link leading to that child. At an active node, if the rate of inflow keeps greater than the one of outflow, its active buffer may be exhausted, resulting in congestion due to the retransmissions of lost packets. Therefore, *RTT* as well as active buffer queue are used for congestion indication in our scheme.

3.2 *RTT* Measurement

For reliable delivery, each active node should set a timer based on the *RTT* between itself and each child for transmission and retransmissions of a packet. At the same time, *RTT* is an important parameter that affects rate adaptation in the phase of congestion voidance and control.

In order to estimate the *RTT*, an active node periodically polls each child by sending a probe packet in which it includes the timestamp of the probe being sent out (referred to as T_{sp}). Upon receiving this probe, a child immediately feedbacks an acknowledgement, it includes in the acknowledgement the time that the probe is received (referred to as T_{rp}) and the time that the acknowledgement is sent out (referred to as T_{sa}). Using the following expression, the parent can then measure its current round-trip time between itself and that child when it receives the corresponding acknowledgement (the time of the acknowledgement being received is denoted as T_{ra})

$$\begin{aligned} RTT_m &= T_{rp} - T_{sp} + T_{ra} - T_{sa} \\ &= (T_{ra} - T_{sp}) - (T_{sa} - T_{rp}) \end{aligned} \quad (1)$$

Here T_{ra} and T_{sp} refer to the clock at a parent, while T_{sa} and T_{rp} refer to the clock at its child. Thus, it does not require a synchronous clock among active nodes for *RTT* measurement. To avoid the fluctuation caused by single *RTT* estimate, the *RTT* is smoothed by using weighted average, that is

$$RTT = 7/8 \times RTT' + 1/8 \times RTT_m \quad (2)$$

Here RTT_m is current *RTT* estimate value, and RTT' is the preceding *RTT* value.

In our scheme, the polling period is set to 2s. If there exists data transmission between active nodes during the polling period, data packet and its acknowledgement are used to take the above necessary information for *RTT* estimate, which thus decreases the unnecessary processing and communication load resulted from additional polling.

3.3 Congestion Detection, Report and Control

Each active node in the multicast tree independently monitors the congestion status of each child using the recent *RTT* value and the feedback received from the child. If an active node finds that the recent *RTT* between itself and a child is greater than the average value of that *RTT*, it deems that there may exists potential congestion on the link leading to that child, then decreases the rate of sending data to that child according to TCP flux model of stable state, that is

$$T = M / (RTT(\sqrt{2p/3} + 4\min(1, 3\sqrt{3p/8})\, p(1 + 32p^2))) \quad (3)$$

Here *T* is sending rate. *RTT* is recent round trip time, *p* denotes packet loss probability, and *M* is the size of packet.

An active node sends a congestion report to its parent when the length of its active buffer queue reaches the predefined threshold. In this congestion reports, requested rate T_{req} is included, which is expressed as follows:

$$T_{req} = \alpha T_{\min} + (1-\alpha)(\sum_{i=1}^{n} T_i - T_{\min})/(n-1) \quad (4)$$

Here $T_{\min}$ is the minimal sending rate among *n* children of the node that sends congestion report, and T_i is the sending rate of child *i*. The parent node regulates its sending rate to the requesting node after getting the congestion report according to the following expression:

$$T_{act} = \min(T, T_{req}) \quad (5)$$

Once the length of active buffer queue at an active node drops below the predefined threshold and the *RTT* between itself and each child restores normal, it sends a congestion release report to its parent. The parent then increases its sending rate to the node that sends this report according to additive increase and multiplicative decrease (AIMD) rate adaptation policy, that is

$$T_a = \min(T, T' + M / RTT) \quad (6)$$

Here T_a is the sending rate after adaptation, T' is its current sending rate, and *T* is derived from Eq. (3).

4. SIMULATIONS AND RESULTS

We ran several ns-2 simulations to verify the performance of our scheme. In these simulations, the data packet size is 1000bytes, initial sending rate is 64kbps and initial *RTT* is 0.1s. To show the result clearly, we average the sending rate in all simulations.

4.1 Basic Function Test

The topology used in this simulation is shown in Fig. 3. Multicast flows from node *S* to receivers (nodes 5-9). The flows originated at node *S* start at the beginning and end at 14th second. In order to simulate congestion, an UDP based CBR flow, which is used as disturbing flow, is injected at 4th second into the link between node 3 and 8, and it lasts for 2 seconds. Thereafter, we record the rate change at node *S*, 1 and 3, and illustrate them in Figs. 4-6. It can be seen from Fig. 6 that node 3 begins to decrease its sending rate to node 8 immediately after the disturbing flow is injected, and starts to increase its sending rate at 6th second because the disturbing flow is released. Fig.5. shows that node 1 decreases its sending rate at 4.06th second because of congestion indication from node 3, and starts to increase sending rate at about 5th second due to the fact that the active buffer queue at node 3 drops. We can also see that the sending rate of node *S* is not affected in the whole process.
This simulation shows that each active node can successfully conduct nearby congestion detection and automatically adjust its sending rate based on congestion detection. Only when necessary, it reports to its parent to avoid making the congestion worse. Therefore, the hop-by-hop congestion control in our scheme can effectively avoid potential congestion, limit the congestion within a part of network and maximizes the throughput of network.

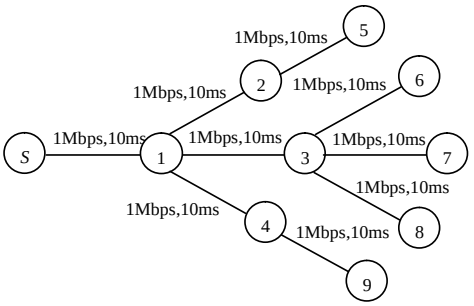


Fig. 3. The topology used for basic function test

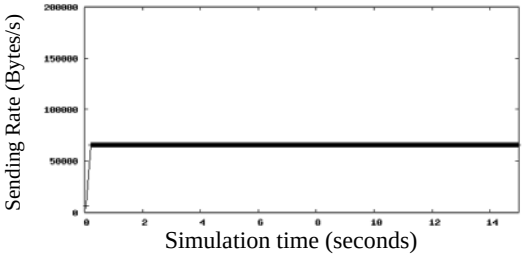


Fig. 4. Basic function test: the sending rate of node *S*

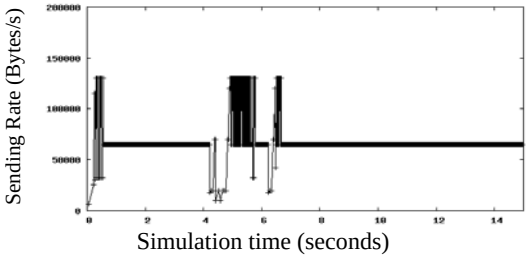


Fig. 5. Basic function test: the sending rate of node 1

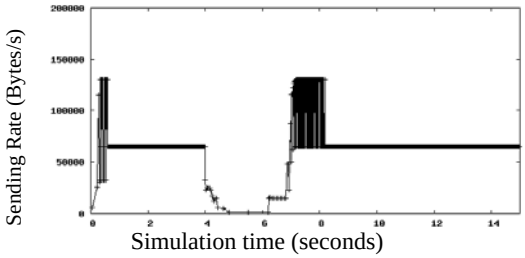


Fig. 6. Basic function test: the sending rate of node 3

4.2 Comparison with End-to-end Control

In this simulation, we compare our control scheme with end-to-end control on the topology in Fig. 7, in which the link bandwidth is 1Mbps and link delay is 10ms.
The following two steps are conducted in the simulation:
1) Multicast flows from node *S* to all 3 receivers (nodes 7-9). The flows originated at node S start at the beginning and end at 10th second. A disturbing flow is introduced into the link between node 5 and node 6 at 4th second, and it lasts for 2 seconds. Thereafter, the change of status at node 4 and node 5 is recorded, and illustrated in Figs.8-9, in which 0, 1, 2 represent congestion release (or normal), congestion avoidance and congestion control respectively.
2) TCP flow originated at node *S* flows to node 9. A same disturbing flow as the one described in step 1) is introduced into the link between node 5 and node 6, and TCP end-to-end control is adopted in this step. We then record the change of status at node *S*, and illustrate it in Fig. 10.
It can be seen from Figs. 8-10 that node 5 enters into the phase of congestion avoidance immediately when the disturbing flow is injected, and restores to normal state once

the disturbing flow is released, while node *S* begins to decrease its sending rate at 4.10th second and restores to normal state at 6.2th second in the case of end-to-end congestion control. It usually takes more time for end systems to detect congestion and comes back from congestion because of link delay, and the more hops from end system to the place of congestion, the more time it takes. Moreover, during the period of congestion, node S keeps in the phase of congestion control in the case of end-to-end control scheme, while in our scheme, node S is not affected and the node nearby the congestion is involved in congestion control for small part of time.

Therefore, it is impossible to process congestion in time for place policy on end system. Compared with end-to-end congestion control, active network based hop-by-hop congestion control responses more quickly, and only limited nodes are affected, while end-to-end congestion control has effect on almost any upstream nodes.

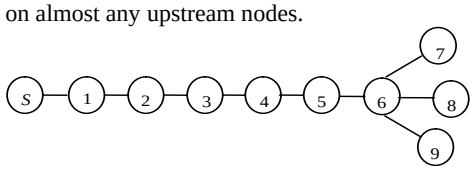


Fig. 7. The topology used for comparison test

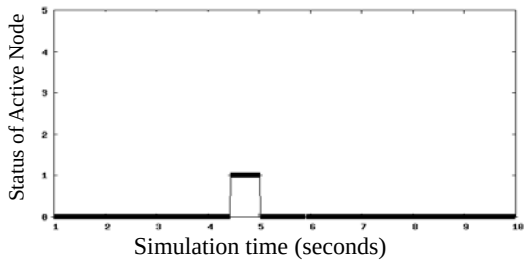


Fig. 8. Active network based control: status change at node 4

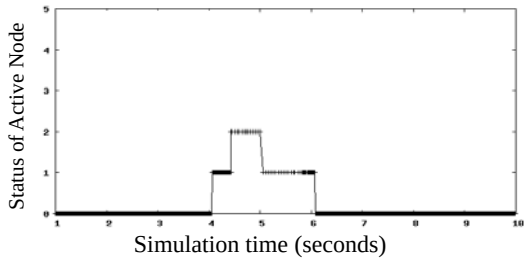


Fig. 9. Active network based control: status change at node 5

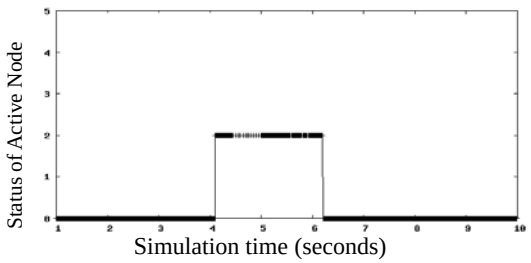


Fig. 10. End-to-end control: status change at node *S*

4.3 TCP-friendliness Test

The topology in Fig. 11 is used for TCP friendliness test. In this simulation, multicast flows from node *S*1 to three receivers (nodes 3-5). TCP flow originated at *S*2 flows to node 5. The link between node 2 and node 5 are thus shared by TCP flow and multicast flow. Multicast flows start at the beginning and end at 10th second, and TCP flows start at 4th second and end at 6th second.

Node *S*1 sends data at an initial rate of 1Mbps (i.e.131072 bytes/s). The change of sending rate of multicast flow at node 2 is recorded and illustrated in Fig. 12. It can be seen that multicast flow and TCP flow share the link between node 2 and node 5 after 4th second, and the average rate of multicast is about 74000 bytes/s, i.e.0.564Mbps, and it implies that the average rate of TCP flow is about 0.436Mbps. Therefore, the bandwidth used by multicast flows is 12.8% more than the one used by TCP flows. It is reasonable that multicast flows use a little more bandwidth because there are multiple receivers in a multicast session [4]. So, our congestion control scheme can be considered as TCP-friendly.

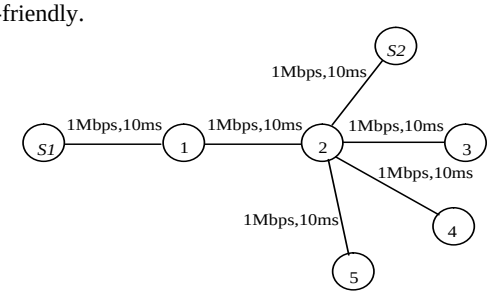


Fig. 11. The topology used for TCP-friendliness test

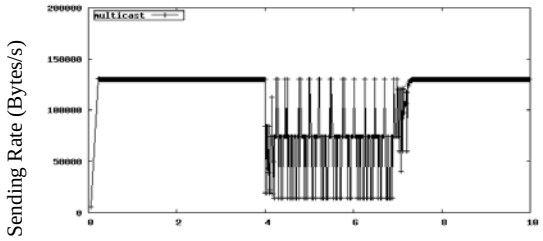


Fig. 12. TCP-friendliness test: sending rate of node 2

5. CONCLUSIONS

We have presented an active network based, multi-rate and TCP-friendly congestion control scheme for reliable multicast. The scheme is designed to effectively handle multiple instances of congestion occurring simultaneously at various parts of multicast tree, and we have obtained the expected results through several simulations. In particular, our results indicate that 1) the proposed congestion control scheme can quickly respond to congestion anywhere in the multicast tree, 2) our scheme supports multi-rate multicasting, 3) our scheme is TCP-friendly, 4) hop-by-hop congestion detection, report and control can effectively limit congestion within parts of network and maximize the utilization of network resources. Thus, we believe that our scheme provides an effective solution to congestion control for large-scale multicast.

6. REFERENCES

[1] L. David, “A Survey of Active Network Research,” IEEE Communications Magazine, January 1997, 35(1): 80-86.

[2] E. Amir, “An Active Service Framework and its Application to Real-time Multimedia Transcoding,”

Proceedings of ACM SIGCOMM 98, September 1998, 178-189.

[3] D.S. Decasper, B. Plattner, "A Scalable High-Performance Active Network Node," IEEE Network，Jan./Feb.1999，8-9.

[4] J. Widmer, R. Denda, et al, "A Survey on TCP-Friendly Congestion Control," IEEE Network, 15 (3): 28-37.

[5] R. Injong, Nallathambi, et al, "MTCP: scalable TCP-like congestion control for reliable multicast", Computer networks, 2002, 38: 553-575.

[6] Q. Dong, J.Y. Xie，"A rate based congestion control algorithm in networks with coexisting unicast and multicast sessions", Journal of university of science and technology Bejing, 2003,10(3): 66-72.

A non-cooperative mixed strategy game framework for pricing in networks

Jing Chen
Department Of Computer Science, Central China Normal University,
Wuhan, 430072 China
School Of Computer Science, Wuhan University
Wuhan, 430079 China
Email: dangcinglulu@mail.ccnu.edu.cn
And
Quan Lu
School of Information Management, Wuhan University
Wuhan, 430072 China

ABSTRACT

A non-cooperative game framework for pricing in high-speed networks is presented in this paper. The approach which adopted is referring to mixed strategy, by which the user can modify his budget according to the market condition and the global synchronization can be avoided. The unique Nash Equilibrium of our game framework was analyzed and the simulation results are given. Then we finish the paper by drawn some conclusions and indicating some directions for future research.

Keywords: non-cooperative game theory, mixed strategy, rate control, pricing.

1. INTRODUCTION

As the internet explodes in size and in the number of users, one of the challenging questions network designers face is how to provide a fair and efficient allocation of the available bandwidth. To this end researchers have proposed many different rate allocation mechanisms. Such mechanisms are delegated to the users that make their decisions independently, according to their individual performance objectives. Since the users are likely to be interested only in maximizing their own benefit, i.e., users are selfish, it is natural to model the problem as a non-cooperative game.

Game theory [1] provides the systematic framework to study and understand the users' behavior. The operating points of a non-cooperative network are the Nash equilibria of the underlying game, that is, the points where unilateral deviation does not help any user to improve its performance. Game theoretic models have been employed in the context of flow control, routing, and admission control; see, e.g., [2, 3] and references therein. For example, Mazumdar et al. [4] first used the NBS in the context of packet-switched (data) networks. The properties of Pareto optimality as well as the development of local optimization procedures which lead to Pareto-optimal solutions (the local procedures being greedy schemes) were studied in a series of papers by Douligeris and Mazumdar [5, 6, 7] in the context of data networks.

Non-cooperative equilibrium are generically inefficient and lead to suboptimal network performance: while each user optimizes its performance objective, the network performance at equilibrium might be suboptimal. Recently, a number of studies have proposed methodologies to overcome this deficiency [2, 4, 5]. In the present study [15] investigate the role of pricing mechanisms as a means to induce efficient allocation of resources in non-cooperative networks. Pricing as an allocation mechanism that makes decentralized decisions compatible with overall efficiency has been studied in the context of queueing systems; see, e.g., [16]. In computer networks, pricing has been receiving increasing attention from both the academic and the corporate world, mostly due to the explosive growth of the Internet that is evolving from a heavily subsidized network to a commercial enterprise [17,18,19,20,21].

Network efficiency is typically defined as the enforcement of a Nash equilibrium that minimizes a social cost function [16, 19]. This assumption leads to the introduction of "shadow prices," that charge each user for the "congestion externality," i.e., the dissatisfaction its actions cause to the other users.

[15] then address the issue of pricing and its relation to bandwidth allocation. [15] is shown that based on a user's budget or willingness-to-pay and its bandwidth demands, a bargaining framework can be developed to allocate the network bandwidths to the users in a way which is optimal in the Pareto sense and is fair to the users.[15] Furthermore, based on this, a pricing scheme based on the congestion in the network for which network revenue is maximized when the network operates at the allocations corresponding to the bargaining solution. This pricing scheme has the following property: a user is never charged more than its declared budget but could be charged less than its budget if the amount of congestion in the network links used by its connection is low. However, once the user determines his budget or his willing to pay in advance, the user's budget cannot be changed in the pricing scheme. This is impractical and does not tally with the market mechanism. But if the user can modify his budget, the global synchronization may occur.

Mixed strategies, limiting an opponent's options by mixing up one's actions, are of fundamental importance in strategic interaction. A mixed strategy is one in which a player plays his available pure strategies with certain probabilities. It indicating that rational play is not in general describable by specifying a single pure strategy. Rather, it is often non-deterministic, with specified probabilities associated with each one of a specified set of pure strategies. This "mixed" security level is always at least as high as the 'pure' one.

In this paper, we propose a new pricing scheme based non-cooperative network framework, which not only guarantee the fairness and efficiency of bandwidth allocation, but also can modify his budget according to the market condition so that the global synchronization can be avoided. The method used in our study is known as mixed strategy. Then, the non-cooperating framework can allocate the network bandwidths to the users based on a user's budget or willingness-to-pay and its bandwidth demands intelligently.

The outline of this paper is as follows: section 2 describe our non-cooperative pricing game framework which using mixed strategy to alleviate global synchronization. Section 3 analyses the unique Nash Equilibrium of our game framework and the simulation result also prove our illustration; conclusions are drawn and some directions of future work are highlighted in section 4.

2. A Non-cooperative Mixed Strategy Game Framework

In this section, we present the non-cooperative mixed strategy pricing game model and the mixed strategy Nash equilibria which are used in the sequel. For details, we refer the reader to the book by Muthoo [8] and the paper by Nash [9].

We consider a network with a set of *J* resources or links and a set of *I* users. Let C_j denote the finite capacity of link $j \in J$. Each user has a fixed route J_i, which is a non-empty subset of J. We define a zero-one matrix A, where $A_{i,j}=1$ if link is in user *i*'s route J_i and $A_{i,j}=0$ otherwise. When the throughput of user is x_i, user *i* receives utility $U_i(x_i)$. We assume that the utility function $U_i(x_i)$ is an increasing, strictly concave and continuously differentiable function of x_i over the range $x_i \geq 0$. Furthermore, we assume that the utilities are additive so that the aggregate utility of rate allocation is $\sum_{i\in I} U_i(x_i)$.Under our model this problem can be formulated as follows.

SYSTEM (U,A,C):

$$\begin{aligned} &\max \sum_{i\in I} U_i(x_i) \\ &st.A^T x \leq C \\ &over \quad x \geq 0 \end{aligned} \tag{1}$$

The first constraint in the problem says that the total rate through a resource cannot be larger than the capacity of the resource. Given that the system knows the utility functions U of the users, this optimization problem may be mathematically tractable. However, in practice not only is the system not likely to know, but also it is impractical for a centralized system to compute U and allocate the user rates, due to the large scale of the network. Hence, Kelly in [10] has proposed to consider the following two simpler problems.

Suppose that given the price per unit flow λ_i, user *i* selects an amount to pay per unit time, p_i, and receives a flow $\frac{p_i}{\lambda_i}$.Then, the user's optimization problem becomes the following[10].

USER$_i$(U_i, λ_i):

$$\begin{aligned} &\max \; U_i(\frac{p_i}{\lambda_i}) - p_i \\ &over \; p_i \geq 0 \end{aligned} \tag{2}$$

The network, on the other hand, given $p=(p_i, i\in I)$,attempts to maximize the sum of weighted log functions $\sum_{i\in I} p_i \log(x_i)$. Then the network's optimization problem can be written as follows [10]:

NETWORK(A,C;p):

$$\begin{aligned} &\max \; \sum_{i\in I} p_i \log(x_i) \\ &st. \; A^T X \leq C \\ &over \; x \geq 0 \end{aligned} \tag{3}$$

Note that the network does not require the true utility functions $(Ui(.), i\in I)$, and pretend that user *i*'s utility function is $p_i \bullet \log(x_i)$ to carry out the computation [9].

We have analyzed the case where users have fixed their target queue sizes $p=(p_i, i\in I)$ (it can also be seen as the users' budget or willing to pay). However, as more intelligence is embedded in end systems in the future, the users may be able to vary the parameters p_i to maximize their net utility given by (2). We assume that the utility functions satisfy the following: $p_i(\lambda_i)$ is decreasing in $\lambda_i \rangle 0$, where $p_i(\lambda_i) = \arg\max_{pi} U_i(p_i/\lambda_i(t^-)) - p_i$ .And $\lambda_i(t^-)$ is the price per unit flow along user *i*'s route at time t^- resulting at the unique stable point of the (p,1)-proportionally fair algorithm. We assume that the price updates take place in a much larger time scale, while users allow their window sizes to converge to a point close to the unique stable point of TCP for fixed($p_i(t), i\in I$). This is a natural assumption since the window sizes typically take only seconds to converge and users are not likely to keep changing their parameters before estimating the current throughput. The intuition behind the updating rule is as follows. At time t based on the price per unit flow at time t^- user *i* computes its optimal price that maximizes its net utility. If the current price per unit flow is too high, user prefers to wait till the price per unit flow is lower. In such a case user needs to probe the network for the available bandwidth and price per unit flow along its route. In order to measure the residual bandwidth not used by the other users, if there is any, and the price per unit flow, user needs to set its window size large enough so that it utilizes all residual capacity not taken by the other users [14]. Hence, [14] assume that user sets the target queue size, which is its willingness to pay in their model, according to:

$$\begin{aligned} p_i(t) &= \arg \max_{p_i} U_i(\frac{p_i}{\lambda_i(t^-)}) - p_i \\ &= \left\{ \begin{array}{c} 0, \quad if \quad \lambda_i(t^-) > U_i{}'(0) \\ p_i \quad such \quad that \quad U_i'(p_i/\lambda_i(t^-)) = \lambda i(t^-) \\ if \quad 0 < \lambda_i(t^-) < U_i{}'(0) \\ Ki \quad , if \quad \lambda_i(t^-) = 0 \end{array} \right\} \end{aligned} \tag{4}$$

We assume that $K_i < \infty$ for all user *i*. This is a reasonable assumption, because at very high rates users' marginal utility is likely to be very small and users are not likely to pay a price larger than a certain limit, for instance, given by a budget constraint.

However, if the price per unit flow is larger than or equal to $U_i'(0)$ for most users, they will receive a negative net utility from any positive p_i. Thus, they should wait till the

price per unit flow λ_i is smaller than $U_i'(0)$, which in turn will cause instant idle network resource. If the price per unit flow of all users are strictly smaller than their $U_i'(0)$, then there exists a unique solution to the problem in (2). This solution is the unique p_i such that $U_i'(p_i/\lambda_i(t^-))=\lambda_i(t^-)$, which maximizes user i's net utility. It's reasonable and inevitable that the user will increase its budget p_i if the price per unit flow along user i's route at time t^- is smaller than the price at the earlier time, which will result in congestion. This is known as global synchronization.

If user modifies its budget with a certain probability, the global synchronization will be alleviated. The approach we adopt here is referred to as mixed strategy, in the term of game theory. Suppose a user has pure strategies $p_{i1}, p_{i2}, p_{i3},...p_{in}$ in a normal form game. A mixed strategy for the player is a probability distribution over $p_{i1}, p_{i2}, p_{i3},...p_{in}$; i.e. a mixed strategy has the form $\sigma_i=\delta_1 p_{i1}+\delta_2 p_{i2}+...+\delta_n p_{in}$ where $\delta_1,...,\delta_n$ are all nonnegative and $\sum_1^n \delta_j=1$.By this we mean that the player chooses p_{ij} with probability δ_j, for $j=1,2,...n$. Thus, every modification of budget for user can be seen as a non-cooperative static game, in which the user is the player, and weather user updates their budget or not is determined with a certain probability. We see any user modification for his budget as a non-cooperative game, and a user plays his available pure strategies with certain probabilities (here is 50% probability).So users update their willingness to pay or budget $p_i, i\in I$, at time t, with mixed strategy, according to:

$$p_i=\left\{\begin{array}{l} if\ \lambda_i(t^-)>U_i'(0):\left\{\begin{array}{l} 0,\ with\ 50\%\ probability \\ p_i(t^-),\ with\ 50\%\ probability\end{array}\right\} \\ if\ 0<\lambda_i(t^-)<U_i'(0): \\ \left\{\begin{array}{l} p_i,\ such\ that\ U_i'(p_i/\lambda_i(t^-))=\lambda i(t^-), \\ \quad with\ 50\%\ probability \\ \\ p_i(t^-),\ with\ 50\%\ probability\end{array}\right\} \\ if\ \lambda_i(t^-)=0,\left\{\begin{array}{l} Ki, with\ 50\%\ probability \\ p_i(t^-), with\ 50\%\ probability\end{array}\right\}\end{array}\right\} \tag{5}$$

3. ANALYSIS AND SIMULATION RESULT

When mixed strategies are applied, payoff must be replaced by expected payoff, and individual rationality may be thought of in terms of mixed strategies. In this case, what to be "guaranteed" is not an actual payoff, but an expectation; the word "guarantee" means that this level of payoff can be attained in the mean, regardless of what other players do. This "mixed" security level is always at least as high as the 'pure' one.

Equilibrium (Nash, 1951) of a strategic game is a (pure or mixed) strategy profile in which each player's strategy maximizes his payoff given that the others are using their strategies.

So the network solves the following maximization problem NB:

$$\begin{array}{l} Max \sum_{i\in I}((p_i\log(x_i)\times 0.5+p_i^{/}\log(x_i)\times 0.5) \\ st.\ A^T x\le c \\ over\ x\ge 0\end{array} \tag{6}$$

Where p_i is the budget of user i which is according to (4), and $p_i^{/}$ is the budget of user i in the earlier time. Because the window sizes typically take only seconds to converge and users are not likely to keep changing their parameters before estimating the current throughput, we assume p_i and $p_i^{/}$ are constants according to(5) during the network the user i's sending rate and suppose that the user modify his willing to pay in terms of 50 percents probability.

Now we will prove there is a unique Nash equilibrium point in our mixed strategy non-cooperative game model which is characterized as follows:

There exist $\mu_j\ge 0(j\in\{1,2,...,J\})$ such that for each $i\in\{1,2,...,I\}$

$$\begin{array}{l} xi=\dfrac{p_i+p_i^{/}}{2(\sum_{j=1}^{J}\mu_j a_{ji})} \\ Ax\le C \\ (Ax-C)_j\mu_j=0\quad j\in\{1,...,J\}\end{array} \tag{7}$$

We all know the aim of network is to maximize the function $f(x)=\sum_{i\in I}((p_i\log(x_i)\times 0.5+p_i^{/}\log(x_i)\times 0.5)$ .Obviously, this function is convex, and the problem is a convex programming problem. Noting that the constraints are linear in {xi} and the function? is convex, it implies that the first-order Kuhn-Tucker[12,15] conditions are necessary and sufficient for optimality.

Let $L(x,\mu)$ denote the Lagrangian where $\mu_j\ge 0; j=1,2,...,j$ associated with capacity constraints denote the Lagrangian multipliers. Then

$$\begin{array}{l} L(x,\mu)=(\sum_{i\in I}((p_i\log(x_i)\times 0.5+p_i^{/}\log(x_i)\times 0.5) \\ \quad -\sum_{j=1}^{J}\mu_j((Ax)_j-C_j)\end{array} \tag{8}$$

Then the first-order necessary and sufficient conditions are given by

$$\begin{array}{l} (\dfrac{p_i+p_i^{/}}{2x_i})-\sum_{j=1}^{J}\mu_j a_{ji}=0; \\ i=1,2,...,I\end{array} \tag{9}$$

and

$$\begin{array}{l} ((Ax)_j-C_j)\mu_j=0; \\ \mu_j\ge 0; j=1,2,...,J\end{array} \tag{10}$$

The Lagraniplier μ_j has the interpretation as the implied cost associated with the network link j. It represents the marginal cost of a rate unit allocated for any connection-crossing link and obviously $\sum_{j=1}^{J}\mu_j a_{ji}$ is to be the equal of the price per unit flow λ_i along user i's route, which we mention above.

Note that the network does not require the true utility functions $(Ui(.), i\in I)$, and pretend that user i's utility function is $p_i\bullet\log(x_i)$ to carry out the computation [10]. It

is shown in [10] that one can always find vetors $\lambda^* = (\lambda_i^*, i \in I), p^* = (p_i^*, i \in I), x^* = (x_i^*, i \in I)$ such that p_i^* solves $USER_i(U_i, \lambda_i^*)$ for all $i \in I$, x^* solves $NETWORK(A, C; p^*)$, and $p_i^* = x_i^* \bullet \lambda_i^*$ for all $i \in I$. Further, the rate allocation x^* is also the unique solution to $SYSTEM(U, A, C)$. The problem of solving $SYSTEM(U, A, C)$ can be achieved by an algorithm that solves $NETWORK(A, C; p(t))$ for a given $p(t)$ at time t on a smaller time scale, and drive p(t) to p^* on a larger time scale. So, our scheme is available and more intelligence.

A mixed strategy equilibrium is a mixed strategy profile $(x_1^*, x_2^*, ..., x_n^*)$ such that, for all $i = 1, 2, ..., n$ $x_i \in \arg\max_{x_i} u_i(x_i, x_{-i}^*)$ where $u_i(.)$ is user i's utility function. Apparently, the above user problem has unique mixed strategy equilibrium. Hence, the result follows as stated.

In [13] also proved that in the case of simple network where there is a single bottleneck link with finite capacity and a finite number of users, there exists a unique Nash equilibrium of the game. Though [13] limit their analysis only to relative simple networks of a single bottleneck and only consider pure strategy space, one would expect the same intuition to be extended to more complicated networks and mixed strategy space. (If each player in an n-player game has a finite number of pure strategies, then there exists at least one equilibrium in (possibly) mixed strategies. (Nash proved this).)

In this section, we also give a numerical example of a simple network and demonstrate the convergence of user's parameters through simulation. Although the convergence results have been proved only for single bottleneck cases, the simulation results indicate the the user converge to the system optimal rates even in general networks with the utility functions which satisfy the assumption: $p_i(\lambda_i) = \arg \max_{pi} U_i(p_i / \lambda_i(t^-)) - p_i$.

We run the following ns2 simulations on the dumbbell network topology showed in Fig. 1. the source destination pairs of the users are given in the figure, and the capacities and delays of the links are indicated by the numbers next to the links. The utility functions of the user1,user2 and user3 are $5\log(x_1 + 1)$, $3\log(x2 + 1)$ and $2\log(x3 + 1)$ respectively. The prices and rates of some of the users are plotted in Fig. 2. The dotted lines in the plots represent the optimal rates and unique equilibrium prices, respectively. These plots clearly demonstrate the convergence of the user prices and, thus, the rates to the system optimum.

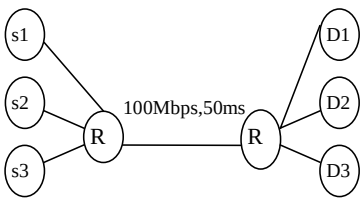


Fig. 1. The dumbbell topology used in the simulations

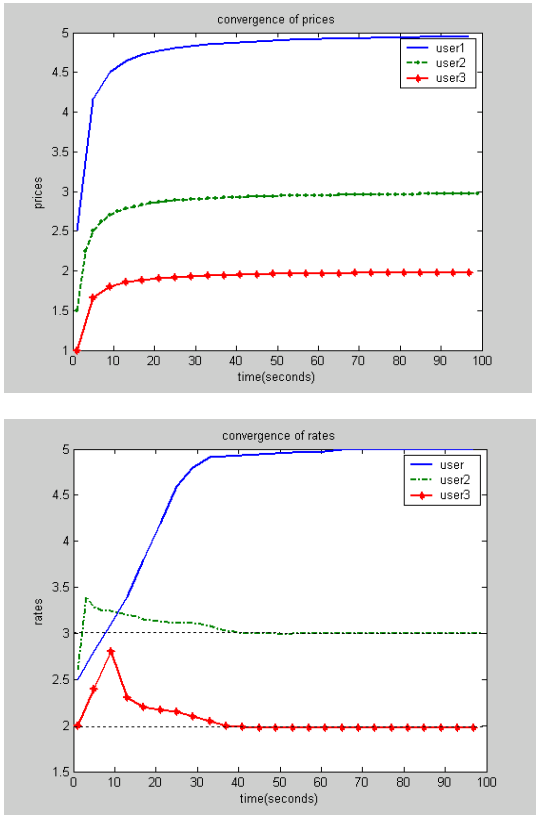


Fig. 2. Convergence of user prices and rates

4. CONCLUSION

All equations must be typed or written neatly in black. They should be numbered consecutively throughout the text. Equation numbers should be enclosed in parentheses and flushed right. Equations should be referred to as Eq. (X) in the text where X is the equation number. In multiple-line equations, the number should be given on the last line.

Currently, most price schemes based on congestion control in the network considering fairness and efficiency have the following property: a user is never charged more than his declared budget but can be charged less than his budget if the amount of congestion in the network links used by his connection is low. However, once the user determines his budget or his willing to pay in advance, it cannot be changed in the pricing scheme. This does not correspond with the reality market mechanism and is not optimal. But if the user can modify his budget, the global synchronization may occur. In this paper, we present a new pricing scheme based on non-cooperative game framework for the allocation of optimal rates to elastic connections which share common bandwidth to not only achieve fairness and efficiency but also avoid the global synchronization. The unique Nash Equilibrium of our game model is analyzed and the simulation result is also given.

The future work will address the issues of the algorithmic implementation in the context when the fact that real situations involve non-static scenarios and evolving game theory will be discussed.

5. REFERENCES

[1] R. B. Myerson, Game Theory: analysis of conflict

[2] Cambridge, Massachusetts: Harvard University Press, 1991.
[3] Y. A. Korilis, A.A. Lazar, et al, Architecting Non-cooperative Networks, IEEE Journal on Selected Areas in Communications, vol. 13, pp. 1241-1251, September 1995.
[4] Z. Dziong, L.G. Mason, Fair-Efficient Call Admission Control Policies for Broadband Networks – A Game Theoretic Framework, IEEE/ACM Transactions on Networking, February 1996，vol. 4, 123–l36.
[5] R.R. Mazumdar, L. Mason, et al, Fairness in network optimal flow control: Optimality of product forms, IEEE Trans. Commun., May 1991，vol. 39, 775–782.
[6] C. Douligeris, R.R. Mazumdar, On Pareto-optimal flow control in an integrated environment, in Proc. 25th Allerton Conf., Urbana, IL, Oct. 1986.
[7] C. Douligeris, R. Mazumdar, More on Pareto-optimal flow control, in Proc. 26th Allerton Conf., Urbana, IL, Oct. 1987.
[8] C. Douligeris, R. Mazumdar, User optimal flow Control in an integrated environment, in Proc. Indo-U.S. Workshop Syst. Signals, Bangalore, India, Jan. 1988.
[9] A. Muthoo, Bargaining Theory with Applications. Cambridge, U.K.: Cambridge Univ. Press, 1999.
[10] J. Nash, The bargaining problem, Econometrica, 1950，vol. 18， 155–162.
[11] F. Kelly, Charging and rate control for elastic traffic. European Transactions on Telecommunications, January 1997，vol.8(1):33–7.
[12] F. Kelly, A. Maulloo, et al, Rate control for communication networks: shadow prices, proportional fairness and stability. Journal of the Operational Research Society, March 1998，vol.49(3):237–52.
[13] M. Minoux, Mathematical Programming: Theory and Algorithms. Chichester, U.K.: Wiley, 1986.
[14] R. J. La, V. Ananthram, Charge-sensitive TCP and rate control in the internet, in Proc. IEEE INFOCOM, vol. 3, Mar. 2000, 1166–1175.
[15] R.J. La, V. Anantharam, Utility-based rate control in the Internet for elastic traffic, IEEE/ACM Transactions on Networking, April 2002， 10（2）： 272-286.
[16] H. Yaiche, R.R. Mazumdar, et al, A game theoretic framework for bandwidth allocation and pricing in broad-band networks, IEEE/ACM Trans. Netw., Oct. 2000， 8（5）： 667-678.
[17] H. Mendelson, S. Whang, Optimal Incentive-Compatible Priority Pricing for the M/M/l Queue, Operations Research, 1990,38（5）: 870–883.
[18] S.H. Low, P.P. Varaiya, A New Approach to Service Provisioning in ATM Networks, IEEE/ACM Transactions on Networking, 1993,1（3）: 547–553.
[19] R.J. Eden, N. McKeown, et al, Billing Users and Pricing for TCP, IEEE Journal on Selected Areas in Communications, September 1995, vol. 13, 1162–1175.
[20] J.K. MacKie-Mason, H.R. Varian, Pricing Congestible Network Resources, IEEE Journal on Selected Areas in Communications, September 1995,vol. 13,1141–1149.
[21] J. Sairamesh, D.F. Ferguson, et al, An Approach to Pricing, Optimal Allocation and Quality of Service Provisioning in High-Speed Packet Networks, in Proceeding of the IEEE INFOCOM'95, (Boston, MA), April 1995.

The Analysis and Design of Turbo Receiver In Downlink Mimo CDMA

Mingxin Tan,
The Department of Information Technology, Central Chinese Normal University
Wuhan City， Hubei Province, 430079, China
Email: tmingxin@yahoo.com.cn
Jia Shen,
Research Institute of Telecom Transmission, CATR of MII, China
Beijing City，100083, China
Email: shenjia@mail.ritt.com.cn
Ning He
The Computer Center, Wuhan University
Wuhan City, Hubei Province, 430079, China
Email: hening@126.com

ABSTRACT
In this paper the performance of the Turbo receiver is evaluated in the downlink space-time Turbo coded CDMA system. Long-code scrambling is considered and the Turbo receiver is designed based on the space-time RAKE receiver. An antenna-specific scrambling scheme is also proposed to improve the receiver performance in some scenarios. Simulation results show that the iterative receiver can achieve higher downlink capacity for this system than the non-iterative one.

Keywords: Iterative Decoding, MIMO, Downlink, CDMA

1. INTRODUCTION

In radio communications multi-user detection (MUD) is an attractive technique in uplink CDMA systems because the base station (BS) naturally needs to detect all users in the cell, and can tolerate the complexity of MUD. On the other hand, using MUD in the downlink is a controversial proposal because MUD is possibly too complex for a mobile terminal. Despite the difficulty of implementation, some research work has been carried out in this field [1]. MUD in downlink CDMA can be employed to combat inter-cell interference as well as intra-cell interference. The former arises between different BSs that use non-orthogonal scrambling codes. Intra-cell interference occurs between the code-division channels. These channels are spread by distinct orthogonal channelization codes (such as Walsh codes) and are always synchronous. However, in a frequency-selective channel, the multipath spread will distort the orthogonality between the channelization codes, and introduce intra-cell interference. A number of MUD approaches have been proposed to suppress the downlink intra-cell interference [2, 3]. However, MUD is not the only solution to the problem. Another, perhaps simpler solution is chip equalization. From the point of a mobile station (MS), the downlink CDMA is actually a point-to-point communication. All code-division channels go through the same channel. Therefore, if the equalization can effectively correct the channel distortion, the orthogonality between code-division channels will be restored. Then a simple despreading process can separate these channels.

However, things become more complicated in a MIMO system, especially a spatial multiplexing (SM), CDMA system. In the system, a BS with n_T antennas should be regarded as n_T independent BSs with K channels respectively. Although they are using the same set of K spreading codes, these codes are distorted by different channels.

The basic idea of chip equalization is to restore the code orthogonality by correcting the channel distortion. Nevertheless this advantage will not remain in MIMO CDMA because the inter-antenna interference (IAI) is the inherent problem of MIMO systems. In order to restore the orthogonality between the code-division channels, equalization needs to be implemented in a MIMO channel. As is well known, MIMO equalization itself is a complex problem, and still under investigation [4]. The iterative MIMO equalizer is proposed in [5]. An alternative solution to this problem is based on MUD approaches. MIMO MUD receivers based on linear detection and interference cancellation (IC) are introduced in [6, 7].

Based on the discription above, two approaches can be adopted in downlink MIMO CDMA systems: space-time MUD and MIMO equalization plus despreading. The latter approach can be found in [8], which is based on non-iterative structures. Reference [9] proposes a hybrid receiver including one-stage chip equalizer and multi-stage PIC detector. The iterative receiver in [10] deals with the different interference compositions separately. It suppresses the MAI and ISI in a linear fashion, while combating the IAI iteratively. This paper investigates the performance of iterative receiver in Downlink MIMO CDMA.

It is arranged as follows: Section 2 describes the system model; Section 3 gives simulation results and analysis. Section 4 concludes the paper.

2. SYSTEM MODEL AND ANALYSIS

2.1 Downlink STTuC MIMO CDMA System

The structure of the downlink STTuC CDMA transmitter is illustrated in Figure. 1. The data in K traffic channels are encoded by separate Turbo encoders and demultiplexed to n_T streams by

Space-Time Demultiplexer. The sub-streams of one channel are spread by the same orthogonal variable spreading factor (OVSF) codes (channelization codes). Spreading factor for all channels is N. The spread sub-streams allocated on an antenna are compounded together and then multiplied by the long scrambling codes before they are transmitted. Here the long codes are modeled with uniformly-distributed random binary sequences. The transmitted data can be written as:

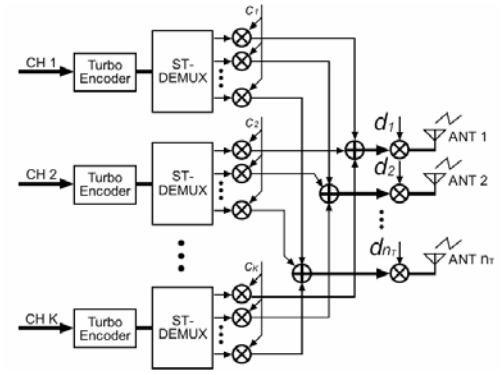


Fig. 1. STTuC CDMA BS Transmitter

$$b = [b[0], b[1], \Lambda\ , b[M-1]]^T$$
$$b[m] = [b_1[m], b_2[m], \Lambda\ , b_K[m]]^T \quad (1)$$
$$b_k[m] = [b_{k,1}[m], b_{k,2}[m], \Lambda\ , b_{k,n_T}[m]]^T$$

Here two scrambling configurations are considered at the base station (BS): BS-specific scrambling (BSS) and antenna-specific scrambling (ASS). The former uses the same scrambling code on all antennas of one BS. This scheme does not change the basic CDMA code allocation, and the receiver wholly depends on the antennas' different channels to distinguish the Space-Time Turbo Codes (STTuC) codewords. The receiver will possibly suffer from severe IAI. In order to effectively reduce the IAI, distinct scrambling codes can be allocated to different antennas. The use of ASS introduces new orthogonality between different antennas and reduces the risk of high "Self IAI". Even when the channels of two antennas are strongly correlated, orthogonality can be guaranteed by the different scrambling codes. This is a practical option because the number of valid scrambling codes in the downlink is always much larger than the number of BSs, although the scrambling codes are a scarce resource in the uplink. It is first assumed that the ASS scheme is used. The spreading sequence for a given symbol in the k-th traffic channel for the m-th symbol period is the multiplication of the k-th OVSF and the corresponding N bits of the long scrambling code.

$$s_{k,m} = [s_{k,m}[1], s_{k,m}[2], \Lambda\ , s_{k,m}[N]]^T \quad (2)$$

where the n-th chip value of $s_{k,m}$:

$$s_{k,m}[n] = c_k[n] d_m[n] \quad (3)$$

and $s_k = [s_k[1], s_k[2], \Lambda\ , s_k[N]]$ is the k-th channelization code and $d_m = [d_m[1], d_m[2], \Lambda\ , d_m[N]]$ is the scrambling code in the m-th symbol period. The K channels are assumed to transmit their signals synchronously and all the STTuC codewords of one channel are aligned. The equivalent spreading code in equation (4) changes for different symbols because of the aperiodic scrambling, and the Channel Impulse Response (CIR) $h_{c,d}$ is the same for all data channels

$$g_{k,m,c,d}(t) = s_{k,m}(t) h_{c,d}(t)$$
$$= \sum_{n=1}^{N} s_{k,m}[n] \sum_{l=1} h_{c,d}[l] q_C(t-(n-1)T_C - \tau_l^F) \quad (4)$$

Since there is no relative delay between data channels($\tau_k = 0$), the signal received by the d-th antenna in the m-th symbol period is expressed as:

$$r_d(t) = \sum_{k=1}^{K} \sum_{c=1}^{n_T} \sum_{m=0}^{M-1} \frac{A_k}{\sqrt{n_R}} b_{k,c}[m] g_{k,m,c,d}(t - iT_S) + \sigma n_d(t) \quad (5)$$

where A_k is the average received amplitude for channel k. Writing equation (4) in discrete time form:

$$g_{k,m,c,d}(j) = \sum_{n=1}^{N} s_{k,m}[n] \sum_{l=1}^{L} h_{c,d} \delta(j-(n-1)-n_l^F) \quad (6)$$

Grouping the values in (6) into a $(N + n_L^F)$-vector (equivalent spreading code):

$$g_{k,m,c,d} = [[g_{k,m,c,d}[0], g[1], ..., g[N + n_L^F - 1]]^T \quad (7)$$

The equivalent spreading codes for traffic channel k in the m-th symbol period are grouped into a spreading matrix:

$$g_{k,m} = \begin{bmatrix} g_{k,m,1,1}[0] & \Lambda & g_{k,m,1,n_R}[0] & \Lambda & \Lambda \\ g_{k,m,2,1}[0] & \Lambda & g_{k,m,2,n_R}[0] & \Lambda & \Lambda \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} \\ g_{k,m,n_T,1}[0] & \Lambda & g_{k,m,n_T,n_R}[0] & \Lambda & \Lambda \end{bmatrix}$$
$$\begin{bmatrix} g_{k,m,1,d}[0] & \Lambda & g_{k,m,1,n_R}[N^F - 1] \\ g_{k,m,2,d}[0] & \Lambda & g_{k,m,2,n_R}[N^F - 1] \\ \mathrm{M} & \mathrm{M} & \\ g_{k,m,n_T,d}[0] & \Lambda & g_{k,m,n_T,n_R}[N^F - 1] \end{bmatrix} \quad (8)$$

Where $N^F = (N + n_L^F)$. Then the equivalent spreading matrix for all channels and all symbols within a fading block can be constructed as in Figure. 2. First, for a specific symbol, all channels' spreading matrices are grouped together as in the left of Figure 2 (each solid line expresses a matrix $g_{k,m}$ in equation (8)). The K equivalent spreading codes are grouped in a synchronous form because the K-th channels are always aligned.

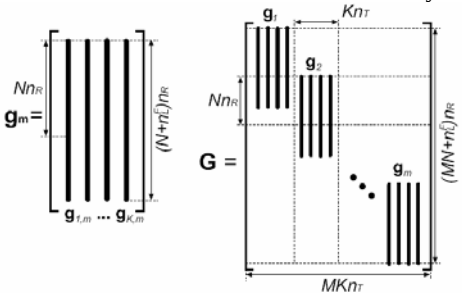


Fig. 2. Equivalent Spreading Matrix For Downlink STTuC CDMA

Naming the spreading matrix for the m-th symbol g_m, and then grouping the equivalent spreading matrix G with g_m s in adjacent shift Nn_R, as in the right of Figure. 2. The equivalent spreading matrix dimension is $((MN + n_L^F)n_R) \times MKn_T$. With the definition of G the received signal can be written in matrix form:

$$r = GAb + n \tag{9}$$

The amplitude matrix is:

$$n = [n[1], n[2], ..., n[(MN + n_L^F)n_R]^T \tag{10}$$

$$r = [r[1], r[2], ..., r[(MN + n_L^F)n_R]^T \tag{11}$$

The combined correlation matrix (CCM) B can be calculated as:

$$B = G^H G \tag{12}$$

2.2 Turbo Receiver for Downlink STTuC CDMA

The Turbo receiver for downlink MIMO CDMA systems shows in Figure. 3. In fact, the last part of figure 3 is Turbo iterative structure, and iterative decoding is the same as reference [11].

We can see Figure. 3. consists of the space-time matched filter (ST-MF), MMSE space-time multi-user detector (MMSE-ST-MUD), space-time multiplexer (ST-MUX) and Turbo decoder. The ST-MF takes charge of combining and dispreading the signals received by multiple antennas. Here figure 4 shows the ST-MF structure. Separate space-time RAKE receivers (ST-RR) are applied to despread and combine the multi-path signals for multiple antennas. Unlike the conventional RAKE receiver, when doing the combining, an ST-RR uses nT sets of channel parameters to compensate the channel distortions corresponding to the nT transmitter antennas separately. Or the ST-RR can be regarded as a bank of conventional RAKE receivers.

Due to the use of aperiodic scrambling codes, the Combined Correlation Matrix (CCM) B has very large dimensions and cannot be inverted directly. The complexity can be reduced using the "sliding window" technique. A sample can be found in [12] that is named "space-time sampling" matrix structure. Since a given symbol only correlates with several adjacent symbols, the large correlation matrix for a K-symbol block can be divided into M small correlation matrices. Each of them contains the correlation between the desired symbol and the correlated symbols. Thus the M matrices can be inverted separately. This makes MMSE detection possible in principle to be used in aperiodic scrambling CDMA. However, the matrix inversion needs to be performed symbol by symbol. This is still too complex for practical implementation. Hence here the MMSE detection is omitted, and only employ the Space-Time Matched Filter (ST-MF) in the first iteration.

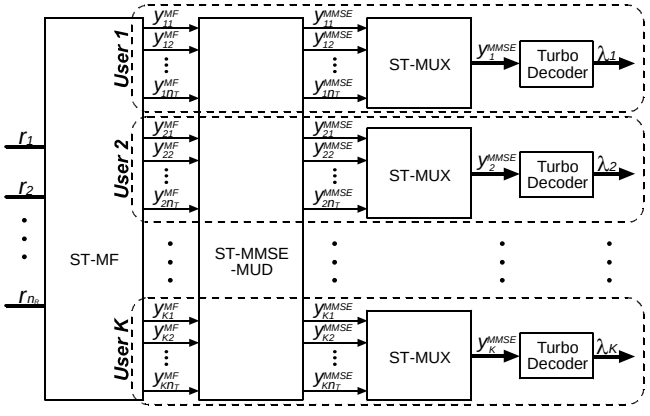


Fig. 3. MMSE Space-Time Receiver For MIMO CDMA

The RAKE processing for the m-th symbol in the k-th channel from the c-th antenna to the d-th receiver antenna is expressed by:

$$g'_{k,m,c,d}(t) = s_{k,m}(t) * h^*_{c,d}(t) \tag{13}$$

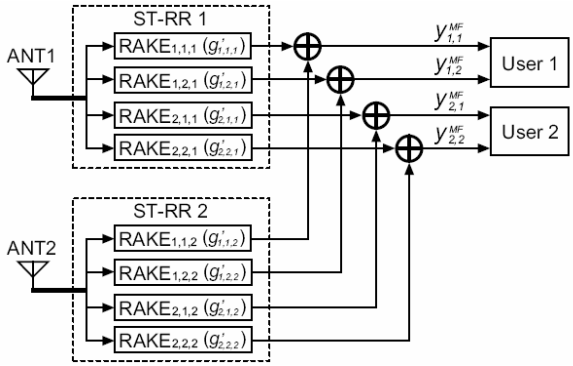


Fig. 4. Space-Time Matched Filter For STTuC CDMA System

Comparing with equation (7), $g'_{k,m,c,d}(t)$ can be defined as the equivalent despreading code. The combined despreading matrix G' can be constructed as the introduced above and illustrated in Figure 2. For CDMA systems using real-value spreading codes, $G' = G^H$. The ST-MF output can be written as:

$$y^{MF} = ABAb + n' \tag{14}$$

The ST-MF outputs are multiplexed by the space-time multiplexer (ST-MUX), and fed to Turbo decoders. If the ST-MF works independently, the Turbo decoding will be performed for as many iteration as needed until a hard decision is made. This receiver can be named "ST-MF receiver". If the ST-MF acts as the first stage of the iterative receiver, only one iteration of Turbo decoding is performed per receiver iteration. The decoders' outputs, the Log-Likelihood Ratio LLR [11] values of data, are converted to their soft estimates, and space-time demultiplexed to reform the STTuC codewords. The mixture of MAI, IAI and ISI will be subtracted from the ST-MF output

3. PERFORMANCE SIMULATIONS AND ANALYSIS

In the simulations, a downlink MIMO CDMA system with 8 traffic channels is considered. The channelization codes are 16-bit OVSF codes and the long scrambling codes are modelled by 16-bit random sequences regenerated for every symbol. The spreading chip period is assumed to be 0.2 μs (chip rate = 5Mcps, data rate =$5 \cdot n_T / T$ = 1.25 Mbps/channel).

3.1 BER Performance: BSS vs ASS

Figure 5 depicts the performance of both the Turbo STTuC CDMA receiver and the ST-MF receiver in a moderate fast urban fading channel. The fading block is 64 symbols. Since the STTuC codeword length is 512 symbols, a frame experiences 8 blocks. The single user Turbo decoder performance in the SISO nonfading AWGN channel serves as the AWGN upper performance bound (no MAI, IAI and ISI exist). The desired data is provided by the hard decision of the decoder output in the last iteration. The average BER value of the 8 channels is taken.

The results in Figure 5 show that, in the moderately fast urban fading channel, the proposed Turbo receiver can effectively suppress the mixture of ISI, IAI, and MAI. It shows desirable convergence characteristics and reaches a covered by one fading block) is about 0.7dB worse than the

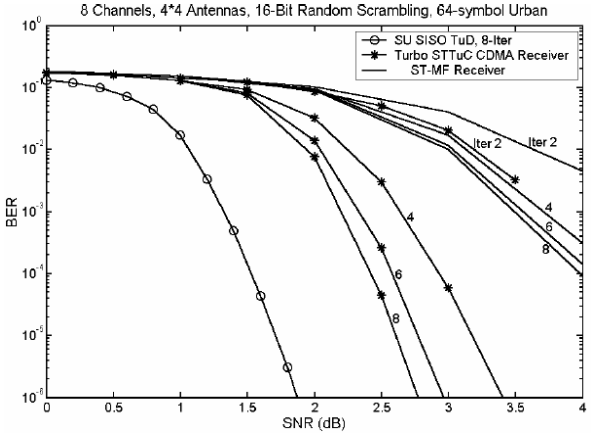


Fig. 5. Performance of Turbo Receivers With BS Specific Scrambling

BER of 10^{-6} at an SNR of 2.75dB. On the contrary, the ST-MF receiver can only approaches a BER of 10^{-4} at an SNR of

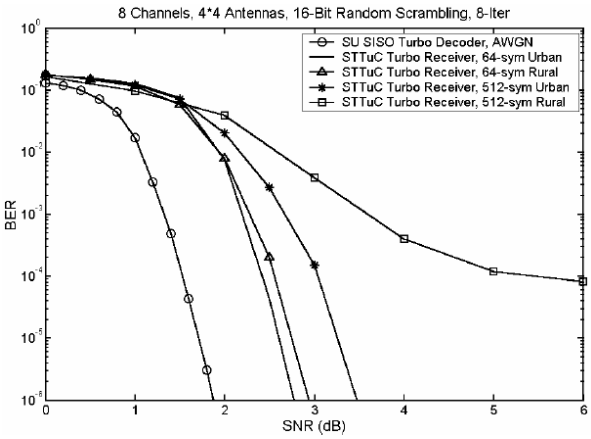


Fig. 6. Performance of Turbo Receiver In different channel model

4.0dB. This suggests that the Turbo receiver can achieve a substantial performance gain over the non-iterative receiver in combating the interference and resisting the multipath distortion in the downlink STTuC CDMA system.

Figure 6 represents the BER performance of the STTuC CDMA receiver in different multipath environ-ments. The 64-symbol rural channel results in only 0.2dB performance degradation over the urban channel with the same fading rate. The performance in the 512-symbol urban channel (the whole STTuC frame is codes for the antennas. This will greatly reduce the risk 64-symbol case. When slow fading occurs together with a small number of paths, the Turbo receiver encounters large performance degradation. In the 512-bit rural channel, the BER performance improves slowly with increasing SNR and can hardly approach a BER of 10^{-6}.

However, this unacceptable degradation can be avoided with the ASS technique. The requirement of more OVSF codes will reduce the system capacity dramatically. However, the scrambling codes are not a very scarce resource in downlink CDMA. Hence it is possible and reasonable to

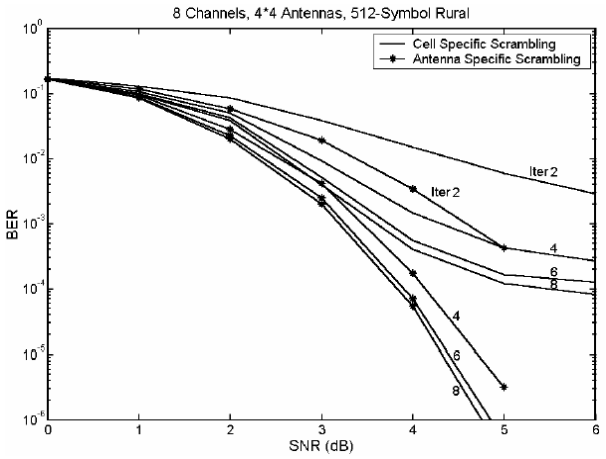


Fig. 7. BS Specific Scrambling vs Antenna Specific Scrambling

allocate different scrambling of high correlation between sub-streams of an STTuC frame. Figure. 7. shows the performance comparison between BSS and ASS systems. It suggests that the Turbo receiver with ASS significantly outperforms that with BSS and can achieve a BER of 10^{-6} at quite satisfactory SNR values.

3.2 BER Performance against Number of Data Channels

In this section the receiver performance with different number of code-division channels is examined. This performance will determine the downlink capacity achieveable in the system.

The system can achieve a larger capacity if more traffic channels are transmitted. However, the more channels transmitted by the BS, the greater the MAI and IAI introduced. A very interference-resistant receiver can minimize the degradation due to the increase in the number of channels (interference), and potentially achieve a large capacity. Here the basic CDMA channel allocation scheme is first considered, in which one user only occupies one traffic channel. If the chip rate is assumed to be 5Mcps, N=16 and n_T =4, the overall symbol rate of one channel is 1.25 Mbps. Then

the system throughput will be1.25K Mbps. For simplicity, it is assumed that all channels are used to transmit data.

Figure 8 represents the BER performance of the ST-MF and the Turbo STTuC CDMA receiver with different numbers of traffic channels in a 64-symbol block fading urban environment.

The simulation results suggest that the Turbo receiver achieves a substantial performance gain over the ST-MF receiver, especially in a heavily loaded system. The performance of the ST-MF receiver substantially degrades with increasing number of channels, whereas the Turbo receiver only suffers a relatively small degradation when the number of users increases. Especially, with the Turbo reception, a 6-channel system has almost the same SNR requirement as a 2-channel system for a BER of 10^{-6}. This is an essential feature for the system to achieve a high capacity.

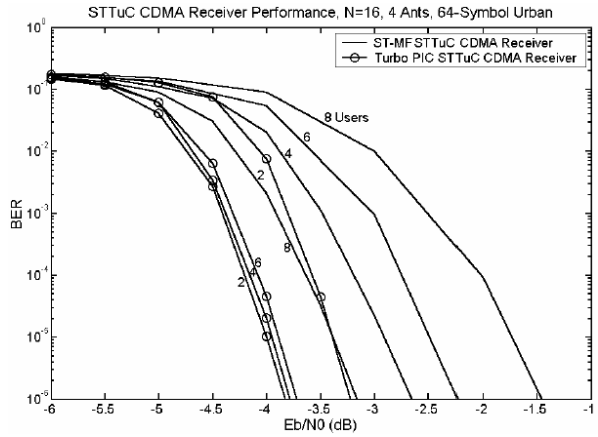


Figure. 8. Performance of STTuC CDMA Receivers against Number of Channels

4. CONCLUSION

In this paper, the performance of the proposed Turbo receiver is evaluated in the downlink STTuC CDMA system. First the Turbo receiver is conceived to adapt to the use of long scrambling codes. Both the space-time RAKE receiver and PIC are performed with a "sliding window" in which the correlation information only of a few adjacent symbols is kept simultaneously. The simulation suggests that the Turbo receiver can achieve satisfactory performance in most circumstances except for the slow fading rural channel. In all circumstances, the Turbo receiver still obviously outperforms the non-iterative receiver. This is perhaps infeasible in uplink because the scrambling codes are a scarce resource. However in the downlink, there are usually sufficient codes for "antenna-specific scrambling".

5.REFERENCES

[1] P. Castoldi, "Multiuser Detection in CDMA Mobile Terminals", ARTECH HOUSE, INC, Norwood, MA, 2002.

[2] R. Wichman, A. Hottinen, "Multiuser detection for downlink CDMA communications in multipath fading channels", in IEEE Vehicular Technology Conference, May 1997, pp. 572 – 576.

[3] B.H. Lee, J.C. Jeon, et al, "Recursive interference canceller using slope optimized soft limiters for DS-CDMA downlink channels in frequency selective fading", in IEEE Vehicular Technology Conference (VTC 2003-spring), Jeju, Korea, April 2003, pp. 2152 – 2156.

[4] A. Duel-Hallen, "Equalizers for multiple input multiple output channels and pam systems with cyclostationary input sequences", IEEE Journal on Selected Areas on Communications, April 1992, pp. 630–639.

[5] M. Gonzalez-Lopez, J. Miguez, et al, "Decision feedback turbo equalization for space-time coded systems", in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03), April 2003.

[6] E. Biglieri, G. Taricco, et al, "Decoding space-time codes with BLAST architectures", IEEE Transactions on Signal Processing, vol. 50, , Oct 2002, pp. 2547 – 2552.

[7] H. Claussen, H. R. Karimi, et al, "High-performance MIMO receivers based on multi-stage partial parallel interference cancellation", in IEEE Vehicular Technology Conference (VTC 2003-fall), Oct 2003, pp. 414 – 418.

[8] A. Burg, M. Rupp, et al, "Low complexity frequency domain equalization ofMIMO channels with applications to MIMOCDMA systems", in IEEE Vehicular Technology Conference (VTC 2003-fall), Oct 2003, pp. 468 – 472.

[9] F. Petre, D. Barbera, et al, "Combined space-time chip equalization and parallel interference cancellation for DS-CDMA downlink withspatial multiplexing", in IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Sept 2002, pp. 1117 – 1121.

[10] T. Muharemovic, E. N. Onggosanusi, et al, "Hybrid linear-iterative detection algorithms for MIMO CDMA systems in multipath channels", in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 02), May 2002, pp. 2621–2624.

[11] C. Berrou, A. Glavieux, et al, Near Shannon limit error-correcting coding and decoding: Turbo-codes (1). ICC'93. 1993, pp.l064-1074.

[12] T. Abe, T. Matsumoto, Space-time turbo equalization in frequency-selective MIMO channels, IEEE Transactions on Vehicular Technology, vol. 52, May 2003, pp. 469–475.

A Model-driven Approach of Web Services Development

Jia Li, Heming Zhang
National CIMS Engineering Research Center, Tsinghua University
Beijing 100084, P. R. China
Email: jial04@mails.thu.edu.cn
Email: hmz@tsinghua.edu.cn

ABSTRACT

The application of the Model Driven Architecture (MDA) to Web Services development has received considerable attention. This paper presents a model-driven Web Services development approach. A MOF compliant meta-model for WSDL is generated, mapped into UML meta-model and integrated into composite web services. The approach can easily be implemented using existing XML Schema tools and UML CASE tools and can facilitate the development of composite web services.

Keywords: Web Services, MDA, Meta-model, WSDL, UML.

1. INTRODUCTION

Web Services are web-based enterprise applications that use XML based standards and transport protocols to communicate with each other in a platform and language independent manner, which are usually described in the Web Services Description Language (WSDL) [1].

There is a growing interest in reusing basic web services in composite web services. In this context, creating compliant meta-models supporting modeling Web Services is an essential factor. Applying Model Driven Architecture (MDA) [4,5,7] to Web Services development has recently received considerable attention [2,3,4]. Extensive work has been done on the model transformation for Web Services to various implementation platforms. In the MDA, all meta-models are based on a unique meta-model called Meta Object Facility (MOF) [5]. Model transformation can be carried out via defining transformation rules between two MOF compliant meta-models [2,3,11].

This paper adopts the MDA strategy for composite web services. We propose a meta-modeling method, by creating MOF compliant meta-model for the WSDL, then mapping WSDL meta-model into UML meta-model, and finally refining the UML models to support modeling for Web Services and can be used as an integration platform for modeling executable composite web services. The whole meta-model generation and transformation process is tool-based and supported by *hyper*Model [6] and UML tools.

This paper is organized as follows. Section 2 is a brief overview of the tools for Web Services development. Section 3 presents the core approach and sketches the implementation. Section 4 is a framework based on the meta-model approach. Finally, the main contributions and future work are concluded in section 5 and 6.

2. RELATED WORK

There are mainly three sets of tools supporting Web Services development, considering their capabilities in relation with the MDA principles [7].

Tools focused on generating the Web Services implementation comprise the first group. The web services description is generated at the implementation phase. There are several tools (IBM WebSphere Studio, Borland Web Service, Microsoft.Net, etc.) capable of generating the WSDL description starting from Java Beans or C++ source code. These tools can also generate the Web Services implementation from WSDL description. This set of resources does not include any modeling capability.

The second group includes tools that support generation of web services description or its implementation (Codagen Architect [12], etc.) by using UML models imported from other tools. These tools do not support the complete development process for lack of modeling ability.

The last set of tools comprises those supporting Web Services Modeling, which use their own notations (XML Spy [13], BEA WebLogic, etc.) or use UML Profile modeling (Oracle Jdeveloper, etc.), which are incapable of collecting complete semantics to define the Web Services.

All these tools deploy the Web Services in a specific technology. So the whole development process depends on the technology in which the Web Services are finally deployed. In order to separate the technology platform from business domain concepts, strong modeling capability should be introduced to support model-driven Web Services development. In this paper, we propose a MDA-based modeling approach of generating the meta-model for WSDL, which can facilitate the development of composite web services.

3. MODELING WEB SERVICES BASED ON THE MDA

Kurtev and K. van den Berg identify four MDA Model Transformation scenarios [8]. Three of the scenarios are studied in [8] to make direct use of the definition of the transformation rules between meta-models. In the context of Web Services, model transformations can be carried out via defining transformation rules between two MOF compliant meta-models. Fig. 1. shows the use of transformation rules for model transformation [2].

In this section, first a brief introduction on concepts involved in the model transformation for XML based languages is presented, and then the tool-based approach of generating meta-model for WSDL is proposed to describe how the MDA meta-model mechanism is used in the

development of Web Services.

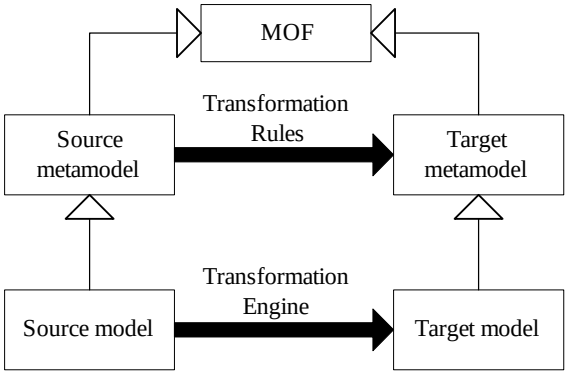


Fig. 1. Using transformation rules in the MDA

3.1 XML, XMI and XSD

The Extensible Markup Language (XML) [5] is a cross-platform, text based W3C standard for interchanging, structuring and representing data. XML can be used as meta-language, allowing the generation of XML-based languages, which may be specialized in Web Services with WSDL [1], UDDI [9], BPEL4WS [9], Ontology with RDF and XML Metadata Interchange (XMI) [5].

The XML Schema Definition (XSD) [14] is an XML language for describing the vocabulary and grammar of XML documents, and to impose various constraints on their context. XSD provides a validating mechanism to check the XML document.

XMI provides a common interchange format between UML and XSD for interchanging models and metadata.

3.2 Transformation between XML and UML

The XMI is designed to facilitate the interchange of data and metadata expressed via the MOF. Consequently, the XMI specification defines a number of mapping rules that specify how to generate XML Document Type Definition (DTD) and XML Schema from UML class diagrams. The XMI also specifies methods of producing MOF models from such input formats. The automatically generated DTDs and XML Schemas are based on the MOF defined rules and allow the MOF-based models to be validated and interchanged among different tools without controversies. This makes XMI a necessary intermediate medium standing between MOF models and XMI representations. Therefore, any transformation from the XML to MOF/UML needs to be based or extend XMI. The transformation from an XML Schema or DTD to an XMI format can be performed using the Extensible Style Sheet Language (XSLT).

One of the key features of the XMI is that it provides parameterized mapping, i.e., by choosing different mapping parameters, it is possible to define different mappings from a UML model to its schema representation.

3.3 A tool-based approach to generating meta-model for WSDL

A meta-model for WSDL defines the model elements of the Web Service Language, specifies the semantics of language and relationship between various model elements. As a result, the modeler often starts from understanding the language description by studying its specification and creating a conceptual model involving the entities of the language and their relationship. Currently, there is no systematic way of creating such conceptual models. Fig. 2. shows the outline of the approach to generating meta-model for WSDL.

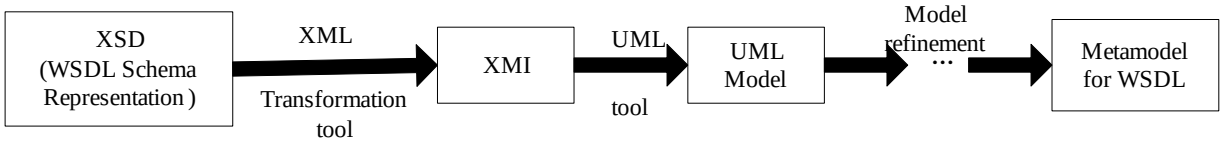


Fig. 2. Approach of generating meta-model for WSDL

To create a MOF compliant meta-model, the approach starts from the XSD schema representing the language. Then, as shown in Fig. 2., the transformation from an XSD Schema to a UML model is a fully automated process, which is supported by UML CASE tools. The UML model presents a clear, high-level view of the involving concepts. At this point, the modeler begins refining the UML model to include WSDL semantic statement.

In this paper, the approach is implemented under the support of *hyper*Model and UML tools. *hyper*Model [6] is an XML schema design and integration tool, offering various UML modeling capabilities. *hyper*Model is a free plug-in to Eclipse Workbench[2,15] allowing the transformation of XML vocabularies and schema into XMI 1.0 format. To implement this method, a XSD document of WSDL is created in *hyper*Model, and an XMI document from the XSD is automated generated. It is possible to view the XMI model as a UML class diagram in *hyper*Model/Eclipse.

In order to have greater flexibility in editing and refining of the model, the XMI document is imported into a separate UML tool, for example Poseidon UML[10]. Through generating the meta-model for WSDL under the support of *hyper*Model and UML tools, the development process of Web Services can be more flexible and convenient. Moreover, meta-model, once created, can be reused in composite web services and can easily generate source code specified for different technology platforms.

3.4 Case Study: A meta-model for WSDL

The Web Service Description Language (WSDL) describes the syntax and semantics necessary to call up the services, which are composed of:

- *Type* — defines abstract data types;
- *Operation*— defines the operation interface the service provides;
- *Message* — describes the message structure;
- *PortType* — presents web service interface;

- *Service* —describes who provides the service;
- *Binding* — the way the service is accessed.

The architecture of WSDL elements can be described as in Fig. 3.. Fig. 4. shows a fragment of the WSDL meta-model created by *hyper*Model.

The XML provides an extensive mechanism for documenting and extensibility features. All the elements in the meta-model have equivalent in MOF by excluding elements specific to XML and stereotypes, which are created from an XML tag representing XSD, attributes.

Input and output are modeled as meta-model attribute ends, which are of type parameters (Param). This corresponds to the following line in the XSD document for the WSDL:

<xs: element name=”input” type=”wsdl: tParam” />

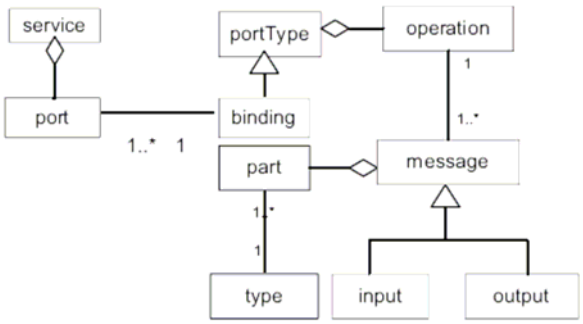


Fig. 3. The architecture of WSDL

Based on this meta-model, a Web Services modeling environment can be constructed, and a generator based on this meta-model can be created to interpret Web Services models to generate WSDL descriptions.

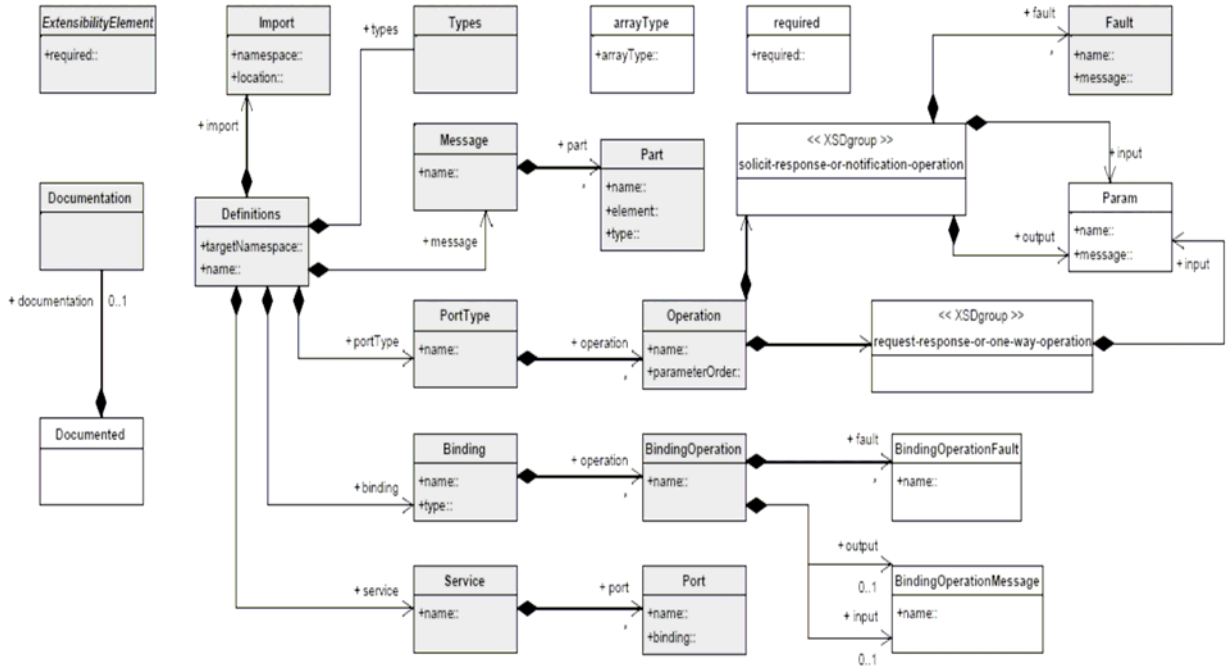


Fig. 4. The WSDL meta-model (fragment)

3.5 Mapping from WSDL to UML

As shown in Figure 1, the transformation between UML and Web service descriptions requires the support of two MOF compliant meta-models. In the transformation process, rules map from UML meta-model into WSDL meta-model. Figure 5 shows the mapping from UML meta-model to WSDL meta-model.

According to this figure, we have the following mappings according to XMI mapping rules:

- The UML Interface is mapped to WSDL PortType, Binding and Service;
- The UML Class and DataTypes are mapped to WSDL Types;
- The UML Operation is mapped to WSDL Message, OperationType and BindingOperationType.

Figure 5 shows the existence of one-to-many mappings, a one-to-many mapping has one element of the meta-model source and many elements of the meta-model target. This is also a consequence of the semantic difference between the source and target meta-model. Sometimes, the semantics of one element of a meta-model is represented using many elements of another meta-model. In the case of the mapping from UML to WSDL meta-model, the UML Interface is mapped into WSDL PortType, Binding and Service. As shown in this section, one benefit of MDA is its capability to allow the generation of several implementations, on different platforms, from a business model. In other words, the business system characteristics are only acquired and modeled once.

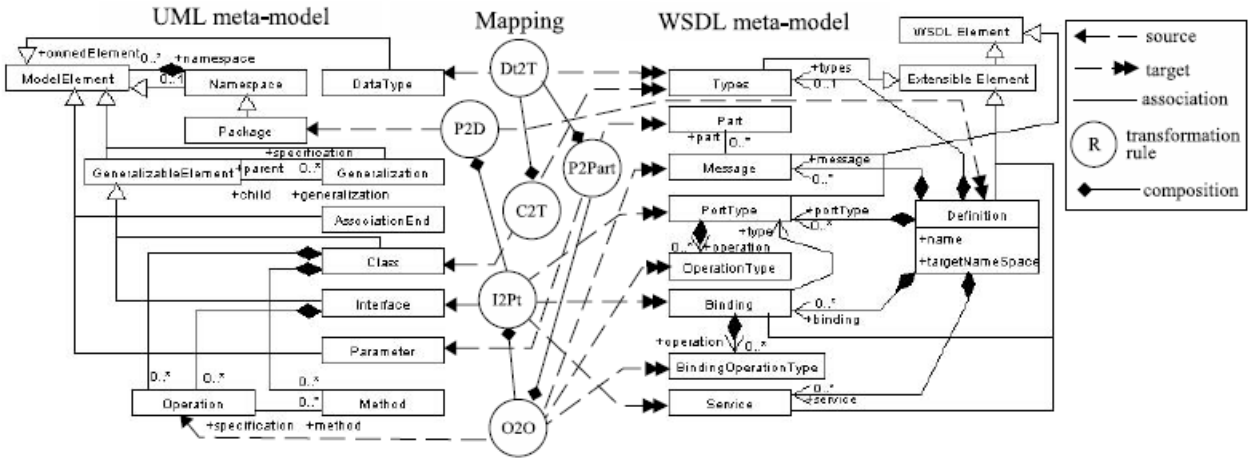


Fig. 5. Mapping from UML to WSDL meta-model

4. A FRAMEWORK

Based on the approach of generating meta-model for WSDL described above, a framework for agile development of composite Web Services based on the MDA approach is presented. Figure 6 shows a UML activity diagram indicating the process of model-driven Web Services development.

In the first step (Discover Web Services) the developer uses a web-browser, a registry client to search and discover candidate web services that may be used in the composite service. The output of this activity is a list of web service descriptions, represented as WSDL documents. Then the necessary web service descriptions are imported into UML by a reverse engineering transformation. The output of this step is one or more UML models of the discovered web services. The developer then uses a UML tool to review and integrate the imported models to form a model of a composite web service. This activity consists of two sub-activities: service modeling and workflow modeling, which focuses on the interface of the service and its internal processes respectively. The output is a new UML model representing the composite web service with its service and workflow. This model can be used to generate the WSDL description of the composite service and to generate the process description of the composite service that can be used to implement the service. Finally the service is published in an appropriate registry, making it available for use.

The process of the model-driven Web Services development is carried out via the automatic generation of WSDL meta-model. The approach of modeling Web Services based on the MOF compliant WSDL meta-model can be easily mapped to UML meta-model, and support pure UML modeling at platform- independent levels, and can enhance the understanding and efficiency of service modeling.

Moreover, the same model may be used as a basis for transformation to more than one target platform (Java, IDL, etc.), to later versions of the same platform. Besides, the high-level, graphical models are easier to understand, as they do not have all the technical details of the target platform.

5. CONCLUSION

This paper presents an approach of Web Services development using the MDA method. A meta-model for WSDL is semi-automated generated, which is specified via XSD. Then an XMI document is created from the XSD specification and imported as a class diagram in a UML tool, provides a high level view of the concepts involved in the language and their relationship. The model-driven approach of Web Services development shows how the meta-model for WSDL can be effectively used for the UML modeling of Web Services. Finally, an MDA tool that supports the complete development process for composite web services is proposed. The proposed approach can facilitate the development of composite web services.

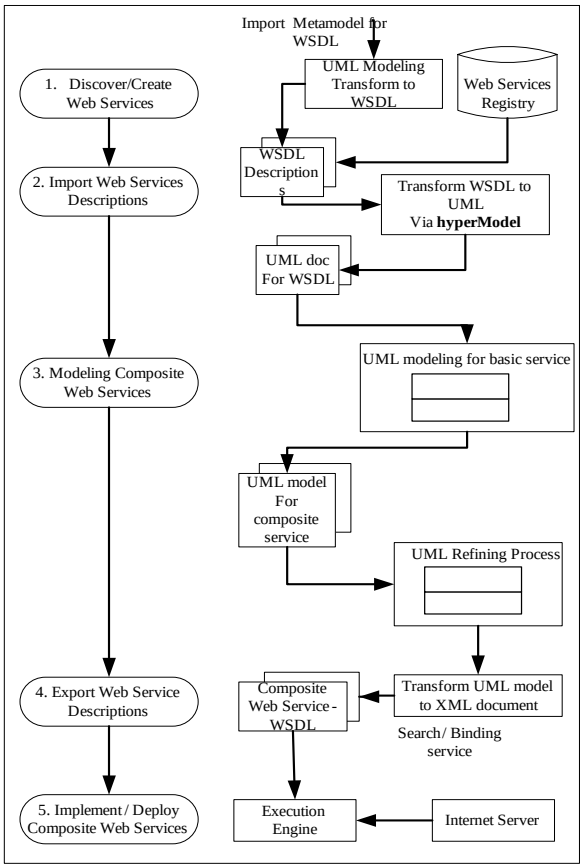


Fig. 6. A Framework of the model-driven Web Services development process

6. REFERENCES

[1] W3C, Web Services Description Language (WSDL) Version 2.0, W3C Working Draft, November 2003, http://www.w3.org.

[2] J. Bezivin, S. Hammoudi, et al, An Experiment in Mapping Web Services to Implementation Platforms, Atlas Group, INRIA and LINA University of Nantes, Research Report, March 2004.

[3] D. S. Frankel, Model Driven Architecture, and Model Driven Architecture, Applying MDA to Enterprise Computing, OMG Press, and January 2003.

[4] D. Lopes, S. Hammoudi, Web Services in the Context of MDA, University of Nantes, and France, 2003.

[5] Object Management Group, http://www.omg.org.

[6] http://www.xmlmodeling.com.

[7] A. Kleppe, J. Warmer, et al, MDA Explained. The Model Driven Architecture, Practice and Promise, Addison-Wesley, April 2003.

[8] Kurtev, K.V.D. Berg, "Unifying Approach for Model Transformations in the MOF Metamodeling Architecture", Proc. of the 1st European MDA Workshop, 2004.

[9] OASIS, Universal Description Discovery & Integration (UDDI), Version 3, Business Process Execution Language for Web Services (BPEL4WS), http://www.oasis-open.org.

[10] Poseidon for UML, http://www.gentleware.com.

[11] J. Bezivin, S. Hammoudi, et al, An Experiment in Mapping Web Services to Implementation Platforms, Atlas Group, INRIA and LINA University of Nantes, Research Report, March 2004.

[12] Codagen, http://www.codagen.com.

[13] XML Spy 5, http://www.xmlspy.com.

[14] W3C, XML Schema Primer.

[15] Eclipse project, http://www.eclipse.org.

VIPS-based Web Cleaning Algorithm

Hongxia Shi, Xun Wang, Yun Pan
College of Computer & Information Engineering, Zhejiang Gongshang University,
Hangzhou, Zhejiang 310035, P.R.China
Email: shihx@mail.zjgsu.edu.cn

ABSTRACT

On the basis of the features and drawbacks of traditional web page cleaning methods being analyzed, this paper presents a new VIPS-based web page cleaning algorithm. This algorithm first categorizes all the web pages with VIPS method, then computes the appearance frequency of each web page block in the website by judging the similarities among page blocks, and finally computes the importance of each web page block according to the appearance frequency, text quantity, position and the number of links, while the web page with the importance below the threshold is considered as "noise content". The algorithm proposed is applied to KNN classification algorithm, and the experimental result proves that this algorithm could effectively improve the accuracy rate of web page classification.

Keywords: Page Cleaning, VIPS, Page Classification.

1. INTRODUCTION

On browsing web pages, we usually find that not all contents meet our demands. For a web page, there are pictures, news and etc. But in general, web pages normally are mainly composed by two typical types of contents: one type is the topical information of the web page, so called topical content; while the other type is the navigation bar, advertisements, copyright information, surveys and etc, so called "noises".

Usually the noises are situated around the topical information, and may sometimes even amid the information. Normally the noises are unrelated to the topical content, and the pages linked to are unrelated to the topical content too. The noises in the pages have caused problems to the web page content-based application systems. So, how to rapidly and accurately distinguish and clear the noises in the web pages is a key technology to improve the accuracy of the processing outcomes of web applications, which is called web page cleaning. Firstly, after page cleaning, there would be no more interferences from the noise contents, thus web applications could process the topical contents in the web pages and improve the processing accuracy. Secondly, web page cleaning could significantly reduce the complexity of the tag structures within the pages and decrease the page sizes, saving much overhead of time and space of later processing. Therefore, web page cleaning has become an important preprocessing link of web information system.

2. COMMON WEB CLEANING ALGORITHM

In web page cleaning, in order to recognize the content of web pages as noises or topical contents, the premier task is to segment the whole page into various blocks. Now, there are many web page segmentation methods, and more popular ones are DOM-based segmentation method [1], location-based segmentation method [2] and VIPS-based segmentation method [3]. After segmenting the page into different blocks, the next task would be to measure the importance of every block. The methods of measuring block value could be categorized into two types. One type is for measuring different pages of the same web site, while another is for measuring an individual page. For the first type, most methods are based on the following considerations: for a web site, noises usually have common contents and styles [4]. In document [5], the researcher defined the common parts of many different web pages as templates. When the web pages are segmented into different blocks according to some rules, then the web page blocks, similar to the predefined templates, that occur over and over are recognized and counted. The web page blocks appear many times are naturally noises. Document [6] first segments a web page into several content blocks according to the "TABLE" tag, then calculates the entropy of every feature terms according to the occurrences of feature terms within the page set, and on this basis calculates the entropy for the page block, and finally uses the entropy to determine whether the block is or is not noise. In document [7], the author defined the style tree (ST) to describe the layout and content of web page. A style tree for a web site is called SST tree. The importance of every node within the SST tree could be calculated to clean the web page. The experimental results show that this de-noises method could significantly improve web classification and clustering. Document [11] describes a typical method focusing on individual web page, while this DOM-based extraction algorithm could clean advertisements and some links.

3. VIPS-BASED PAGE CLEANING ALGORITHM

Introduction of VIPS Algorithm

VIPS-based web page segmentation method is mainly adopted to extract the content structure of a certain web page. VIPS could fully exploit the layouts of web pages, such as fonts, colors and sizes, to extract suitable nodes from the DOM tree, and then locate the segmentation symbols for these nodes. Segmentation symbols are the horizontal and vertical lines in the page that do not cross any nodes. From these segmentation symbols, the content structure could be constructed. VIPS could efficiently segment all relevant contents into the same block, since it segments the page on the semantic level. Every content block within VIPS is represented as a node, while the nodes of upper levels have sub-nodes. Following is a detailed explanation of VIPS.

For any web page, there exists a corresponding DOM tree. In VIPS, the leaf nodes that are no longer dividable of DOM tree is considered as "fundamental objects". In the VIPS-based content structure, the nodes are called "layout blocks", namely a "fundamental objects" or a group of "fundamental objects". The nodes within content structure are not obliged to have corresponding nodes within the

DOM tree.

The basic model of VIPS-based content structure of a web page could be described as follows. A web page Ω could be defined as a tri-tuple $\Omega = (O, \Phi, \delta)$, where $O = (\Omega_1, \Omega_2, \ldots, \Omega_N)$ is the finite set of sub-pages (a web page could be segmented into many blocks, while every block could be regarded as a sub-page of the web page). $\Phi = \{\varphi^1, \varphi^2, \ldots, \varphi^T\}$ is the finite set of segmentation symbols, including horizontal and vertical segmentation symbols. Every segmentation symbol has a weight, representing its visibility in the browser, while the symbols of same Φ have identical weights. δ is the relation of web page blocks within O, and could be denoted as:

$$\delta = O \times O \to \Phi \cup \{NULL\}$$

Suppose that Ω_i and Ω_j are adjacent pages within O, $\delta(\Omega_i, \Omega_j) \neq \textbf{\textit{NULL}}$ means that Ω_i and Ω_j are segmented by the symbol $\delta(\Omega_i, \Omega_j)$. Since Ω_i is the sub-page of original web page Ω, Ω_i and Ω have similar models of content structure. Suppose Ω_s^t is the *t*th sub-page of level S, then:

$$\Omega_s^t = (O_s^t, \Phi_s^t, \delta_s^t)$$

$$O_s^t = \{\Omega_{st}^1, \Omega_{st}^2, \ldots, \Omega_{st}^N\}$$

$$\Phi_s^t = \{\varphi_{st}^1, \varphi_{st}^2, \ldots, \varphi_{st}^T\}$$

$$\delta_s^t = O_s^t \times O_s^t \to \Phi_s^t \cup \{NULL\}$$

where N_{st} is the number of sub-pages of O_s^t and T_{st} is the number of segmentation symbols of O_s^t.

For every page block, a DoC (Degree of Coherence) is defined to describe the coherence degree of the content within the page block. Suppose part of page block 1 contains information about car and part of it contains information about planes; while the content of page block 2 contains only information about car, then the Doc of block 2 would be larger than that of block 1. DoC value has three properties: (1) domain [0,1], (2) the larger the DoC value, the higher coherence of contents of page block, (3) in the content block tree, the DoC of sub-node should be no less than that of its parent-node. Predefining a Doc to control the depth of the content block tree, thus if the DoC of a certain content block surpasses the predefined DoC, then stop the segmentation. The smaller the predefined Doc, the coarser the constructed page content structure would be.

The VIPS-based page segmentation process is as follows. Firstly, construct a web page DOM tree and extract the visual information. Secondly, extract the visual web page blocks for the current level from the DOM tree, starting form the root node. Check whether every node within the DOM tree could form an independent block, and if not, and then apply the same processing method on its sub-nodes. After the extraction of all blocks of the current level, these blocks are stored in a database. Recognize the VIPS-based segmentation symbols from these blocks and set a weight according to the features of their adjacent web page blocks within the same level. After having constructed the content structure of the level, check the DoC values of the new VIPS-based page blocks to see whether they have passed the predefined DoC. If not, then continue segmenting web page block. Once all the web page blocks are processed, output the final VIPS-based content structure.

Description of Cleaning Algorithm

Input: all or part of the web pages of a web site.

Output: non-noise contents of the given web site.

Step 1: use VIPS algorithm to extract the content structure of all or part of the web pages of a web site. The leaf nodes are web page blocks and are stored in a database. Every web page block has some properties, mainly including: text, length of text, Doc, the number of links, position of web page blocks and etc.

Step 2: calculate the weight of the web page blocks. Following matters should be taken into considerations: (1) noises normally appear repeatedly in a web site, (2) the text of noises is relatively shorter when compared with the whole web page, (3) noises are normally situated at the top, bottom, left or right of the web page, (4) the topical contents contain less links. From these matters, we could conclude the following rules: (1) the more occurrences of a certain web page block within a web site, the less important the block is, (2) the more text contained in a web page block, the more important it is, (3) the more closer to the center of the web page, the more important the block is, (4) the smaller the ratio of the text linked by a web page block, the more important it is.

The procedure to calculate the weight is follows:

(1) calculate the occurrences of web page blocks

To calculate the occurrences of web page blocks, it is necessary to determine whether these blocks are identical. Vector space model (VSM) is adopted to calculate the text similarity between two web page blocks. If the similarity exceeds a certain threshold, then these two blocks are considered to be identical. The detailed algorithm is as follows:

① extraction of feature terms. Words with huge information are extracted as feature terms. Document [8] summaries the current available feature term extraction methods, such as IG, χ^2 statistics (CHI)、information entropy and etc.

② use eigenvector to represent the text of web page block. The eigenvector web page block D could be represented as $D(W_1, W_2, \ldots, W_N)$, and the formula to calculate W_i is:

$$w(t, \vec{d}) = \frac{tf(t, \vec{d}) \times \log(\frac{N}{n_t})}{\sqrt{\sum_{t \in \vec{d}} [f(t, \vec{d}) \times \log(\frac{N}{n_t})]^2}}$$

where $W(t, \vec{d})$ is the weight of feature term *t* within web page block $\vec{d}$, $tf(t, \vec{d})$ is the word frequency of word *t* within web page block $\vec{d}$, *N* is the total number of web page blocks, n_t is the number of page blocks with the occurring of word *t*, and the denominator is the normalization factor.

③ calculate the text similarity of two web page blocks. The formula is:

$$\mathrm{Sim}(d_i, d_j) = \frac{\sum_{k=1}^{M} W_{ik} \times W_{jk}}{\sqrt{(\sum_{k=1}^{M} W_{ik}^2)(\sum_{k=1}^{M} W_{jk}^2)}}$$

where d_i is the eigenvector of the *i*th web page block in

the database, M is the dimension of the eigenvector, and W_k is the *k*th dimension of the vector.

④ if the text similarity between two web page blocks exceed a certain value C, then the contents of these two web page blocks are considered to be identical

If the same web page block occurs repeatedly in the same web page, then count it just once.

(2) calculate the relative text length OR_l of web page block

$OR_l = \frac{O_l}{P_l}$, where O_l is the text length of web page block, and P_l is the text length of the entire web page.

(3) calculate the relative position (OR_x, OR_y) of the web page block

From Step 1, we could obtain the starting position, width and height, and define them respectively as (O_x, O_y), O_w, O_h. Define the center of the web page block as (OX, OY), where $OX = O_x + O_w/2$, $OY = O_y + O_h/2$

Suppose the width and height of the original web page are P_w and P_h, then:

$$OR_x = OX / P_w = (O_x + O_w/2)/P_w$$
$$OR_y = OY / P_h = (O_y + O_h/2)/P_h$$

(4) calculate the relative length of the linked text OLR_l within the web page block

$OLR_l = OL_l / O_l$, where OL_l is the length of the linked text of the web page block.

(5) calculate the weight of web page blocks

For the above-mentioned four rules, the weight of every rule is not the same. For example, the first rule is more capable in recognizing the web page noises than the fourth rule. So, set different weights for these four rules to indicate the importance of that rule on noise recognition.

Suppose the weights of these four rules are W_1, W_2, W_3, W_4, and define the domain as [0,1].

Define n the occurrences of web page block B, N the number of web pages of the web site, then the formula for calculating weight is:

$$W(B) = [\, w_1(1 - \frac{n}{N}) + w_2 \cdot OR_l + w_3(1 - \frac{\sqrt{(OR_x - 0.5)^2 + (OR_y - 0.5)^2}}{\sqrt{2}/2}) + w_4(1 - OLR_l)\,] / (w_1 + w_2 + w_3 + w_4)$$

where $\sqrt{(OR_x - 0.5)^2 + (OR_y - 0.5)^2}$ represents the distance between (OR_x, OR_y) and the relative center(1/2, 1/2).

Step 4: given a threshold, when the weight of a certain page block exceeds that threshold, then the block would be regarded as useful information, otherwise, the block would be treated as noises and be cleaned.

4. WEB CLEANING EXAMPLE

Download the web pages containing sports news from http://sports.sohu.com. Because of the huge amounts of web pages, 40 web pages containing topical contents are randomly selected for testing. Firstly, use VIPS algorithm to segment these pages. Fig.1 and Fig.2 are the segmentation for a specific web page, where the web page block A in Fig.1 is a noise block and web page block B in Fig.2 is a topical block. Above-mentioned cleaning algorithm is used to calculate the weight of these two blocks, in order to verify the efficiency of this algorithm.

Fig.1 and Fig.2 are the segmentation for a specific web page, where P_l=2811, P_w=764, P_h=2948. The values of w_1, w_2, w_3, w_4 are: 1, 0.8, 0.2, 0.2.

Table 1 The properties of web page block A

n	40	O_l	102
OL_l	81	O_x	610
O_y	30	O_w	164
O_h	2841		

The relative position of A is:

$$OR_x = OX / P_w = (O_x + O_w/2)/P_w = 0.91$$
$$OR_y = OY / P_h = (O_y + O_h/2)/P_h = 0.49$$

Table 2 The properties of web page block B

n	1	O_l	1558
OL_l	101	O_x	10
O_y	259	O_w	600
O_h	1255		

The relative position of B is:

$$OR_x = OX / P_w = (O_x + O_w/2)/P_w = 0.41$$
$$OR_y = OY / P_h = (O_y + O_h/2)/P_h = 0.3$$

With the formula for calculating weights and above data, the weights of web page blocks A and B are: Weight(A)=0.07 and Weight(B)=0.79. So this algorithm could effectively distinguish topical content from noise block.

Fig.1 Noise block A

Fig.2 Topical block B

5. Experimental results

This cleaning algorithm is applied to the preprocessing of KNN classification algorithm, to test the influence of the proposed VIPS-based cleaning algorithm on web page classification. The detailed algorithm procedures are as follows:

Step 1: select the text of "text classification material database" of Chinese natural language processing on open platform www.nlp.org.cn as the training text for the KNN algorithm (a typical text classification algorithm). The texts are classified into 10 groups, which are sports, military, economy, education, environment, transportation, medicine, arts, political, computer.

Step 2: download 50 pages for each topic, such as computer, military, economy and education, from www.163.com

Step 3: apply KNN algorithm to classify these web pages

Step 4: apply VIPS-based cleaning algorithm to clean these web pages

Step 5: apply once more the KNN algorithm to classify theses cleaned web pages

Directly the KNN classification algorithm on the testing web pages and the experimental results are shown in Table 3.

Table 3 The classification results before web cleaning

	Sports	Military	Economy	Education
Testing pages	50	50	50	50
Classified as group	40	20	90	44
Correct classification	39	19	47	43
Recall rate	78.0 %	38.0 %	94%	88 %
Precision rate	97.5 %	95%	52.2%	97.7 %

The average recall rate is 74.5% and the average precision rate is 85.6%, where six web pages are classified into wrong groups.

If the testing pages are preprocessed by the VIPS-based cleaning algorithm, and KNN classification algorithm is adopted to classify these pages, then the classification results are shown in Table 4.

Table 4 The classification results after web cleaning

	Sports	Military	Economy	Education
Testing pages	50	50	50	50
Classified as group	35	34	68	39
Correct classification	35	33	45	38
Recall rate	70.0 %	66.0 %	90%	76 %
Precision rate	100 %	97.1 %	66.2%	97.4 %

The average recall rate is 75.5% and the average precision rate is 90.2%, where twenty-one web pages are classified into wrong groups.

From the experimental results, the VIPS-based cleaning algorithm has not much influence on the recall rate, but could significantly improve the precision rate after the web page cleaning.

6. Conclusions

VIPS-based web page cleaning algorithm could perform the web page cleaning, but with low time efficiency. The algorithm expenses much time on the comparison of the similarities among web page block. Large web sites normally contain plenty of web pages, thus these web pages should be classified into several sets, and then every set could be processed as a single sub-site. For example, every 50 pages of an identical web site could be classified as a single set. When performing cleaning, every time only 50 pages are processed.

7. REFERENCES

[1] Chen, J., Zhou, B., Shi, J., Zhang, H.-J. and Qiu, F.,Function-Based Object Model Towards Website Adaptation[C], in proceedings of WWW'01, Hong Kong, China, May 2001, pp.587-596.

[2] Kovacevic, M., Diligenti, M., Gori, M. and Milutinovic, V.,Recognition of Common Areas in a Web Page Using VisualInformation: a possible application in a page classification[C], in proceedings of ICDM'02, Maebashi City, Japan, December 2002, pp.250.

[3] Yu, S., Cai, D., Wen, J.-R. and Ma, W.-Y., Improving Pseudo-Relevance Feedback in Web Information retrieval Using Web Page Segmentation[C], in proceedings of WWW 2003,Budapest, Hungary, May 2003, pp.11-18.

[4] Yi, L. and Liu, B., Web Page Cleaning for Web Mining through Feature Weighting[C], in proceedings of IJCAI-03, Acapulco, Mexico, August 2003, pp.43-50.

[5] Bar-Yossef, Z. and Rajagopalan, S., Template Detection via Data Mining and its Applications[C], in proceedings of WWW 2002, May 2002, pp.580-591.

[6] Lin, S.-H. and Ho, J.-M., Discovering Informative Content Blocks from Web Documents[C], in the proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (SIGKDD'02), 2002.

[7] Yi, L. and Liu, B., Web Page Cleaning for Web Mining through Feature Weighting[C], in the proceedings of Eighteenth International Joint Conference on Artificial Intelligence (IJCAI-03), Acapulco, Mexico, August 2003.

[8] Yiming Yang and Jan O. Pedersen. A comparative study on feature selection in text categorization [C]. In: Douglas H. Fisher ed. Proc. of the 14th International Conference on Machine Learning ICML97. San Francisco: Morgan Kaufmann Publishers, 1997, pp.412-420.

[9] Gupta, S., Kaiser, G., Neistadt, D. and Grimm, P., DOMbased Content Extraction of HTML Documents[C], in the proceedings of the 12th World Wide Web conference (WWW 2003), Budapest, Hungary, May 2003.

[10] Zhang ZhiGang, Chen Jin, Li XiaoMing, A HTML web page cleaning method[J], Journal of intelligence, 2004. 4(23): pp.387-393.

[11] Gupta, S., Kaiser, G., Neistadt, D. and Grimm, P., DOMbased Content Extraction of HTML Documents[C], in the proceedings of the 12th World Wide Web conference (WWW 2003), Budapest, Hungary, May 2003.

[12] Dong Zhendong. HowNet Knowledge Database [EB/OL]. http://www.keenage.com/, 2006-1.

An Enhanced Genetic Algorithm by Fuzzy Selection and Local Search for Multicast Routing Problem

Yang Tao, Wei Xiong
The Center of Software Technology, School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications,
Chongqing 400065,China
Email: swaol2008@yahoo.com.cn

ABSTRACT

An enhanced Genetic Algorithm by fuzzy selection and local search (GAFL) was proposed in this paper, and it was applied to multicast routing problem. For an NP-hard problem in multicast routing, the fuzzy selection can lead GAFL search into proper direction and find globule optimization as quickly as possible. Moreover combined with its effective local search the search ability of GAFL has been enhanced more, avoiding being plunged into local optimization. Finally simulation results demonstrated GAFL's effectiveness in the routing problem for different scale networks.

Keywords: Genetic Algorithm, fuzzy selection, local search, multicast routing, globule optimization, local optimization.

1. INTRODUCTION

An original description of the Genetic algorithm (GA) was proposed by John Holland [1] in 1975. It was popularly used in all kinds of optimization problems for its feature of robust global search ability. Norio Shimamoto [2] applied GA to a minimum cost Steiner tree problem but without constraints. Then C.W.Tsai [3] proposed an MSGA algorithm to solve the routing problem under delay constraints. Though the efficiency of these methods is better than that of traditional GA more or less, and this showed that the heuristic genetic algorithm would greatly improve the search ability of genetic algorithm. However, when it come to the routing problem, the MSGA still can't be satisfied for meeting some real case.

Therefore, we proposed an enhanced Genetic Algorithm by fuzzy selection and local search (GAFL) and applied it to a real case of multicast routing problem in Section 2. Then in Section 3 we illustrated and analyze the simulate results of GAFL compared with MSGA. In Section 4, we make a brief conclusions and future work.

2. APPLYING GAFL ALGORITHM TO PRACTICAL PROBLEM

We modified a traditional genetic algorithm using both the fuzzy selection and the local search to analyze a multicast routing problem. The GAFL algorithm will find a multicast tree with minimum cost under delay constraint.

2.1 Practical Problem Description

Suspended a network is an undirected and connected graph *G(V, E, w)*. Then n is defined as the number of network nodes, equal to $|V|$, m as the number of network links, equal to $|E|$, s as a source node of the multicast network and *D* as a set of destination nodes, which is the rest of nodes except the source s. A link $e=(v_i, v_j)$ is a connection from a node $v_i \in V$ to another node $v_j \in V$, $i \neq j$. Each link contains two weight values: cost $C(e)$ and delay $D(e)$.

We will find an optimum multicast routing tree, *T(s, D)*. Let $P_T(s, d_i)$ denote a unique path in a tree *T(s, D)* from the source node s to a destination node d_i, $d_i \in V$. Firstly, the total cost from s to all destination nodes is defined as Eq.(1). Secondly, the delay of a path from a source node s to a destination node d_i is defined as Eq.(2). Finally the minimum cost tree under the delay constraint is defined as Eq.(3), in which $\triangle d$ is a delay upper bond of a path.

$$C(T(s,D)) = \sum_{e \in T(s,D)} C(e) \quad (1)$$

$$D(P_T(s,d_i)) = \sum_{e \in P_T(s,d_i)} D(e) \quad (2)$$

$$Min\{C(T(s,D))\} \text{ and } D(P_T(s,d_i)) \le \Delta_d, \forall d_i \in D \quad (3)$$

2.2 GAFL in Multicast Routing Problem

1) Chromosomes Encoding and Initialization Population: For mapping a multicast routing tree into a chromosome, GAFL adopt a encoding method similar to that of [4,5], and then use depth-first search method to generate a link from a source node s to next node k, using randomized selection, until to find a path to a destination node *d(d ∈D)*. When all of paths to the destination nodes in set *D* are found, a multicast routing tree is constructed. Finally an initialization population is just a set of the entire tree found. And according to the encoding method each one of multicast routing trees is correspond to a chromosome.

2) Evaluate Fitness: In GAFL it uses a fitness value as Eq.(4) to estimate a multicast tree represented by a chromosome [6].

$$F(T(s,D)) = \sum_{e \in T(s,d_i)} C(e) \prod \Theta(D(P_T(s,d_i))) - \Delta_d, \forall d_i \in D, \ \Theta(f) = \begin{cases} 1 \ldots f \le 0 \\ 2 \ldots f > 0 \end{cases} \quad (4)$$

3) Fuzzy Selection: Selecting a fixed number of selected chromosomes in each population are very popular. But the major difference of GAFL from the traditional genetic algorithm is fuzzy selection that can adjust the number of chromosomes selection according to the fitness values of chromosomes in each population, moreover can control the size of every population.

Firstly, it is to form fuzzy membership. Let f_i be the average fitness value in a population *i*, and T_j is be a fitness value of a chromosome. Initially, each population has selected a fixed number of chromosomes *N*. We select a few numbers of populations *M* to estimate the average fitness value *K* as

Eq. (5) and to find the minimum and maximum fitness value K_{min} and K_{max} of the fuzzy linguistic function, and then we define four critical points K_1, K_2, K_3 and K_4 as Eq. (6). The points are used to form three trapezoid membership functions.

According to the three shapes of trapezoid membership functions the number of chromosomes that will be selected in the next generation, will be determined as Fig.1. The three types of chromosomes selection linguistic variables are *low(L), middle(M),* and *higher(H)*. The fuzzy function $F_K(x)$, represented by α ($\alpha \in [0,1]$) in Fig. 1, is considered as a linguistic variable with values in the field of x [7].

$$K = \sum_{i=1}^{M} f_i / M,\ f_i = \sum_{j=1}^{N} T_j / N \tag{5}$$

$$\begin{cases} K_1 = K_{min} + (K - K_{min}) \times 2 / 4 \\ K_2 = K_{min} + (K - K_{min}) \times 3 / 4 \\ K_3 = K + (K_{max} - K) \times 1 / 4 \\ K_4 = K + (K_{max} - K) \times 2 / 4 \end{cases} \tag{6}$$

$$K \in [K_{min}, K_{max}]$$

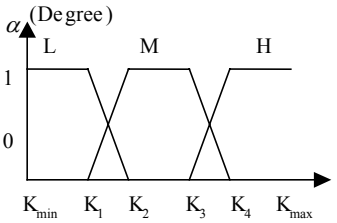


Fig. 1. The membership function plot

Secondly, it is to form fuzzy membership function. Now we can get the membership functions of *low (L), middle (M)*, and *high (H)* are defined as follows.

$$\alpha = F_K^L(x) = \begin{cases} 1 & For\ K_{min} \le x \le K_1 \\ \dfrac{x - K_2}{K_1 - K_2} & For\ K_1 \le x \le K_2 \end{cases},$$

$$\alpha = F_K^M(x) = \begin{cases} \dfrac{K_1 - x}{K_1 - K_2} & For\ K_1 \le x \le K_2 \\ 1 & For\ K_2 \le x \le K_3 \\ \dfrac{K_4 - x}{K_4 - K_3} & For\ K_3 \le x \le K_4 \end{cases}, \tag{7}$$

$$\alpha = F_K^H(x) = \begin{cases} \dfrac{x - K_3}{K_4 - K_3} & For\ K_3 \le x \le K_4 \\ 1 & For\ K_4 \le x \le K_{max} \end{cases}$$

Finally, it is to generate select the optimal chromosomes. The $F_K^L(x)$, $F_K^M(x)$ and $F_K^H(x)$ are corresponding correspond to linguistic function of the low middle, and high linguistic function. After define the membership functions of chromosomes selection, each new population's average fitness value will be mapped to one of the linguistic functions and transformed to the number of chromosomes will be selected. If the chromosomes selection is low, it will select the number of L_i chromosomes will be selected, and so on for to M_i and H_i. L_i, M_i and H_i are defined as Eq. (8), in which W_L, W_M, W_H are weights.

$$L_i = N \times W_L,\ M_i = N \times W_M,\ H_i = N \times W_H \tag{8}$$

4) Crossover: A single-point crossover is used in the paper. For the single-point crossover, two chromosomes T_1 and T_2 are randomly selected from the Gene pool. A cutting point from T_1 and T_2, which is at the same location, respectively, is then randomly selected. The cutting point is called the crossover site. After applying the crossover operator to T_1 and T_2, two new offspring chromosomes T_3 and T_4 are generated.

5) Mutation: The mutation operator is used to change the gene value in the chromosome for next generation. The types of mutation use a bit-flip to exchange the value of a gene of a chromosome from 0 to 1 or from 1 to 0 and to generate a new offspring chromosome by the happen probabilities of Mp.

6) Local Search: We also combined a genetic algorithm with local search to enhance the results. After crossover operations, we randomly chose two chromosomes *Ta* and *Tb* to be parents, and then calculate the respective cost and delay for the two chromosomes of all paths from source node *s* to various destinations *D*. Chromosome *Ta* is calculated by $Ta=\{pa(s,d_1), pa(s,d_2), \ldots, pa(s,d_n)\}$, and chromosome Tb is calculated by $Tb=\{pb(s,d_1), pb(s,d_2), \ldots, pb(s,d_n)\}$. The sub-fitness function $p(s,d_i)$ is defined as:

$$P(s,d_i) = C(s,d_i) \times \Theta(D(s,d_i)),\quad \Theta(f) = \begin{cases} 1 \ldots . f \le 0 \\ 2 \ldots . f > 0 \end{cases} \tag{9}$$

Next, we compared both sub-fitness paths with the same destination from the tow chromosomes respectively, and then added the one with the higher sub-fitness value to the chromosome of the offspring. After all of the higher sub-fitness paths have been added to the chromosome of the offspring, a new offspring is generated. For example $Ta = \{2 \rightarrow 7,\ 2 \rightarrow 7 \rightarrow 9 \rightarrow 5,\ 2 \rightarrow 7 \rightarrow 9 \rightarrow 5 \rightarrow 4\}$ and $Tb = \{2 \rightarrow 7,\ 2 \rightarrow 7 \rightarrow 8 \rightarrow 4,\ 2 \rightarrow 7 \rightarrow 8 \rightarrow 4 \rightarrow 5\}$ are parents. And a new offspring of chromosomes $Tc = \{2 \rightarrow 7,\ 2 \rightarrow 7 \rightarrow 9 \rightarrow 5,\ 2 \rightarrow 7 \rightarrow 8 \rightarrow 4\}$ is generated. Then Local search compares the fitness values of the three chromosomes with the fitness values of the other chromosomes in the gene pool. If one of three chromosomes has a higher fitness value than the other chromosomes, it will replace the chromosome with the lower fitness value.

7) Check and Repair Operation: After performing these genetic operations above, the check and repair operation is going to avoid producing an illegal multicast tree. First, compare the sub-graph of chromosome with original network graph. If it is not a tree connection from a source node to all the destination nodes, repair it. We apply the graph search randomly connect the destination nodes until all destinations include in. Another case, if the sub-graph has cycle situation happened, we randomly selected a subset of nodes and breaks the cycle into legal a tree by removing the links that are incident to the selected node. Finally, above two situations, the check and repair were repeated until a multicast tree is constructed.

3. SIMULATION RESLUTS AND ANALYSIS

In the simulations, we first generated a variety of network graphs with different numbers of nodes ($|V|$=16,32,60), inter-connected to each other. And then selected one of the nodes as a source node s and randomly a number of destination nodes ($|D|$=8,16,32) for experiment input data.

Then according to the simulation results the comparisons of effectiveness between Tsai's GA (MSGA) and GAFL for the multicast routing were made later, respectively.

Fig. 2 shows comparison of minimum cost between our proposed GAFL with MSGA. It was shown us that with the number of nodes increasing, the value of minimum cost and average cost of both methods go very fast respectively. But

GAFL can achieve better minimum cost and average cost than MSGA for all test networks, especially in the case of large number of network nodes.

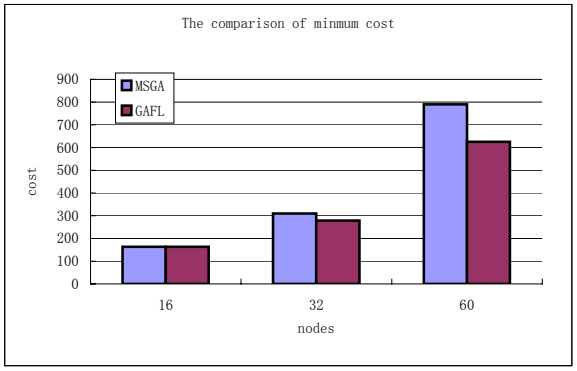


Fig. 2. The comparison of minimum cost between GAFL with MSGA

Fig. 3 and Fig. 4 respectively makes a comparison of the convergence process between GAFL and MSGA with 16, 64 nodes. Both figures show the GAFL can find the global optimization quickly than MSGA, especially in the large-scale test networks.

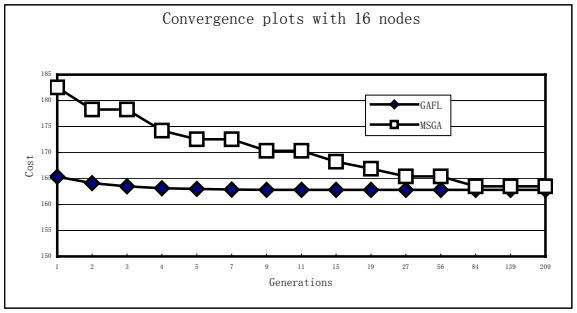


Fig. 3. The comparison of convergence process between GAFL and MSGA with 16 network nodes

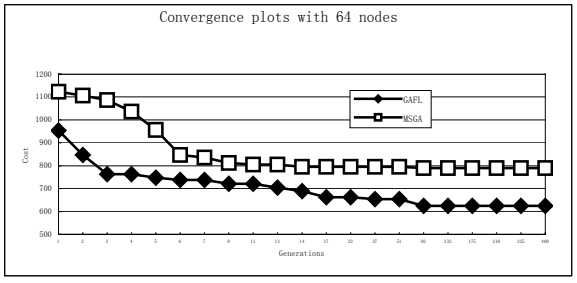


Fig. 4. The comparison of convergence process between GAFL and MSGA with 64 network nodes

Fig. 5 shows the running time of using MSGA and GAFL for different nodes. In Fig. 5, all of the data obviously indicate that running time of both algorithms goes very fast as the number of the network nodes increase. What it is worse, the running time of the GAFL algorithm increases faster than MSGA. But the GAFL algorithm can always obtain the minimum cost than MSGA in time.

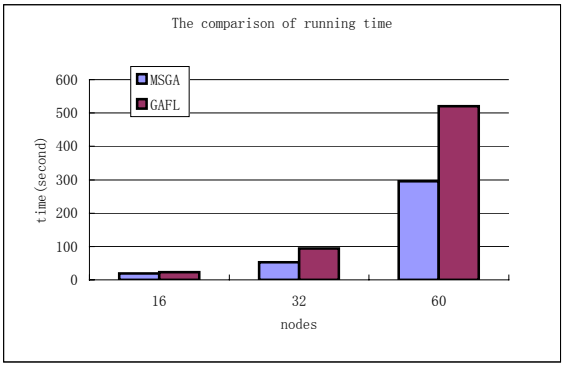


Fig. 5. The comparison of running time between GAFL with MSGA

4. CONCLUTIONS

Since finding the minimum cost multicast tree with delay constraint is an NP-hard problem, we modified traditional GA by using the fuzzy selection and local search to solve it. Then the primary simulation results demonstrated that GAFL could quickly obtain global optimal solutions compared with MSGA for different nodes and can find better minimum cost tree with delay constraint within the same generations.

However with the network nodes and edges increasing, the GAFL may need more running time than MSGA. Therefore in future work, we will put our emphasize on introducing some proper heuristics algorithm into population initialization or using improved crossover or mutation operations in order to reduce running time of GAFL.

5. REFERENCES

[1] J. Holland, Adaptation in Neural and Artificial Systems, University of Michigan Press 1975.

[2] N. Shimamoto, A. Hiamatsu, et al, “A Dynamic Routing Control Based on A Genetic Algorithm”, in Proceedings of IEEE International Conference on Neural Network, 1993,1123-1128.

[3] C. F. Tsai, C. W. Tsai, et al, “A Novel Algorithm for Multimedia Multicast Routing in A Large Scale Network”, The Journal of Systems and software 72, 2004,431-441.

[4] Z.Wang, B. Shi, et al, “Bandwidth-Delay-Constrained Least-Cost Multicast Routing Based on Heuristic Genetic Algorithm”, Computer Communications, Vol.24, 2001,685-692.

[5] A. T. Haghighat, K. Faez, et al, “GA-Based Heuristic Algorithms for QoS Based Multicast Routing”, Knowledge-Based System 16, 2003,305-312.

[6] A.T.Haghighat,K.Faez,et al,“Genetic Algorithm-Based Heuristic Algorithms for Bandwidth-Delay-Constrained Least-Cost Multicast Routing”, Computer Communications 27, 2004, 111-127.

[7] P. Chen, T. l. Dong, “A Fuzzy Genetic Algorithm for QoS Multicast Routing”, Computer Communications 26, 2003,506-512.

Distributed Web log Mining Based Collaborative Filtering Recommendation Algorithm

Xun Wang, Yun Ling, Biwei Li
College of Computer & Information Engineering, Zhejiang Gongshang University
Hangzhou 310035, China
Email: wangxun@mail.zjgsu.edu.cn

ABSTRACT

For distributed large commercial mirror sites, this paper presents a distributed web log mining based collaborative filtering recommendation algorithm. Based on user-browsing preference, the user access matrixes of different mirror sites are constructed, then synthesized and standardized. Taking the standardized synthesized user access matrix as raw data, this paper proposes utilizing web page similarities to predict the rating for pages not having been rated, thus increasing the pages that have been jointly rated among users. This method could effectively solve the sparsity of user ratings in collaborative filtering and improve the accuracy of the calculation of the nearest neighbor of the target user. The experimental results show that this algorithm is applicable to the popular distributed web server clustering architecture, avoids the inaccuracy and complexity of web page's manual ratings, effectively solves the sparsity of user rating data of the traditional collaborative filtering algorithm and enhances the recommendation quality.

Keywords: Distributed Web log mining, Multi-Agent, recommendation system, collaborative filtering, preference.

1. INTRODUCTION

With the wide application of Internet and E-commerce, web has been turned into an important approach for information acquiring. There is pressing demands on the recommendation systems, which could actively provide users with individualized information services. Through applying clustering technology, collaborative filtering system could locate n most similar neighbors, predict user interests on the basis of the other similar users' interests and recommend information, which is a relatively successful method of current individualized recommendation systems [1,2,3]. Compared with the content-based filtering system, collaborative filtering system could automatically filter the information that the system could not analyze and represent, and has the capability to recommend up-to-date information. But there also exists the problem of sparsity. On one hand, the number of information items greatly exceeds what the user could absorb; while on the other hand, generally, user rarely gives one's interest ratings for the information browsed, making the user access matrix quite sparse, so as hard to discover similar users and provide accurate recommendations.

With the expansion of web sources, a single server could not adapt to the web development, such as on storage, accessing speed, web bottleneck riddance, thus distributed web server architecture has emerged. Now, the distributed architecture of most big web sites is based on the clustering architecture shown in Fig 1. One feature of clustering architecture is that the web pages of web sites are mirrored, so the web logs of user access are stored on the respective servers.

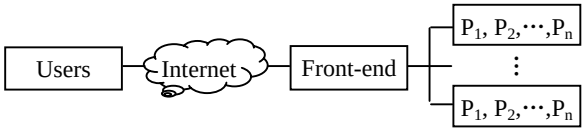


Fig. 1. Web clustering architecture

Currently, many documents have presented the solutions to single serve web log mining [5-6]. Based on these researches, this paper proposes a distributed web log mining model aiming at the architecture shown in Fig 1. This model takes comprehensive considerations of the information such as user browsing time and web page length and adopts multi-agent technology to execute distributed parallel mining, so as to obtain the user browsing preference of the pages of mirror sites. With user rating replaced by preference, this model constructs the user access matrix for every mirror site and combines the records of identical users into the synthesized user access matrix. Then, the similarities of information items are used to predict the ratings for web pages not having been rated. On this basis, correlation similarity is used to compute the nearest neighbor of the target user and engenders recommendation.
The rest of the paper is organized as follows. Section 2 presents the architecture of the multi-agent based distributed web log mining system, section 3 proposes the preference based on Web-log mining and constructs the user access matrix. Section 4 proposes the collaborative filtering recommendation algorithm based on web page similarities to predict the ratings for pages not having been rated. Section 5 gives out the simulation experiment of the collaborative filtering recommendation system and analyzes the results. Section 6 is the conclusion, describing the problems remained and further research directions.

2. DISTRIBUTED ARCHITECTURE

Multi-agent based distributed web log mining adopts three-layer architecture, which is composed by user access layer, agent layer and web server layer respectively as shown in Fig 2.

User access layer is the interface for user browsing. Web server layer is composed by multiple Web servers, in which every web server provides its corresponding web log, and interacts with the agent layer through a remote Adapter. Agent layer is the bridge between user access layer and web server, designed to solve the problem on parallel mining of heterogeneous data and ensure the integrity of data. The detailed functions of various agents are described as follows: User Agent: analyzes the user requests from user access layer, transmits the requests to the decision Agent, and simultaneously returns the results of decision agent processing to the user access layer.

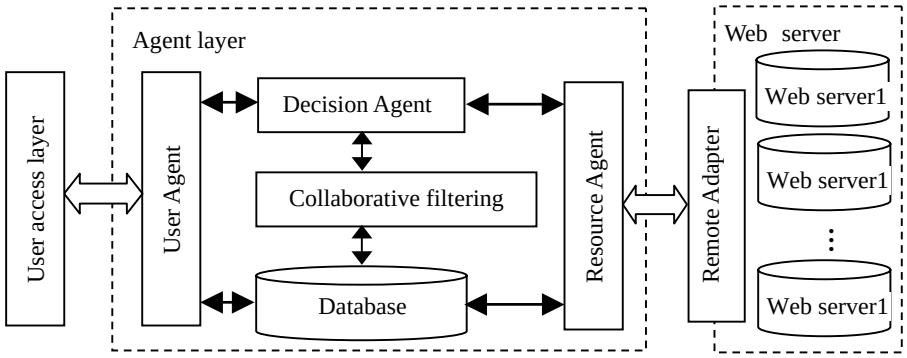


Fig. 2. System architecture

Decision Agent: the center of all agents, which is responsible for accepting the requests from user agent, transforms these requests into tasks, guides the functioning of the collaborative filtering agent and resource agent, and later returns the results to user agent.

Collaborative filtering Agent: accepts the commands from decision agent, synthesizes user access matrix of different mirror sites obtained by resource agent, adopts relevant algorithm to compute the nearest neighbor of the target user based on the synthesized user access matrix, generates the collaborative filtering recommendation based on correlation similarities and returns the recommendation to decision agent.

Resource Agent: accepts the commands from decision agent, exerts the control over the remote agent on the Web server, analyses and processes the files after remote agent's preprocessing, constructs and stores user access matrix of corresponding mirror sites.

Remote Agent: located on the Web server, which accepts the commands from resource agent, realizes the preprocessing procedures of Web mining and provides the results after remote agent preprocessing.

3. USER ACCESS MATRIX

Web log of mirror sites contains information about the pages, time, ID when user browses web pages. By mining those web log data, a user access matrix could be constructed to further discover user-browsing interests. Web log mining includes data preparation and preference calculation.

Data preparation

Data preparation focuses on transforming web log into reliable data suitable for mining, including data reduced, user identification, session identification, transaction identification, and etc.

Data reduced is the process to eliminate redundant information in the web log, which is irrelevant to data mining, such as the records with suffixes of GIF, JPEG and CGI, request method, error code, transmission protocols, and etc. After the processing, only the information relevant to user's browsing interests is retained, such as the IP address, user ID, requested URL, access time and bytes.

Since the involvement of user privacy, user identification could not simply depend on the ID number. Considering the existence of firewall, local cache and proxy server, currently user identification is realized by using cookies and the IP address and distant agent binding technology. Besides those methods, website topology and web page classification information are fully exploited to improve the efficiency of user identification.

Session identification refers to the continuous web page requests from the same user, and it aims at categorizing the pages every user has visited into segments of user browsing actions. Currently, time stamp is widely used to process the session identification. If the time difference of user page accessing exceeds a certain time stamp, then it is regarded as the beginning of a new session.

Construction of user access matrix

After synthetically taking the browsing frequency and browsing time into consideration, document [7] proposed an algorithm of correctly mining user interest path. But research discovers that user-browsing time could not accurately reflect user interests since user browsing time is directly related to the length of the text on web page. So this paper comprehensively takes browsing frequency, browsing time and text length into consideration, and proposes a improved preference method to accurately reflect user browsing interests.

Definition 1（Preference） Denote U as the set of URLs of all web pages, and denote W as the set of all browsing sub paths. If there exists $w \subset W$, for $\forall x \in w$（x is the browsing sequence of pages composed by $\forall u \in U$, where the jth browsing pages is called j bit), the first m bits of those pages are identical, but since there are n different browsing pages for the m+1 bit, called n different choices for bit m, and the preference of kth (k=1, 2, …, n) choice is defined as:

$$P_k = (C_k \cdot V_k) \Big/ ((\sum_{i=1}^{n} C_i) \cdot (\sum_{i=1}^{n} V_i) / n^2)$$

Where Ci represents the support of the ith choice, meaning the number of times entering next page through the ith choice. Vi is defined as: $V_i = \lfloor URL.T_i / URL.S_i \rfloor$, where URL.Ti represents the sum of discretized time entering next page through the ith choice. URL.S represents the length of the text browsed.

Discretized time refers to applying discretization technology to user browsing time representation, dividing time into intervals and using the index of intervals to denote real time. Suppose that the time spent on a certain page could be categorized into four types: passing, simple viewing, normal viewing and interest viewing, URL.T could be discretized into:

$$URL.T=\begin{cases}1, & 0<t\le URL.T_{\max\_pass\,in\,g}\\ 2, & URL.T_{mzx\_pass\,in\,g}<t<URL.T_{\max\_simple\_viewing}\\ 3, & URL.T_{mzx\_simple\_viewing}<t<URL.T_{\max\_normal\_viewing}\\ 4, & URL.T_{\max\_normal\_viewing}<t\end{cases}$$

where URL.Tmax_passing, URL.Tmax_simple_viewing, URL.Tmax_normal_viewing represent the predetermined maximum passing time, maximum simple viewing time and maximum normal viewing time respectively.

Take Pk as the rating for the corresponding URL that user has accessed and construct the user access matrix.

$$\begin{bmatrix} & URL_1 & URL_2 & \cdots & URL_n\\ USER_1 & P_{1,1} & P_{1,2} & \cdots & P_{1,n}\\ USER_2 & P_{2,1} & P_{1,1} & \cdots & P_{1,n}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ USER_m & P_{m,1} & P_{m,1} & \cdots & P_{m,n}\end{bmatrix}$$

where the row vector represents the ratings made by the same user for different web pages, while the column vector represents the ratings made by different users for the same web page.

Synthetization and standardization of user access matrix

Making the user access matrix of mirror sites in the database as object, collaborative filtering agent synthesizes the records of identical users into the comprehensive user access matrix, representing the ratings of all users for the whole web site. To standardize user access matrix, the data will be compacted into the closed interval [0, 10]. The mean value and the standard deviation rated by the jth user for n web page in the user access matrix would become:

$$\overline{p}_j=\frac{1}{n}\sum_{i=1}^{n}p_{i,j},\quad s_j=\left|\frac{1}{n}\sum_{i=1}^{n}(p_{i,j}-\overline{p}_j)^2\right|^{1/2}$$

Then the entries in the user access matrix would be standardized to:

$$p'_{i,j}=\frac{(p_{i,j}-\overline{p}_j)}{s_j}$$

Applying the standardization formula, the standard data would be compacted into the closed interval [0, 10],

$$\overline{p}_{i,j}=10\times\frac{p'_{i,j}-p'_{\min,j}}{p'_{\max,j}-p'_{\min,j}}$$

Where $p'_{\min,j}$ and $p'_{\max,j}$ are the minimum and maximum values of $p'_{1,j}, p'_{2,j}, \cdots, p'_{n,j}$ respectively. After standardization, the user access matrix becomes:

$$\begin{bmatrix} & URL_1 & URL_2 & \cdots & URL_n\\ USER_1 & \overline{P}_{1,1} & \overline{P}_{1,2} & \cdots & \overline{P}_{1,n}\\ USER_2 & \overline{P}_{2,1} & \overline{P}_{1,1} & \cdots & \overline{P}_{1,n}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ USER_m & \overline{P}_{m,1} & \overline{P}_{m,1} & \cdots & \overline{P}_{m,n}\end{bmatrix}$$

4. RATING PREDICTION BASED COLLABORATIVE FILTERING RECOMMENDATION

Computation of similarity

For large E-commercial web sites, the items with user rating are normally no more than 1%, with even lower rating for an item to have been rated by two users. Similar to that situation, user-rating data for mirror sites is certainly to be extremely sparse, even after synthetization and standardization, comprehensive user access matrix is normally much sparse. Since traditional methods of cosine and correlation similarity measurement are ineffective in computing the nearest neighbor of the target user, the quality of collaborative filtering algorithm could not be fully guaranteed. Document [1] proposed using the mean value of ratings, which could to a certain extent enhance the recommendation accuracy of collaborative filtering, but could not completely solve the problems with traditional similarity measurement method under the circumstances of sparsity of user rating data. Document [8] proposed using singular value decomposition (SVD) to reduce the dimensions of item space, so as to make every item have user ratings for every space after dimension reducing, which significantly improves the extensibility of the recommendation system. Since dimension reducing could result in information loss, under the circumstances of very large item dimensions, the effect of dimension reducing is not guaranteed [9].

Based on the documents [3,10], this paper proposes the calculation of web page similarities, predicting the ratings for pages from the ratings made by the users for similar pages and increasing the pages that have been jointly rated among users. Thus could effectively overcome the drawbacks of traditional method of similarity measurement under the circumstances of sparseness of the user access matrix and make the calculation of the nearest neighbor of the target user more accurate.

This paper adopts correlation similarity to compute the similarities between users. Iij is used to denote the set of pages that have been rated by both users i and j, and the similarity sim (i, j) between users i and j could be measured by the Pearson correlation coefficient:

$$sim(i,j)=\frac{\sum_{c\in I_{i,j}}(R_{i,c}-\overline{R}_i)(R_{j,c}-\overline{R}_j)}{\sqrt{\sum_{c\in I_{i,j}}(R_{i,c}-\overline{R}_i)^2}\sqrt{\sum_{c\in I_{i,j}}(R_{j,c}-\overline{R}_j)^2}}$$

Ri,c represents the rating made by user i for page c, while $\overline{R}_i$ and $\overline{R}_j$ represent respectively the mean of ratings made by users i and j for pages.

Searching nearest neighbor

In computing the similarity between users i and j, first the set U_{ij}, the union set of pages rated by users i and j, is

calculated. Denote I_k the page set rated by user k, then $U_{ij}=I_i \cup I_j$.

Denote N_i the page set within web page space U_{ij} that user i has not rated, then $N_i=U_{ij}-I_i$.

For any given page $p \in N_i$, the following method is used to predict the rating p_i, p for page p by user i.

Step 1: similar to the calculation of user similarity, compute the similarities between page p and other pages. First obtain all the ratings made by users for pages i and j, then compute the similarity between pages i and j using the method of correlation coefficient measurement.

Step 2: select certain pages with the highest similarities as the set of neighboring pages for page p, namely searching the whole web page space for the set $M_p=\{I_1, I_2, \cdots, I_v\}$, for $p \notin M_p$, the page I1 and the page p has the highest similarity $sim(p, I_1)$, while the page I2 and the page p has the second highest similarity $sim(p, I_2)$, and so on.

Step 3: after obtaining Mp, predict the rating P_i, p for page p by user i using the method proposed by Sarwar [12]:

$$P_{i,p}=\frac{\sum_{n\in M_p} sim_{p,n}\times R_{i,n}}{\sum_{n\in M_p}(\left|sim_{p,n}\right|)}$$

After taking these three steps, all the pages within set U_{ij} have been evaluated by both users i and j, namely for any given page $p \notin U_{i,j}$, the rating made by user i for page p would be:

$$R_{i,p}=\begin{cases} r_{i,p}, & if \quad user \quad i \quad rated \quad page \quad p \\ P_{i,p}, & if \quad user \quad i \quad not \quad rated \quad page \quad p \end{cases}$$

Then based on the union set Uij, the similarity between users i and j could be computed through the method of correlation similarity measurement.

Finally, for every user u, searching the whole user space for user set $C=\{C_1, C_2, \cdots, C_k\}$, for $u \notin C$, C1 and u has the highest similarity sim (u, C_1), while C_2 and u has the second highest similarity sim (u, C_2), and so on.

Collaborative filtering recommendation

After obtaining the nearest neighbor of target user, the following scheme could generate the recommendation. Denote NBS_u the set of the nearest neighbor of user u, then the rating prediction $P_{u,i}$ for page i by user u could be obtained from the rating made by user u for the set NBS_u, using the following formula [1]:

$$P_{u,i}=\overline{R_n}+\frac{\sum_{n\in NBS_u} sim(u,n)\times(R_{n,i}-\overline{R_n})}{\sum_{n\in NBS_u}(\left|sim(u,n)\right|}$$

sim(u, n) represents the similarity between users u and n, while $R_{n,i}$ represents the rating made by user n for page i, and $\overline{R_u}$ *and* $\overline{R_n}$ represent respectively the mean of rating made by users u and n for pages.

5. EXPERIMENTAL RESULTS AND ANALYSIS

To test the efficiency of the proposed algorithm, we had randomly selected 600 papers published on core journals from CNKI. Those papers focused on artificial intelligence and were published between 2000 and 2005. The participants were 38 undergraduates taking the course. The simulation experiment is processed in two steps.

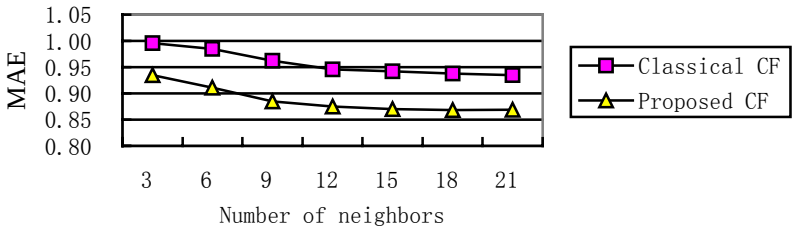


Fig. 3. Comparison of accuracy of recommendation

Step 1: quality test on the rating prediction based collaborative filtering algorithm. In the experiment, participants were required to rate at least 25 randomly selected papers from those 600 papers, so as to get 1129 ratings. Making 500 papers as the training set and the rest 100 papers as testing set, the experiment adopted MAE (mean absolute error) and correlation similarity as recommendation quality and similarity measurement criteria respectively to compare the traditional collaborative filtering algorithm and rating prediction based collaborative filtering algorithm. The experimental results are shown in Fig 3.

where MAE is defined as: $MAE=\frac{\sum_{i=1}^{N}\left|p_i-q_i\right|}{N}$, while $\{p_1, p_2, \cdots, p_N\}$ represents the set of the predicted user ratings and $\{q_1, q_2, \cdots, q_N\}$ represents the corresponding set of actual user ratings, $N=3,6,\cdots,21$.

As seen from Fig 3, under conditions of selecting various nearest neighbors, proposed rating prediction based collaborative filtering algorithm has smaller minimal MAE when compared with the traditional collaborative filtering algorithm, which suits the previous prediction. Under the circumstances of extreme sparsity of user rating data, the items jointly rated by two users are quite rare, so the accuracy of traditional collaborative filtering algorithm of

computing correlation similarity is hard to be guaranteed. The rating prediction based collaborative filtering algorithm could predict the rating for pages from the ratings made by the users for similar pages, thus increasing the pages that have been jointly rated among users, and adopts the correlation similarity measurement method to compute the similarities among users and effectively improve the recommendation quality of the recommendation system.

Step 2: simulation test on comprehensive recommendation quality. Based mainly on these 600 papers, an experimental system was constructed and every participant was assigned a unique account. The experiment required every participant browse these papers according to ones own interests for no less than forty hours within a semester. A certain measure is taken to guarantee the realness of the simulation test. (The experiment demanded each participant write a reading report according to ones own interests, who would be a main part of the results of rating for the course.)

At the end of the semester, the reading time of every participant has exceeded 40 hours. 24 participants were selected to log in the system using the original account and rated the first five papers recommended in 10 point scale. The average rating is 7.95, which means that the proposed algorithm is applicable and has excellent recommendation quality.

6. ACKNOWLEDGEMENT

This work is supported by in part the foundation of science and technology department of Zhejiang province under grant No. 2005C34002, Zhejiang natural science foundation under grant No. Y105579.

7. CONCLUSIONS

With the rapid increasing of Internet application and scale, commercial web sites with a single server could no longer meet the demands, so most large web sites have adopted the architecture of distributed multiple Web servers. This paper proposes a distributed web log mining based collaborative filtering recommendation algorithm aimed at the architecture of distributed mirrored commercial web sites. Multi-agent technology is used to preprocess the web log of mirror sites, and user rating is replaced by preference to construct the user access matrix. Then the agent layer synthesizes and standardizes the records of identical users of mirror sites into the comprehensive user access matrix. Based on the comprehensive user access matrix, this paper proposes the calculation of web page similarities, predicting the ratings for pages from the ratings made by the users for similar pages and increasing the pages that have been jointly rated among users. Thus could effectively solve the problem of sparsity of user ratings of collaborative filtering and improve the accuracy of the calculation of nearest neighbor of target users.

The proposed algorithm is applicable only to the architecture of distributed mirror sites and adopts the centralized collaborative filtering recommendation, so the performance of the algorithm is affected. If collaborative filtering recommendation is executed independently on each mirror site web log and the recommendation results are synthesized according to a certain threshold, then the results might be better. Furthermore, this algorithm is tested only in the simulation, thus lacking the test under distributed net environment, so the reliability and performance of the algorithm needs to be further proved. The commercial web sites with architecture of distributed server clustering without mirroring occupy certain proportion, so how to apply efficient mining on the web log of this architecture is the focus of further research.

8. REFERENCES

[1] J. Breese, D. Hecherman, et al, “Empirical analysis of predictive algorithms for collaborative filtering”, In: Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI98), 1998, 43-52.

[2] S. Badrul, G. Karypis, et al, “Analysis of Recommendation Algorithms for E-commerce[C]”, Proceedings of EC’00, 2000.

[3] S. Badrul, G. Karypis, et al, “Item-based Collaborative Filtering Recommendation Algorithms”, In: Proceedings of the 10th International World Wide Web Conference, 2001, 285-295.

[4] M. Sayal, P. Scheuemann, “Distributed Web Log Mining Using Maximal Larger Itemsets [J]”, Knowledge and Information Systems, 2001, 3: 389-404.

[5] D.L. Yang, S.H. Yang, et al, “An efficient web mining for session path patterns”, In: Proceeding of International Computer Symposium2000, Workshop on Software Eng. And Database Systems, Taiwan, 2000, 107-113.

[6] D.SH. Xing, J.Y. Shen, et al, “Discovering preferred browsing path from Web logs [J]”, Chinese journal of computers, 2003, 26(11): 1518-1523.

[7] J. Srivastava, R. Cooley, et al, “Web usage mining: Discovery and applications of usage patterns from Web data [J]”, SIGKDD Explorations, 2000, 1(2): 12-23.

[8] S. Badrul, G. Karypis, et al, “Application of dimensionality reduction in recommender system—A case study”, In: ACM Web KDD 2000 Workshop, 2000.

[9] C.C. Aggarwal, on the effects of dimensionality reduction on high dimensional similarity search, in: ACM PODS Conference, 2001.

[10] A.L. Deng, Y.Y. Zhu, et al, “A collaborative filtering recommendation algorithm based on item rating prediction [J]”, Journal of Software, 2003, 14(9): 1621-1628.

Research and Implementation of Scheduler LFB Supporting DiffServ in ForCES Architecture*

Li Lin, Weiming Wang,
College of Information and Electronic Engineering，Zhejiang Gongshang University
149 jiaogong Road, Hangzhou, P.R.China, 310035
Email: linli@pop.zjgsu.edu.cn, wmwang@mail.zjgsu.edu.cn
And
Fenggen Jia
National Digital Switching System Engineering & Technological R&D Center
Zhengzhou, Henan province, China
Email: fgjia@mial.zjgsu.edu.cn

ABSTRACT

QoS is the key technology in the research of next generation network (NGN). DiffSrev QoS is widely recognized since its advantage in the scalability for QoS deployment. ForCES exploit the router architecture of NGN, which supports QoS. A Scheduler LFB model supporting DiffServ in ForCES router is researched. Based on analysis of ForCES router architecture and the associated scheduler policy supporting DiffServ, the scheduler LFB is also implemented. The implementation is based on the IXDP2401 network processor. The experiments have shown the feasibility of the scheduler model and the associated scheduler strategies.

Keywords: ForCES, LFB, DiffServ, Scheduler policy, PHB.

1. INTRODUCTION

ForCES (Forwarding and Control Element Separation) are a working group under the routing area of IETF, which devotes on the open programmable IP router architecture for Next-generation network. The ForCES group proposed an IP router architecture consisting of multiple Forwarding Elements (FEs) and Control Elements (CEs). The ForCES protocol [1] defines the interaction between CEs and FEs.

Next-generation IP networks are expected to provide guaranteed end-to-end QoS service. DiffServ [2] (differentiated service) is regarded as one of the main solutions to resolve the QoS problem of IP network. In this paper, based on the analization of ForCES router architecture, we present an implementation of Scheduler LFB that supports Diffserv in ForCES router.

2. ARCHITECTURE OF FORCES ROUTER

An IP network element is composed of numerous logically separate entities that cooperate to provide a given functionality (such as a routing or IP switching) and still appear as a normal integrated network element to external entities. Two primary types of network element components exist: control-plane components and forwarding-plane components. ForCES group defines architecture and protocols to standardize the information exchanged between CEs and FEs, which allows CEs, and FEs to become physically separated standard components.

According to the ForCES requirement and framework protocol, open programmable IP networks locally consist of Network Elements (NE), which comprise multiple(up to hundreds of) FEs and at least one CE (Note: there possibly exist redundant CEs for better reliability). A CE is usually located on a different network node form FEs.

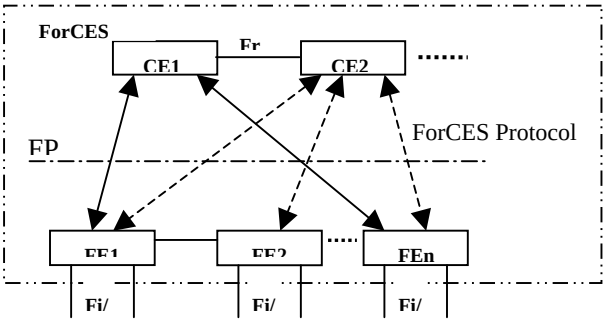


Fig. 1. Architecture of ForCES router

Architecture of ForCES router is shown in fig.1. CE primarily processes ForCES messages and is responsible for the management of LFBs in FEs. FE primarily processes and forwards data packets in line-rate. FE forwards the majority of data packets and some packets involved with routing and protocol messages are redirected to CE for further processing.

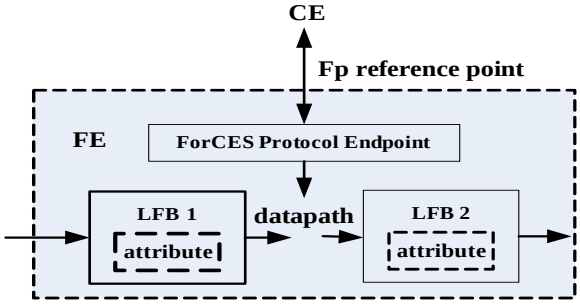


Fig. 2. FE Architecture

The Architecture of FE is defined by FE Model [3], in which the resources of FE are modeled as various kinds of LFBs, each of which has a single function, e.g., classifier, scheduler, IPv4 or IPv6 forwarder. Those LFBs are

* Project sponsored by National 863 Project (2005AA121310), NSF China (60573116), and Zhejiang Key Sci&Tech Project (2005C21013)

connected by datapath, forming specified topology and realizing dynamic configuration of resources to provide various IP services. The topology and attributes of LFBs are controlled and managed by corresponding messages defined by ForCES protocol.

The requirements of LFB are defined in ForCES router are defined in ForCES requirement protocol [6], e.g., (1) LFB should support attributes such as dynamically adding, modifying and deleting; (2) supporting query of LFB topology; (3) supporting query of LFB attributes; (4) supporting LFB dynamically joining and leaving system, etc. In this paper, we design and implement LFB dynamic configuration and loading module on the basis of those LFB requirements above.

3. SCHEDULER STRATEGY SUPPORTING DIFFERSERV

Diffserv is an architecture brought forward by IETF to assure QoS guarantee for IP networks. The whole procedure can be depicted as follows: when customer's flows enter DiffServ Domain, the domain border routers will first classify customer's flows into different level of service bases on Service Level Agreement (SLA) between customers and Internet Service Providers (ISPs), and then mark them with different DSCP values; intea-DS routers classify customer's flows into different queues based on DSCP field and process them with different PHB (Per-Hop Behaviors). Customer flows traverse the differserv domain hop by hop, receiving the same PHB on each router and thus we can provide different service to customers. IEFT has presently define three types of PHB: EF (Expedited Forwarding) [4], AF (Assured Forwarding) [5] and BE (Best Effort, default PHB). A specifically PHB can be implemented through different scheduler algorithm and active queue management mechanism.

3.1 Standard PHB

Expedited Forwarding (EF): The expedited forwarding per-hop behavior assures that any traffic class with EF's related DSCP is given highest priority and is not queued. EF was designed to be used to build a low-loss, low-latency, low-jitter, assured bandwidth service. A packet that is marked with EF DSCP will receives guaranteed low-drop precedence as the packet traverses diffserv-aware networks en route to its destination.

Assured Forwarding (AF): Assured Forwarding PHB is suggested for provider DS domain to offer different levels of forwarding assurances for IP packets received from a customer DS domain. Four AF classes are defined, where each AF class is in each DS node allocated a certain amount of forwarding resources (buffer space and bandwidth).

Best Effort (BF): Best Effort PHB provides the best-effort services. Packet that not belongs to above mentioned PHBs, will be treated as Best Effort traffic.

3.2 PHB Mechanism

Base on the characteristics of DiffServ and considering the scheduler algorithm existing, we choose active queue management mechanism WRED and the SP (Strict Priority) and DRR (Deficit Round Robin) combined scheduler algorithm to implement DiffServ in ForCES Router.

In our ForCES router, the QM module supports 16 virtual ports, each of which supports 16 virtual queues. WRR scheduling algorithm is employed among the 16 virtual ports. As Fig.3 shows, we separate queues in one virtual port into two groups. SP is implemented between the two groups, while DRR is implemented among the queues of each group. Group1 contains 3 high priority queues, i.e., queue 0 to queue2. Group2 contains the other 13 low priority queues, i.e., queue3 to queue15. Among the low priority queues, the first 12 ones employ WRED queue management strategy,

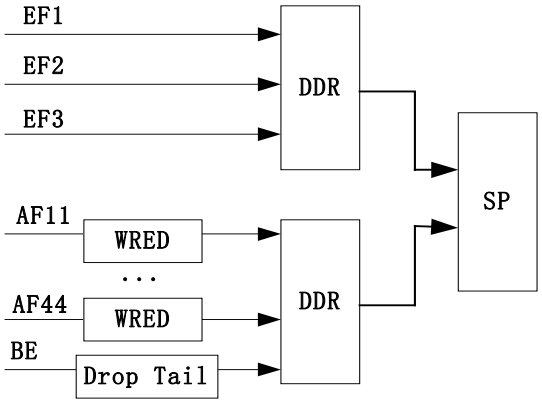


Fig. 3 Scheduler strategy supporting DiffServ

while the last one employ tail drop queue management strategy.

We setup the high priority, queue 0 to queue2, as the EF queue. The SP algorithm first scheduler the high priority queues. It scheduler the low priority queues only when the high priority queues is empty. Therefore, the high priority queues can provide low-loss, low-latency, low-jitter, assured bandwidth service, satifying the EF PHB requirement. The first 12 queues with low priority implement DRR and WRED are setup as the AF queues. With properly designed credit value, DRR can allocate different bandwidth to different queues. In this scheduler strategy, we designed 4 levels credit value, corresponding 4 kinds AF (AF1, AF2, AF3, AF4), each of which setup 3 kinds drop precedence of the packet and implemented the WRED, finally make off 12 AF queues. The last queue with the low priority is setup as the BF queue (default queue), which use the drop-tail drop policy. Then, this scheduler strategy support three kinds PHB behavior, and satisfy the DiffServ requirement.

4. IMPLEMENTATION OF SCHEDULER LFB IN FORCES ROUTER

Besides the essential capabilities of scheduler, the Scheduler LFB must satisfy the LFB requirement. Therefore, we analyze the process of Scheduler LFB from two aspects: Dynamicity and Scheduler Algorithm. Dynamicity figures that Core Component in Xscale layer can dynamically control Microblock's behavior running in Micro Engine. And Scheduler Algorithm is used to support DiffServ. All function of CE is programmed by VC++, and the development platform of FE is Intel IXDP2401.

4.1 Dynamicity Implementation of Scheduler LFB

In Intel SDK programming model, control block is provided to Core Component in Xscale layer for dynamicity implementation. In fact the control block is a block of memory in SRAM shared by Xscale and MEs. CC dynamically configures control block via CE's instruction, and the various LFBs monitor its change and determine whether add or delete LFB in the packet handle path. In order to achieve this purpose, CC needs to pass the base address of control block in SDRAM to associated MEs. In addition, the dynamical configuration of control block value should be implemented.

The dynamicity implementation of Scheduler LFB is composed of two parts: one is CC core which includes the primary function of Schedule LFB such as patch symbols in scheduler Microblock and the other is scheduler application model which realize the configuration and query of scheduler argument. The two parts is associated with message and we define two types of message:

IX_CC_SCHEDULER_MSG_SET_CONTROL_BLOCK
IX_CC_SCHEDULER_MSG_GET_CONTROL_BLOCK

for handling the configuration to control block. In message handling function of schedule CC we add process to the new message types and configuration order to control block in associated config file:

EnableSchedulerMB () /*adding Scheduler LFB on the packet handling path*/
DisableSchedulerMB () /*deleting Scheduler LFB on the packet handing path*/

at the same time we provide setting function associated with scheduler argument such as:

SetPortWeight "<portId> <newWeight>",
SetQueueCredit "<queueId> <newCredit>"

CE sets DSCP Classifier LFB state --- LFB state from OFF to ON via LFB's attribute tree for dynamically loading the LFB. The setting above is encapsulated into ForCES message via CE and sent to FE. After decapsulating the message via protocol LFB, FE analyzes its lading Scheduler LFB and then sends configuration message to Scheduler LFB and then realize the topology rebuilding of associated function LFBs. After starting Scheduler LFB, LFB topology is changed and makes ForCES router supporting QoS function. We can modify the arguments of Scheduler LFB such as weight via CE tree node and handle packets on different scheduler policy.

4.2 Algorithm Implementation of Scheduler LFB

The implementation of Scheduler module is shown in Fig.4. The Scheduler runs on a single microengine. It including 3 threads: control thread、 scheduler thread and QM message handler thread.

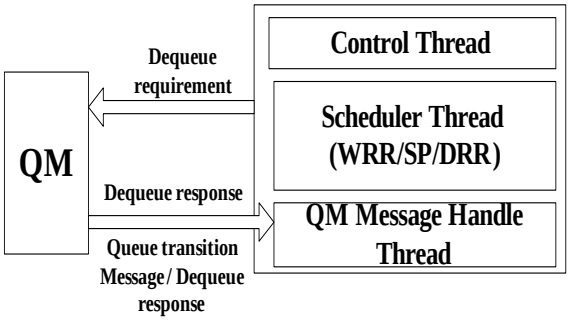


Fig. 4 . Scheduler Me

The control thread is responsible for initialization threads and monitor the SRAM where deposited the attribute such as weight, credit. Once the attribute value of SRAM changed, it will amend the attribute value of local memory to support the dynamic configuration of LFB.

The scheduler thread is responsible for actually scheduling a queue and sending a dequeue request to the QM microengine. The scheduler does WRR on the ports and the does DRR on the queues within the port. SP is implemented among the first 3 queues and other 13 queues within the same port. All schedulers use bit vectors to maintain information about which ports/queues have data and credit. Once the eligible queue is found, a dequeue request is sent by the scheduler thread to the QM.

The QM message handler thread handles messages coming back from the QM microengine. The QM microengine receives dequeue requests from the scheduler. For each request, it sends a transmit message to the TX microengine and a dequeue response to the scheduler. This response has the length of the packet dequeued and an indication if the queue went from non empty to empty (dequeue transition). If the scheduler issued a dequeue to a queue that had no data (invalid dequeue), then the packet length returned is 0. The QM may also send an enqueue transition message to the scheduler when a queue goes from non-empty to empty.

1) Port scheduler (WRR)

The process of port scheduler as follows:

1. Check vectors for port with data, positive credit and round robin mask set.
2. Update the port mask (Ensure Round robins go to the next port).
3. Extract port data structure from local memory, Decrement current credit for the port.
4. Check current credit OK. If current credit is = 0, reset the credit to initial credit and Clear the appropriate bit in the port credit vector.
5. Check packet in flight OK. If packet in flight > Max

limit, swap out, move to next port.

2) Queue scheduler (DRR)

The process of queue scheduler as follows:

1. Check vectors for queue with data, positive credit and round robin mask set
2. Update the queue mask (Ensure Round robin go to the next queue).
3. Send dequeue request to QM
4. Experimental Result and Analysis

5. TEST RESULT AND ANALYSIS

In order to testify that the Scheduler LFB we designed and implemented can satisfy the demands of LFB in ForCES and t the scheduler policy can support DiffServ, we construct the experimental platform for our testing.

5.1 Experimental environment

As shown in Fig. 5, CE connects with FE's control port via HUB, which runs CE platform software. Network tester SmartBits, which are installed in analytical host, connects with FE, which can send various data streams on demands to FE. And FE receives and handles those packets and then sends back to SmartBits.

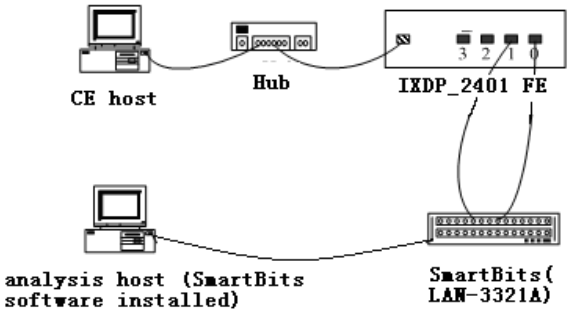


Fig. 5. Experimental Platform

5.2 Experimental process and Result

We set three flows via SmartBits which DSCP values are 1, 2 and 3. They are sent from 1000M ports and receive from the same 100M ports. The total flow is 9% of kilomega at begin and then increase by 2% until at last allocating the total flow on average. In CE's DiffServ configuration interface the below is set: three pieces of DSCP rules, here assigning queue 1 to high priority queue and assigning queue 2 and 3 to low priority queue; setting credit value in order of 1000, 8000 and 4000; the same WRED rule associated with 2 AF queue. We observer the delay of flow and the rate of loss as Fig. 6, Fig.7 below:

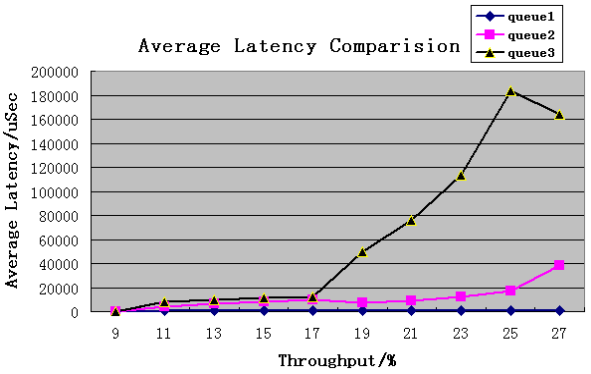


Fig. 6. Frame Loss

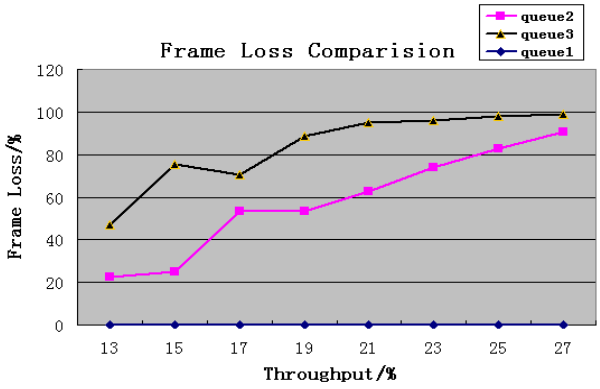


Fig. 7. Average Latency

In our testing, CE can dynamically control the behavior of FE via LFB attributes in FE. As seen from Fig. 6, the queue 1 assigned high priority has no obvious packet loss and delay but the loss of packets and average delay in queue 3 is greatly higher than queue 2. The testing result shows that the Scheduler LFB can implement the dynamical loading, the scheduler policy can satisfy the demands on low loss of packets and low delay in EF queue and the assignment of forward resources in AF queue which ensure the demands of forward services.

6. CONCLUSION

This article first put forward a kind of scheduler strategy which support DiffServ, simultaneously, proposed the realization method of the Scheduler LFB in ForCES router which implements this strategy. The experimental result proves that the scheduler LFB possesses flexibility and expansibility and its scheduler policy satisfies the demand of DiffServ.

7. REFERENCES

[1] A. Doria (co-editor), "ForCES Protocol Specification draft-ietf-forces-protocol-08.txt"[DB/OL],http://www.ietf.org/internet-drafts/draft-ietf-forces-protocol-08.txt, work in process, Sep. 2006

[2] S. Blake, D. Black, et al. "An Architecture for Differentiated Services".RFC2475, Dec.1998

[3] L. Yang, J. Halpern, et al., “ForCES Forwarding Element Model”, work in process, Sep. 2006, http://www.ietf.org/internet-drafts/draft-ietf-forces-model-06.txt
[4] V. Jacobson, K. Nichols, et al. “An Expedited Forwarding PHB”. RFC2598, Jun.1999
[5] J. Heinanen, F. Baker, et al. “Assured Forwarding PHB Group”.RFC2597, Jun.1999
[6] H. Khosravi, T. Anderson (editor), "Requirements for Separation of IP Control and Forwarding", RFC 3654, Nov. 2003

A study on HBA-based Multi-path Input/Output For Storage Area Networks

Xuelan Zhang, YuAn Tan
Department of Computer Science and Engineering, Beijing Institute of Technology
Beijing, 100081, China
Email: xlzh_9@yahoo.com.cn, victortan@yeah.net

ABSTRACT

Multiple I/O paths between the host and storage device were constructed to provide load balancing and high availability in storage area networks (SANs). A virtual host bus adapter (vHBA) layer was introduced to utilize the redundancy capability of hardware configuration, and create the path-independent device names. The I/O requests received by the vHBA's device object were forwarded to the physical host bus adapter. The paths operate simultaneously to increase the I/O throughput and reduce the I/O response time. The I/O operations were routed to the survival path after a path fails, to improve the availability of data access. The experiment shows that the proposed HBA-based solution reduces the failover time compared to the disk-based solution.

Keywords: Fault Tolerance, High Availability, Load Balancing, SCSI, Windows.

1. Introduction

In recent years, installed disk storage capacity has seen an almost exponential growth at customer sites, coupled with the movement away from direct attached storage to storage area network (SAN) or network-attached storage (NAS)[1]. Multiple concerns are addressed when developing storage solutions. A few of those concerns include multiple data paths, failover, transaction and batch elapsed time, the amount of CPU required to sustain transaction and batch processes, and concurrency. These metrics quantify a storage network and its performance.

1.1 I/O Path Management

Multiple, or redundant, data paths between a server and a storage subsystem is extremely important for optimum system performance and availability of the storage subsystem. If a failure occurs in the data path between a system server and a storage subsystem, automatic switching of the input/output (I/O) to another path is essential for availability of the system overall [2,3,4,5]. To improve the performance of data access, the host machine must be able to spread I/O workload over multiple active paths, which can eliminate bottlenecks that occur when many I/O operations are directed to a disk device across the same path. However, an operating system (OS) sees separate I/O paths to the same device and the OS is unaware that the separate bus paths are in fact connected to the same device. For example, when a path fails, an I/O error would be generated for that path instead of failing over to another path. So there is a need to provide I/O path management for redundant data paths to allow access to devices through multiple paths, the host may have driver software equipped with fail over and load balancing features.

1.2 Multipath Configuration

One way of ensuring high availability is literally to have two of everything. One can have a clustered server instead of a single server to ensure availability in case of a server failure. Two or more servers can construct a cluster and each server has multiple host bus adapters (HBAs). Each HBA is connected to a switch or Hub, which has dual paths to the dual-ported storage device. The cluster software and RAID system prevents the failure of the host server, RAID controller or disk drive from disrupting data access. A multipath I/O architecture is required on the server to utilize the redundancy hardware.

For example, the RAID system is configured to serve 8 LUNs (logical unit) of data designated as LUNs 0-7. LUNs 0-3 are assigned to Server-A and LUNs 4-7 are assigned to Server-B. For each LUN, there are two paths between the assigned server and RAID system. In the event of a path failure, I/O operations from a failed path would be automatically rerouted to a remaining path (i.e., failover), optimizing the I/O data paths. Besides the high availability, the multipath software provides load balancing of data flow and prevent a single path from becoming overloaded, causing I/O congestion that occurs when many I/O operations are directed to common devices along the same I/O path.

The multipath software would eliminates a single path element as a single point of failure, thus increasing availability of data through multiple redundant data paths and also could improve system performance by balancing the subsystem's I/O workload across the multiple paths.

2. Disk-based Layer Approach

A layered device driver approach is deployed to implement the failover and load balancing for multiple I/O paths [6]. As shown in Fig. 1., two disk device objects are created to represent two paths of the LUN. The multipath driver intercepts the I/O requests for disk objects and calls the low-level disk driver perform the actual disk I/Os. While this approach has advantages of easy implementation and good portability between different OS, it has such inefficiencies, including multiple disk device objects and the time wasted retrying the I/O request along a failed path.

The presence of multiple disk device instances for a single LUN can lead to wastefulness of system resources, include kernel memory and device number space. For example, the minor number space available for "sd" (SCSI disk) devices limits the Linux to 256 single-pathed drives. Each additional path to a storage device decreases this by a factor of 2. Another issue is that system administrators, as well as applications, are faced with a challenges when

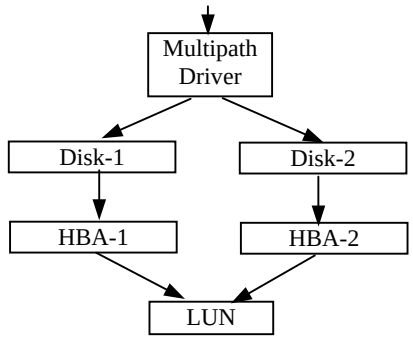


Fig. 1. A disk-based layer approach

attempting to understand and manage multipath configurations, there is no way to determine duplicate disk devices, or which disk devices are in fact the same LUN.

In Fig. 1. if an I/O request is sent via the multipath driver through disk-1 to HBA-1, and it turns out that path to LUN fails, then this is communicated back up to the disk driver, which will typically execute additional tries, until a timeout or a predetermined number of retries occur. Each try may be very time-consuming, taking up to several minutes to execute. These tries is wasted time; eventually, the disk driver stops retrying and a device error is returned to the multipath driver, then it will try the other path by trying this I/O request through disk-2 to HBA-2. Although this I/O request will succeed after the failover, the application suffers from these futile retries on the failed path, because the multipath driver has no control over the I/O request after submitting it to the disk-1 until the driver gives up.

3. HBA-based Layer Approach

A new architecture for naming and managing multipath devices is introduced. It eliminates the need for the multiple layered implementations shown as Fig. 2. In this architecture, a multipath device is represented as a single device instance, rather than having one instance per physical path. Multipath devices are attached to a pseudo bus nexus driver, called as virtual Host Bus Adapter (vHBA) driver. Here, "virtual" refers to a reconfigurable structure, in contrast to these physical HBAs.

A bus driver is a driver that manages a "bus" and its associated physical devices. The vHBA driver simulates a normal Fibre Channel Protocol/Small Computer System Interface (FCP/SCSI) HBA driver, only a single instance of vHBA driver exists. Each managed LUN is treat as a child device of this vHBA, regardless of the number of physical paths between the server and the LUN. The vHBA driver creates a path-independent name, to eliminate multiple path-dependent names to a device.

3.1 Path-dependent Device Name

Tab. 1 shows an example of the path-dependent names without vHBA driver. While all physical paths are active at startup, HBA-1 discovers LUN-0、LUN-1、LUN-2、LUN-3, disk-1、disk-2、disk-3、disk-4 are created for them. Then HBA-2 discovers these 4 LUNs with inverse order, and another 4 disks are created. In case HBA-1 or Switch-1 fails at startup, HBA-1 sees none of LUNs, server-A creates 4 disk device objects with the order HBA-2 reports 4 LUNs. Without vHBA, there are multiple path-dependent disk instances, leads to disk name confusion and resource waste.

Table 1. Path-dependant device names.

LUN	Both HBAs are up	HBA-1 fails	HBA-2 fails
LUN 0	disk-1, disk-8	disk-4	disk-1
LUN 1	disk-2, disk-7	disk-3	disk-2
LUN 2	disk-3, disk-6	disk-2	disk-3
LUN 3	disk-4, disk-5	disk-1	disk-4

The LUNs have different disk device names due to the device discovery sequence changes in different cases.

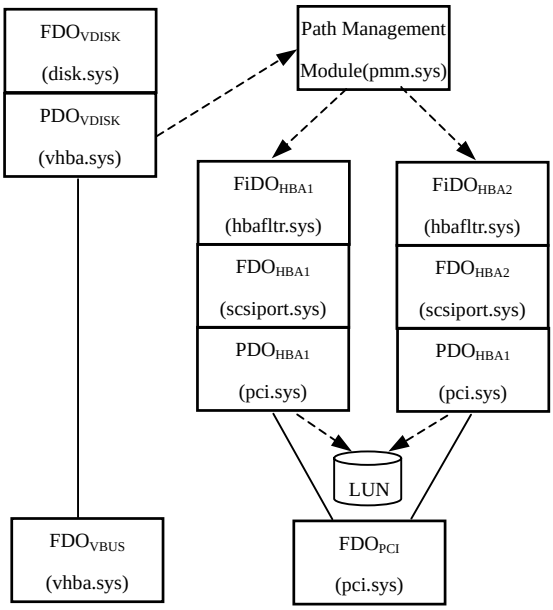


Fig. 2. Device object tree with vHBA driver

3.2 I/O Architecture with a vHBA Driver

Fig. 2. Illustrates an architecture having a vHBA driver (vhba.sys). In Windows, a driver creates device objects for each device it controls; the device objects represent the device to a device driver. These device objects include physical device objects (PDOs), functional device objects (FDOs), and filter device objects (FiDOs).

The vHBA driver exposes the managed LUNs with predetermined sequence, so the device name is not changed in case of path failure. The virtual disk number is associated with each LUN and it is path-independent.

A HBA filter driver (hbafltr.sys) is attached to all physical HBA instances to prevent them from reporting LUNs to the upper layer, so no additional disk device objects are created for these LUNs that vHBA driver has exposed.

4. I/O Fail over

The path management module (pmm.sys, PMM) is responsible for managing failover, and it maintains a path set for each virtual disk. The path set is aggregations of paths that bind specific HBA devices to the physical LUN. To identify the paths of physical LUN, PMM use the serial number page (Vital Product Data page 80h) or device identification page (Vital Product Data page 83h) defined by the SCSI command set. For a physical LUN, all paths return same serial number or device identification field, regardless of different path elements.

The vHBA driver routes I/O requests of virtual disks to

the PMM, which then forwards them to a filter HBA object that is providing transport services to the device, with the appropriate physical device information.

Since PMM is above the HBA layer, in case of a path failure, any I/O request that comes back uncompleted is retried from PMM, which has information about other available paths. As a result, futile retries illustrated in Fig. 1. can be avoided, because the level that detects the failed path is the same as the level that has information about alternative paths, PMM effectively reroutes the uncomplete I/O requests on failed paths to the good paths. PMM can immediately (after a single failure) fail over to another path, so it requires only two tries (one failed and one successful) to complete the I/O request, resulting in a significant time saving.

A failover operation proceeds as follows. PMM routes the I/O request to a selected path, if this path such as HBA-1 fails, the I/O request is returned to PMM. PMM marks the path as non-functional to prevent it from being selected for the subsequent I/O requests. It then determines whether to retry the I/O request. If retry is necessary, the request is rebuilt and transmitted to a different device path. In case that all paths to the device have failed and a I/O hold timeout occurs, there is no need to retry, a final device error is returned to the system and the I/O request is reported as uncomplete.

PMM probes the non-functional paths periodically by issuing a test unit ready (TUR) SCSI command over them. If the command succeeds, then PMM marks the path as functional, then the path can be selected to service I/O requests.

5. Experiment Results

To study the effects of load balancing and failover, extensive experiments are conducted. All the experiments were run on the same PC and Windows 2003.

3 test cases of path failure are designed to verify the effective failover policy of HBA-based multipath solution over disk-based one. The first is to remove LUN map entry, the second is to pull out the fibre cable between the host and storage device, the lost is to reset the RAID controller. The elapsed time that 500MB data is written to the LUN is recorded in normal condition and a test case, and then the time difference is failover time. Each test case is conducted 10 times and the average failover time is shown in Tab. 2.

Table 2. Average fail over time of failure conditions.

Test case	Disk-based	HBA-based
LUN map removal	4.45s	4.48s
Cable removal	52.61s	7.82s
Reset RAID controller	67.23s	38.26s

There is slight difference between these two solutions in case of LUN map removal, because the disk driver doesn't retry the I/O request on the failed path due to the RAID controller returns "ILLEGAL REQUEST" as the error code. The failover time is reduced distinctly by the HBA-based solution since it avoids the futile retries on the failed path by its counterpart. For the last test case, the RAID controller reset lasts about 30 seconds, so both paths are in active and the retries on them doesn't succeed. After the controller is ready, the HBA-based solution requires shorter time to reroute the I/O requests to any of the recovered paths.

6. Conclusions

A HBA-based multipath solution is developed on Windows operating system to eliminates inefficiencies of disk-based one, including path-dependant disk number, long failover time, by creating single-instance multipath disk device and avoiding futile retries on failed path. The cost is that HBA-based solution is complicated and difficult to port to other operating systems, since the HBA driver interface is boundled with OS kernel closely compared with the disk driver interface. This cost is worthy because the reducing of failover time can improves the system availability evidently [7].

7. REFERENCES

[1] S. Wilson, Managing a Fibre Channel Storage Area Net work, Storage Networking Industry Association, 1998.

[2] QLogic Corporation. QLogic Fibre Channel Software: Enhancing Disk Adapter Performance, Function, and Control, 2000.

[3] IBM Corp. Multi-path I/O for AIX 5L Version 5.2, 2002.

[4] VERITAS Software Corporation. VERITAS Dynamic Multipathing (DMP), 2001.

[5] EMC Corp. PowerPath Product Description Guide, 2002

[6] Y. A. Tan, Y. D. Cao,“Design and Implementation of a RDP Device Driver for Windows 2000”, Journal of Beijing Institute of Technology, 2003,12(2): 260-264.

[7] Y. A. Tan, Y. D. Cao, et al, “Virtual disks and fault tolerant paths in windows”, Transaction of Beijing Institute of Technology, 2005, 25(12): 1057-1061.

Method to Find Community Structures from Mobile Communication Data

Peihua Fu
College of Electrical Engineering, Zhejiang University
Hangzhou, 310029, P.R. China
College of Computer and Information Engineering, Zhejiang Gongshang University
Hangzhou, 310035, P.R. China
Email: fph@mail.zjgsu.edu.cn

ABSTRACT

Recently, there has been an increased amount of work on finding communities in network. In this paper, we present a method for identifying communities of practice from mobile communication bill within an organization. We use an improved edge-clustering algorithm that can rapidly find communities within a graph representing information flows. We apply this method to a batch of mobile communication data with 12251 recorders collected over a two-month span, and show that the method is effective at finding true communities. These studies are complemented by a qualitative evaluation of the results in the field.

Keywords: Data Mining, Complex Network, Mobil Communication.

1. INTRODUCTION

Many systems take the form of networks, sets of vertices joined together in pairs by edges. Examples include social networks such as acquaintance networks and collaboration networks, technological networks such as the Internet, the World Wide Web, and biological networks such as neural networks and metabolic networks. Recent research on networks has focused on a number of distinctive properties that most real networks seem to share. One such property is "community", a subsets of nodes highly linked among themselves but loosely connected to the rest of the network [1,2].

Finding community structure in a network could clearly have practical applications. Communities in a social network might represent real social groupings, perhaps by interest or background; communities in a citation network might represent related papers on a single topic; communities in the Web might represent pages on related topics. Being able to find these communities could help us to understand and exploit these networks more effectively [3].

In this paper, we propose a method for finding community structure within an organization. The method uses mobile communication bill as the indicator of interpersonal communication to construct a social network, and then finds the community structures by partitioning this network. When applied to a batch of mobile communication data, it gives promising results, which help us understand better the structure of the organization.

2. DEFINITION OF NETWORK COMMUNITY STRUCTURE

Many networks, it is found, are inhomogeneous, consisting not of an undifferentiated mass of vertices, but of distinct groups. Within these groups there are many edges between vertices, but between groups there are fewer edges, producing a structure like that sketched in Fig. 1. [3].

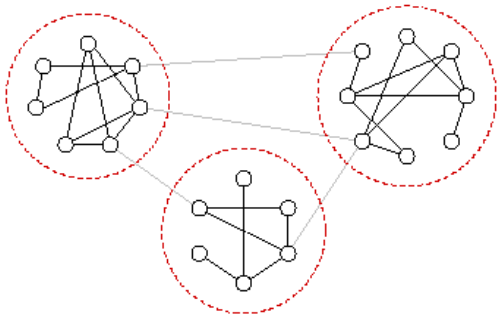


Fig. 1. In this network there are three communities of densely connected vertices (circles with solid lines), with a much lower density of connections (gray lines) between them.

Network can be denoted as graph $G=(V,E)$, where V denotes vertices (individuals in network); E denotes the edges join vertices (relation among vertices). $A_{i,j}$ is used to describe the adjacency matrix of G. For the i th vertex in G, $k_i=\sum_j A_j$ is the degree. Considering a sub-graph $W\subset G$, if $i\in W$, the degree of vertex i can be divided into two parts. One part is the edges, which join the vertices within W: $k_i^{in}(W)=\sum_{j\in W} A_{i,j}$. Another is the edge, which join the vertices outside W: $k_i^{out}(W)=\sum_{j\notin W} A_{i,j}$. We call W one of the communities in G, when Equation (1) holds. A larger α corresponds to more obvious community characteristic [4].

$$k_i^{in}(W)>\alpha k_i^{out}(W)\ \forall i\in W,\alpha\geq 1. \quad (1)$$

3. NETWORK PARTITIONING ALGORITHM BASED ON EDGE CLUSTERING COEFFICIENT

The problem of finding communities is not new and is closely related to the problem of graph partitioning there are many traditional methods such as Spectral Bisection and Hierarchical Clustering[5,6]. Spectral Bisection divides the network into two parts. When there are quite a number of communities, the results are not approving. Hierarchical Clustering cannot give the number of communities directly, and it is not suitable for a large-scale network.

In recent years, many algorithms for detecting communities have been proposed, starting with the seminal work by Girvan and Newman[7,8]. These authors proposed a method based on the progressive removal of links with the largest betweenness, a quantity proportional to the number of

shortest paths passing through a given edge. This method generates very good results. Unluckily, it has a main disadvantage: its computational demand is very high. For instance, for sparse networks with N vertices, the computation time grows like N^3.

Radicchi et al. introduced a local algorithm based on counting short loops of edges in the network, loops of length three, in the simplest case [4]. Consider an edge that runs between two vertices i and j having degrees k_i and k_j. The maximum number of triangles to which such an edge can belong, assuming that there is at most one edge between any pair of vertices, is min $[(k_i - 1), (k_j - 1)]$. These authors defined what they call the edge-clustering coefficient $\widetilde{C}_{i,j}^{(3)}$, which is roughly the fraction of these triangles that are actually realized:

$$\widetilde{C}_{i,j}^{(3)} = \frac{z_{i,j}^{(3)} + 1}{\min[(k_i - 1),(k_j - 1)]} \tag{2}$$

where $z_{i,j}^{(3)}$ is the measured number of triangles to which the edge belongs. The extra +1 in the numerator is included to avoid penalizing too heavily edges that belong to zero triangles, but which join vertices of low degree.

The edge-clustering coefficient $\tilde{C}_{i,j}^{(3)}$ is a measure of how inter-communitarian a link is. The edges among different communities have lower coefficient while the edges inside the communities have higher coefficient. In Fig. 2, two obvious communities in network are connected with edge AB. Edge AB gains a lowest coefficient:

$$\widetilde{C}_{A,B}^{(3)} = \frac{0+1}{\min[(4-1),(6-1)]} = \frac{1}{3} \tag{3}$$

for edge AC which inside a community, the coefficient is higher.

$$\widetilde{C}_{A,C}^{(3)} = \frac{2+1}{\min[(4-1),(3-1)]} = \frac{3}{2} \tag{4}$$

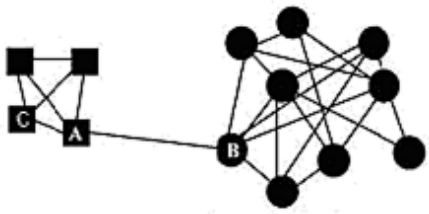


Fig. 2. This graph consists of two well-defined communities: the four vertices denoted by squares, including vertex *A*, and the nine denoted by circles, including vertex *B*. Edge *AB* has the lowest coefficient. If we were to remove it, the graph would split into two connected components.

This property can be used to distinguish whether the edges are inside the communities or inter-community. If we remove *AB*, the network will be divided into two independent communities.

The algorithm is simply stated as follows:

1) Calculate the edge-clustering coefficient for all edges in the network.
2) Remove the edge with the lowest coefficient.
3) Recalculate edge-clustering coefficient for all edges affected by the removal.
4) Repeat from step 2 until no edges remain.

The algorithm is very fast, since calculating the clustering coefficient can be done with local information only. Consider a network with N vertices and M edges, the computational time of the algorithm is expected to be linearly dependent on M for small systems and to cross over to an M^2 regime for large sizes.

4. METHOD FOR FINDING COMMUNITY STRUCTURE

We improve the former clustering algorithm, which based on edge clustering coefficient and propose a method for finding community structure within mobile communication data. This method is consists of four basic steps: collecting data, processing data, partitioning network and validating result.

4.1 Collecting Data

Data collection is the foundation of our method for finding community structure. At first, we select a network (eg. an organization), then ask each member in the network to provide mobile communication bill within a selected period. At present, China Mobile and China Unicom, two mobile communication companies in mainland of China, offer bill download service. The download file is the txt format. If members in network can cooperate actively, data collection is convenient.

4.2 Processing Data

At first we filter the raw data in accordance with the principle of protecting privacy. We delete all the communication record except receiver or sender in the network, at the same time; we encrypt the sender (caller's number) or receiver (receiver's number) of each record.

In order to analyze the relationship within members in network quantitatively, we define traffics for different mobile communication modes

The relationship of member i and j in network can be denoted by their total traffics F_{ij} :

$$F_{ij} = n_{sms} + 2n_{c1} + 4n_{c2} + 6n_{c3} \tag{5}$$

Where n_{sms} denotes the message number between member i and j, n_{c1}, n_{c2} and n_{c3} stand for the times of talk between member i,j less than 3 minutes, 3 to 10 minutes and more than 10 minutes respectively. Traffics for different mobile communication modes are in Table 1. A larger traffic corresponds to a closer relationship between i and j.

We should determine the threshold of traffic to create the adjacency matrix. If the traffic between two members is under the threshold, we say there is no relation between them, and there is no edge connecting the relevant vertices. It is very important to determine this threshold. If it is too low, most members have relationship, and it is not easy to find small community. If it is too high, most members have no relationship, and vertices are isolated. We determine the threshold F_H by data filter technology.

$$F_H = 2\beta\left(\sum_{i<j} F_{ij} - \sum_{\max 5\%} F_{ij} - \sum_{\min 5\%} F_{ij}\right) / n(n+1) \tag{6}$$

Where n is the number of member in network. $\sum_{\max 5\%} C_{ij}$ is the sum of traffic of the member pair whose traffic is on top 5%, $\sum_{\min 5\%} C_{ij}$ is the sum of traffic of the member pair whose traffic is on bottom 5%. β is used to adjust the threshold.

Table 1. Traffics for different mobile communication modes

Communication mode	Traffic
1 message(SMS)	1
A talk lasting less than 3 minutes	2
A talk lasting 3 to 10 minutes	4
A talk lasting more than 10 minutes	6

4.3 Partitioning Network

Network partitioning is the key step of finding community structure. As a clustering algorithm, edge clustering algorithm remove the edge with the lowest edge-clustering coefficient until no edge left. But it doesn't describe the community structure, and further analyzing is needed. To improve edge clustering, a condition should be determined to stop the edge-remove process automatically. And when it is true, the edge-remove process can divide the complex network into several communities with tight relationship. We set threshold C_H as the signal to stop removing edge process. When edge-clustering coefficient of all edges in network are larger than C_H, the edge-remove process stop.

$$C_H = \lambda(\sum_{i \neq j} \tilde{C}_{ij}^{(3)}) / m \tag{7}$$

Where m is the number of edges λ is used to adjust the threshold.

The algorithm we propose is simply stated as follows:

1) Calculate the edge-clustering coefficient for all edges in the network.

2) Remove the edge with the lowest edge-clustering coefficient.

3) Recalculate edge-clustering coefficient for all edges affected by the removal.

4) Repeat from step 2 until edge clustering coefficient of all edges in network are larger than C_H

4.4 Validating Result

In order to validate the results of the communities identified by our algorithm, we conducted a field study, consisting of personal interviews with individuals within network. In these interviews, we presented the subjects with the community in which they were placed by the algorithm, and then invited them to comment on these results. The main questions of interview include:

1) Do you think you belong to the community, which you are labeled?

2) Is there somebody redundant in your community?

3) Is there somebody absent in your community?

4) Please describe the meaning and characteristic of your community.

5) What is your evaluation of this algorithm as a whole

5. TEST OF THE METHOD

According to the method of finding community structure, the special software system is developed, which has functions of data gathering, data processing, network partitioning and result validating. The software can find community structure in network automatically. We test our algorithm and software system with a batch of mobile communication data collected over a two-month span.

In data preprocessing, we find some records of call duration is not correct, and are short of comparability. So the traffic of one call is set to 3. We get the threshold 22. If the traffic between two members is not lower than 22, we think the two members have relationship. A pair of vertexes in graph is connected with one edge.

The network is obtained with 33 vertexes and 48 edges after data processing, which is shown in Fig. 3.

Using our algorithm, the interpersonal network community structure was found after 18 edge-remove processes as shown in Fig. 4. The whole network has been divided into 5 communities: community 1{02，10，18，26}, community 2{09，12，17，19，24}, community 3{11，31} ，community 4{03，05，06，08，

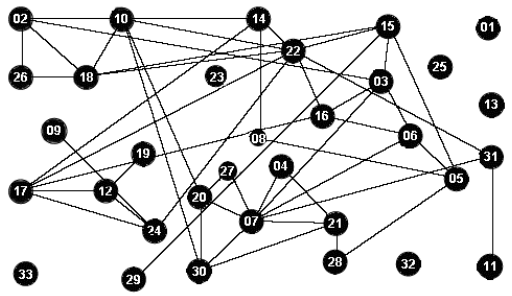


Fig. 3. The original interpersonal network with 33 vertexes and 48 edges

14，15，16，22}, community 5 {04，07，20，21，27，28，30}. In addition , there are 6 isolated vertexes which are labeled 01，13，23，25，32，33。

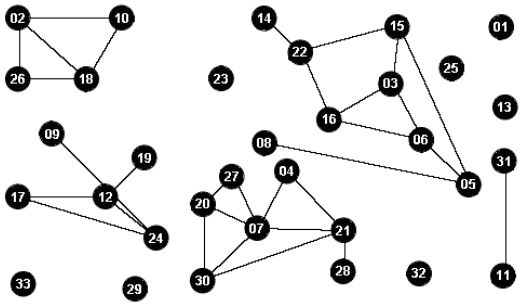


Fig. 4. The network after 18 edge-remove processes. It has been divided into 5 communities

32 students in the class were interviewed for validation. The results show: 93% students thought they belong to the community, which they were labeled. 8% students thought there was somebody redundant in their community. 6% students thought there was somebody absent in their community. The average score of the algorithm was 90.

Every interviewee described the meaning and characteristic of their community concretely. Some examples are as follows:

1) In community 1, all 4 students live in the same room.

2) In community 2, all 5 students have same interest and often discuss questions and take activities together.

3) In community 3, 2 students are good friends. In the period of data collecting, they were preparing for a important exam and had less contact with other students.

4) In community 4, 5 students in 8 are cadres of class; the other 3 care much about the class, and take an active part in

the class activities.

5) In community 5, all 6 students live in the same room.

From the above analysis we conclude that the network partitioning is according with the fact, and the method of finding community structure from mobile communication data has a high veracity.

6. FUTURE WORK

For convenience, our method is based on unweighted networks. This means we divide the interpersonal relation into two states: "related" and "un-related", and a lot of useful information is lost. Our future research will focus on weighted networks which can well conform to the fact. In addition, we hope to test our method with larger networks with huge amount data.

7. REFERENCES

[1] A.L. Barabasi, R. Albert. "Statistical mechanics of complex networks", Rev. Mod. Phys., 2002, 47–97.

[2] M. E. J. Newman. "The structure and function of complex networks", SIAM Review, 2003, pp.167–256.

[3] M. E. J. Newman. "Detecting community structure in networks". Euro.Phys. J. B, 2004. 321–330.

[4] F. Radicchi, C. Castellano, et al, "Defning and identifying communities in networks". Proc. Natl. Acad. Sci. USA 101, 2004,2658-2663.

[5] A. Pothen, H. Simon, et al, "Partitioning sparse matrices with eigenvectors of graphs", SIAM J. Matrix Anal. Appl. 11, 1990, 430-452.

[6] J. Scott, Social Network Analysis: A Handbook. Sage, London, 2nd edition, 2000.

[7] M. Girvan, M. Newman, "Community structure in social and biological networks", Proc. Natl. Acad. Sci. USA 99, 2002, 8271-8276.

[8] M. E. J. Newman. " Scientific collaboration networks. i. network construction and fundamental results", Phys. Rev. E, 64(016132), 2001.

A Conversion of DCT Coefficients to AVS Transform Coefficients

Xinchen Zhang
College of physical science and Technology, Central China Normal University
Wuhan, Hubei 430079, P.R. China
Email: zxc9501@163.com
Zhen Han, Qiong Liu, Zhongyuan Wang
The Key Laboratory of Multimedia and Network Communications Engineering, Wuhan University
Wuhan, Hubei 430079, P.R. China
Email: zhangxc@phy.ccnu.edu.cn

ABSTRACT

AVS, the latest video coding standard developed by AVS working group of China, uses an integer transform, which is different to Discrete Cosine Transform (DCT). We propose an effective method to convert DCT coefficients to HT coefficients entirely in the transform domain. According to the transform kernel matrix, we provide a fast algorithm and an integer approximation of the transform. The results show that the proposed transform outperforms the conventional pixel domain approach. It is expected to have applications in the MPEG-2 to AVS transform domain video transcoding.

Keywords: AVS, Video Transcoding, Transform Coefficients, Transform Domain, DCT.

1. INTRODUCTION

Video transcoding deals with converting a previously compressed video signal into another with different format, such as different bit rate, frame size, frame rate, or even compress standard [1]. It is one of key enabling technologies for multimedia interoperability.

MPEG-2 [2] is the most common video standard adopted in the multimedia industry. It has been already widely used in the field of digital broadcast, HDTV and DVD applications. But the MPEG-2 codec was developed about 10 years ago with the available algorithmic tools and hardware development constraints.

The Advanced Audio Video Coding Standard (AVS) [3] is established by Audio Video Standard Working Group of China, for digital television, digital videodisk and broadband network multimedia applications. It achieves higher coding efficiency by employing techniques. When compared to MPEG-2, the compression rate of AVS is 2-3 times of that of that of MPEG-2 [4]. So this new generation video standard, with its significant bandwidth saving, is expected to replace the use of MPEG-2 video compression in digital video systems.

The transcoding of MPEG-2 video to AVS format is particularly interesting given the wide availability and use of MPEG-2 video today. Because AVS and MPEG-2 have different transform kernels, MPEG-2 uses an 8x8 Discrete Cosine Transform (DCT) and AVS uses an 8x8 integer transform (IT), one important step in the transcoding is to convert the coefficients from the DCT domain to the integer transform (IT) domain.

To get good video quality and acceptable complexity, we designed a video transcoder based on Pixel-Domain Close-Loop architecture, which can perform the bit-rate reduction, frame skipping, frame size conversation, and error resilient conversation by changing several switches. It reduces the complexity by reusing the MV and MB mode and keeps the video quality by correcting the MV and the residuals. We implement this MPEG2-to-AVS transcoder on PCs with speed optimization using SIMD technology. On the multi-processor PC Server, it can satisfy the needs of the real-time video transcoding applications. But on normal PC, it cannot achieve the target. In order to reduce the complexity more, we want to implement the MPEG2-to-AVS transcoder in the transform domain. Avoiding the reconstructing the image pixels, the transform domain transcoding could be more efficient than the pixel domain transcoding because complete decoding and re-encoding are not require [5, 6]. Therefore, this is a need to have an efficient method to perform the DCT-to-IT conversion in the transform domain.

In this paper, we first propose a transform domain DCT-to-IT conversion. Section 3 discusses the fast algorithm and the integer approximation for the conversion respectively. In section 4, we give the simulation results. Section 5 makes a conclusion.

2. BACKGROUND

Our proposed conversion, which converting DCT coefficients to AVS integer transform coefficients in transform domain, is called Z-transform in this paper. The input of Z-transform (Y_{in}) is DCT coefficients of MPEG2 video streams. And the output (Y_{out}) is AVS Integer Transform coefficients. The conversion can be characterized in a matrix form:

$$Y_{out} = Z \times Y_{in} \times Z^T \qquad (1)$$

Where Z is the kernel matrix of Z-transform, and Z^T is the transpose of Z . We shall analyze the AVS-IT and DCT transform of MPEG2 to derive the matrix Z .

AVS-IT can be represented in the following form:

$$Y_{out} = T_8 \times X \times T_8^T \qquad (2)$$

Where

* This paper was supported by NSFC (NO. 60472040).

$$T_8 = \begin{pmatrix} 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ 10 & 9 & 6 & 2 & -2 & -6 & -9 & -10 \\ 10 & 4 & -4 & -10 & -10 & -4 & 4 & 10 \\ 9 & -2 & -10 & -6 & 6 & 10 & 2 & -9 \\ 8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 \\ 6 & -10 & 2 & 9 & -9 & -2 & 10 & -6 \\ 4 & -10 & 10 & -4 & -4 & 10 & -10 & 4 \\ 2 & -6 & 9 & -10 & 10 & -9 & 6 & -2 \end{pmatrix}$$

Denote the eight-point DCT matrix by S, with

$$S(i,j) = \frac{c(i)}{2}\cos\left(\frac{2j+1}{16}\cdot i\pi\right), 0 \le i, j \le 7 \quad (3)$$

and

$$c(i) = \begin{cases} 1/\sqrt{2}, i = 0 \\ 1, otherwise \end{cases} \quad (4)$$

Then we can write the operation of DCT in the following matrix form:

$$Y_{in} = S \times X \times S^T \quad (5)$$

According to (5), we have

$$X = S^T \times Y_{in} \times S \quad (6)$$

Eq.(2) can be rewritten as:

$$Y_{out} = T_8 \times S^T \times Y_{in} \times S \times T_8^T \quad (7)$$

Compared with (1),

$$Z = T_8 \times S^T \quad (8)$$

Therefore, we have the Z matrix:

$$Z = \begin{pmatrix} a' & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b' & 0 & -d' & 0 & -e' & 0 & -f' \\ 0 & 0 & c' & 0 & 0 & 0 & g' & 0 \\ 0 & e' & 0 & b' & 0 & f' & 0 & d' \\ 0 & 0 & 0 & 0 & a' & 0 & 0 & 0 \\ 0 & d' & 0 & -f' & 0 & b' & 0 & -e' \\ 0 & 0 & -g' & 0 & 0 & 0 & c' & 0 \\ 0 & f' & 0 & -e' & 0 & d' & 0 & b' \end{pmatrix} \quad (9)$$

Where

a' = 22.627
b' = 21.015
c' = 21.539
d' = 0.43697
e' = 0.43788
f' = 0.021982
g' = 0.26263

The coefficients in T_8, the kernel matrix of AVS-IT, have been shifted left 5 bits previously. Therefore, the values of a' through g' also have been scared up to 32 times. So divided by 32, the value of f is very close to 0, while d is very similar to the value of e. For that, we set f=0, and d=e to modify the matrix Z to reduce the unnecessary computations.

After modified, the Z matrix is:

$$Z = \begin{pmatrix} a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b & 0 & -d & 0 & -d & 0 & 0 \\ 0 & 0 & c & 0 & 0 & 0 & e & 0 \\ 0 & d & 0 & b & 0 & 0 & 0 & d \\ 0 & 0 & 0 & 0 & a & 0 & 0 & 0 \\ 0 & d & 0 & 0 & 0 & b & 0 & -d \\ 0 & 0 & -e & 0 & 0 & 0 & c & 0 \\ 0 & 0 & 0 & -d & 0 & d & 0 & b \end{pmatrix} \quad (10)$$

Where

a = 22.627
b = 21.015
c = 21.539
d = 0.43697
e = 0.26263

After that, we can convert the DCT coefficients to AVS coefficients directly by using the Z-transform.

3. FAST INTEGER TRANSFROM CONVERSION

3.1 Fast Conversion

Furthermore, as suggested by (1), the 2-D transform can be separated to 1-D transform, i.e. column transform after row transform or row transform after column transform.

In the following, we will take the second case for instance, to describe the 1-D transform.

Let y be an eight point column vector of, and Y be the 1-D transform of y.

$$Y[0] = a \times y[0]$$

$$Y[1] = b \times y[1] - d \times y[3] - d \times y[5]$$

$$Y[2] = c \times y[2] + e \times y[6]$$

$$Y[3] = d \times y[1] + b \times y[3] + d \times y[7]$$

$$Y[4] = a \times y[4]$$

$$Y[5] = d \times y[1] + b \times y[5] - d \times y[7]$$

$$Y[6] = -e \times y[2] + c \times y[6]$$

$$Y[7] = -d \times y[3] + d \times y[5] + b \times y[7]$$

The 1-D transform can be described as the Fig. 1.

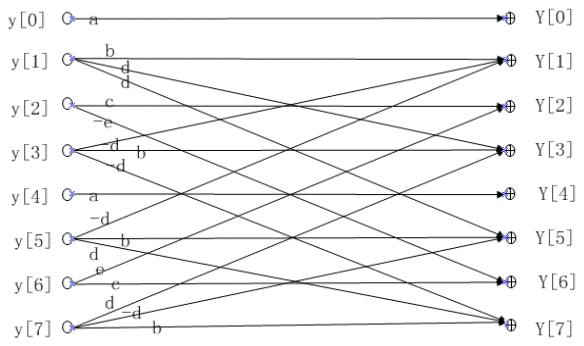


Fig. 1. Fast algorithm for the 1-D Z-transform

3.2 Integer Approximation of the Conversion

The parameters, a through e, are floating-point numbers. However, the floating-point operation is generally more complex than integer operation. Thus, we apply a method to convert the coefficients to integer numbers.

We multiply Z matrix by an integer, which is a power of two. So the coefficients are scaled up to integer in this step. And then, we use the integer matrix to perform the Z-transform. After that, the coefficients will be scaled down in the following steps to get the correct values.

Thus, the key problem is which integer we select to multiply. Theoretically, the larger it is, the better accuracy of results. However, the ability of CPU is limited. Choosing a too large multiplicator will cause overflow and wrong results. We should select a proper integer so that using 32-bits microprocessor, which is generally used, can perform computations.

Firstly, the DCT coefficients are in the range of -2048 to 2047, which needs 12 bits to represent. The maximum sum of absolute values of any row of Z is 22.627, so the maximum range of the Z-transform is $22.627^2 \approx 511.98 \approx 512 = 2^9$, that needs 9 bits to represent. Thus, 21 bits needed to represent the final transform results. So we still have 11 bits. The multiplicator have to smaller than the square root of $\sqrt{2^{11}} = 2^{5.5} = 45.25$. Therefore, the integer we select to multiply should be 32.

The kernel integer matrix is $IZ = round\{32 \times Z\}$, IZ has the same form of (10),but the parameters a through e is :

$a = 724$
$b = 672$
$c = 689$
$d = 14$
$e = 8$

4. EXPERIMENT RESULTS

In this section, we will compare the real Z-transform with the reference transform, and the integer transform with the reference. The reference transform is IDCT-IT (IDCT followed by IT). The experiment system is setting as Fig. 2 shows. The input x, an 8x8 block partitioned from the test sequence, is DCT-transformed, quantized (Q_1) and inverse-quantized (IQ_1) used in MPEG-2. The reconstructed DCT coefficients, X, are sent to three processing systems. Each of the three systems map X into the AVS IT domain, which are then reconstructed through quantization (Q_2), inverse-quantization (IQ_2), and inverse-AVS-transform (IT^{-1}). The DCT-to-IT conversation schemes used are: the reference IDCT-IT, the real-arithmetic Z-transform (Z), and the integer- arithmetic Z-transform (Z'). The pixel-blocks reconstructed form the three systems are denoted as x_r, x_z, x_z' respectively. We use Peak Signal-to-Noise Ratio (PSNR) to measure the distortion between the source and the reconstructed pixel-blocks.

In our experiment, the first 100 frames of three sequences (BASKETBALL, HORSERIDING, TEMPETE) with different types of motion were used. The size of the frame is 720×576 format. We take the average of three sequences. We use QP2 form 2 to 62 as AVS Q2 input parameter, which can include full application area. We use QP1 = 4, 8, 16, 20.

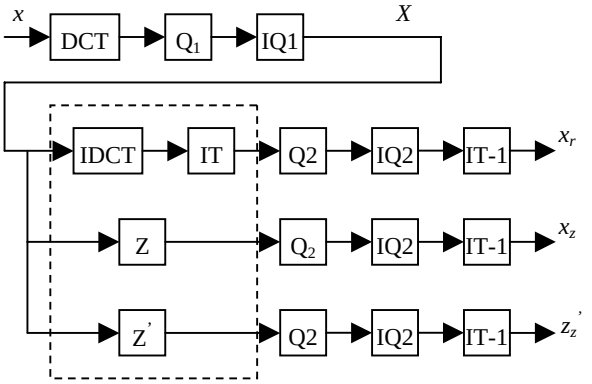


Fig. 2. Experiment setting

Fig. 3. shows the PSNR difference between using the real Z-transform and the reference. The real Z-transform outperforms the reference one. The improvement could be as much as 0.037dB for QP_1 =4 and QP_2 = 2. When QP_2 and QP_1 increases, the gain diminish as the quantization error dominates the distortion.

Fig. 4. shows the PSNR difference between using the integer Z-transform and the reference. The results show that the integer Z-transform is very close to the reference. The average distortion is about 0.005dB. The maximal distortion does not exceed 0.025dB. This distortion decreases as QP_2 increases.

According to the results, the proposed integer transform domain conversion can achieve better external picture quality and less complexity than the conventional IDCT-IT transform at both high and low bit-rate applications.

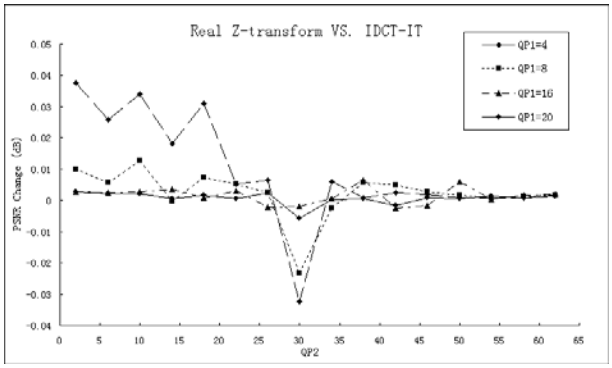


Fig. 3. PSNR different of real Z-transform vs. reference

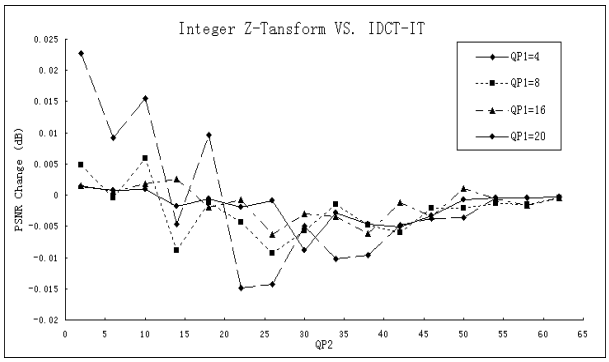


Fig. 4. PSNR different of integer Z-transform vs. reference

5. CONCLUSIONS

In this paper, we introduce a transform domain conversion of DCT coefficients to AVS IT coefficients. This conversion can be used in MPEG2-to-AVS transform domain transcoding. We derive the transformation kernel matrix and develop efficient algorithms for computing the transform. Also an integer approximation of the transform is proposed. In the experiment, the results show that the proposed integer *Z*-transform achieves equal (or better) PSNR performance to the conventional pixel domain implementation while requiring reducing computational complexity. Therefore, it can be applied to the transcoding of MPEG-2 (include other DCT-based video standards) to AVS video. Currently, we are developing the transcoder in transform domain.

6. REFERENCES

[1] V. Anthony, C. Charilaos, "Video Transcoding Architectures and Techniques: an overview", IEEE Signal Processing Magazine, pp. 18-29, Mar 2003.

[2] "Generic Coding of Moving Pictures and Associated Audio Information Part 2: Video", ITU-T and ISO/IEC JTC 1, ITU-T Recommendation H.262 and ISO/IEC 13818-2 (MPEG-2), 1994.

[3] AVS Video Expert Group, "Information Technology Advanced Audio Video Coding Standard Part 2: Video", in Audio Video Coding Standard Group of China (AVS), Doc.AVS-N1063, Dec. 2003.

[4] J. Wang, L Shi et al, "An AVS-to-MPEG2 Transcoding System", in Proc. Of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong, Oct. 2004,302-305.

[5] P. Assunção, M. Ghanbari, "A frequency-domain video transcoder for dynamic bit-rate reduction of MPEG-2 bitstreams", *IEEE* Trans Circuits Syst. Video Technol, Dec 1998,953-967.

[6] C.W. Lin, Y.R. Lee, "Fast algorithms for DCT-domain video transcoding", in Proc. *IEEE* Int. Conf. Image Processing, Thessaloniki, Greece, vol. 1, Sept. 2001,421-424.

QoS Multicast Routing Algorithm Based on Layered Structure*

CHEN Niansheng[1,2] , LI Layuan[2] , CHENG Chuanhui[2,3]
[1] Department of Computer Science, Hubei Normal University
Huangshi, Hubei, China
Email: hschenns@163.com
[2] School of Computer Science, Wuhan University of Technology
Wuhan,Hubei,China
[3] School of Information, Zhongnan University of Economics and Law
Wuhan,Hubei,China

ABSTRACT

With the rapidly expanding application of mobile Ad Hoc networks, the multicast routing technology with QoS constraint becomes an important research subject in the research field of networks and distributed systems. QoS multicast routing algorithm based on layered structure (QMALS) as a QoS Multicast Routing Algorithm is put forward in this paper. It deals with problems by cluster partition with weight coefficient based on the thought of multilevel hierarchy manage architecture. With QoS constraint QMALS can produce QoS Multicast tree with constraints of bandwidth, surplus energy, delay and delay jitter. Simulation experiments show that this algorithm can meet QoS restricted multicast requirements with good performance.

Keywords: Ad Hoc Networks; Cluster; Multicast Routing; QoS Routing Algorithm.

1. INTRODUCTION

Mobile Ad Hoc network (MANET), also called self-organized network, is a multi-hop temporary autonomy system organized by a group of mobile nodes with wireless receiving and sending devices. No center, self-organization, fast deployment, mobile and multi-hop are characteristic of MANET[1]. It can be used in any occasions where temporary communication networks needs building up, such as national defense combat readiness, rush to deal with an emergency, scientific research and exploration. MANET is applied in the form of free group communication with the support of the multicast technology. However, Ad Hoc networks also have some negative characteristics such as unconstraint network topology architecture, limited bandwidth and batteries energy of mobile network itself. Therefore, how to design an efficiency and fast multicast routing is a key technique in multicast applications. At present there are different routing protocols put forward by many researchers such as AODV[2]、ZRP[3]、DSR[4]、TBP[5] aiming at Ad Hoc mobile networks. However, none of them supports multicast technology except AODV protocol.

With the development of network technology, multicast applications have to offer not only best effort services but some applications with several QoS constraints. Thus how to implement multicast routing technology in Ad Hoc networks with QoS constraints simultaneously becomes an important project in network field[6,7,8]. For example, in Ref [7] the definition of long-life links avoids the dynamic change of network topology caused by the nodes movement and puts forward a QoS routing protocol satisfying both the delay and the bandwidth. But it doesn't offer any method or approach to support multicast. The contribution in Ref. [8] is QoS multicast routing protocol (QMRP) with constraints of bandwidth, delay, moving speed of node and the surplus electrical energy of node. However, impreciseness still exists in the definition of feasible path, besides only virtually considering the addable delay constraint.

This paper points out a QoS routing algorithm named QoS Multicast Algorithm Based on Layered Structure(QMALS) which support multicast communication according to the characteristics and requirements of mobile Ad Hoc networks application requirements. QMALS is set up by the idea that Ad Hoc mobile networks use multilevel hierarchy manage architecture. Firstly, QMALS do layered procedure of mobile Ad Hoc networks by cluster partition with weight coefficient. Then virtual backbone network only with key nodes is formed. At last, based on this virtual backbone network, ideal multicast tree is produced by some related procedures. The multicast tree can not only meet the requirements of multicast technology but also achieve QoS constraints of bandwidth, energy, delay and delay jitter. We say that it meets the multicast techno-requirements with QoS constraints.

2. DESCRIPTION OF NETWORK MODEL AND ROUTING PROBLEM

2.1 Network Model

In this paper, we assume some presuppositions: (1)Every mobile node in the Ad Hoc networks has an unique mark and is signed with GPS；(2)Since the available launch distance between every two nodes is the same, if two nodes are located within their mutual launch arrange, they are called Adjacent and connected by a link；(3)Adjacent nodes are marked by BEACON messages and the link situation can be known[6];(4)Only the node delay, delay jitter are discussed this paper because of the equivalence of QoS constraint for delay and delay jitter between node and link. Therefore a weighted Graph $G=(V,E)$ can depict the mobile Ad Hoc networks. V is a network node set, and E is a set of double links whose nodes communicate mutually with each other. If R_+ means positive real number set, R^+ means non-negative real number set, then for any network node $n \in V$ the QoS eigenvalue is: delay function $delay(n): V \to R_+$, delay jitter function $delay\_jitter(n): V \to R^+$, surplus energy restricted function $energy(n): V \to R_+$;for any link $e \in E$,

* Supported by the National Natural Science Foundation of China under Grant No.90304018; The Grand Research Project of Hubei Province Department of Education in China under Grant NO. D200622003

we can define bandwidth function $bandwidth(e): E \to R_+$ besides delay function and delay jitter function.

2.2 The problem of QoS multicast routing with multiple QoS constraints

Definition 1 In an existing Ad Hoc network $G=(V,E)$ in order, to calculate a multicast routing with QoS constraints existing between a resource node and multi destination nodes, we have to define $s \in V$ as multicast resource node, $M \subseteq \{V-\{s\}\}$ as a non-empty set of multicast destination nodes, and the multicast tree $T(s,M)$ is made of "*s*" and "*M*".

Definition 2 In an existing Ad Hoc network $G=(V,E)$, multicast tree $T(s,M)$ has some relations as follows:

(1) $delay(P_T(s,t)) = \sum_{n \in P_T(s,t)} delay(n)$.

(2) $delay\_jitter(P_T(s,t)) = \sum_{n \in P_T(s,t)} delay\_jitter(n)$.

(3) $bandwidth(P_T(s,t)) = \min\{bandwidth(e), e \in P_T(s,t)\}$.

(4) $energy(P_T(s,t)) = \min\{energy(n), n \in P_T(s,t)\}$.

$P_T(s,T)$ is the routing path between source node "*s*" and destination node "*t*" of multicast tree $T(s,M)$.

Definition 3 Supposing B_{min} is the lowest bandwidth constraint of multicast tree and E_{min} is the node surplus energy constraint, a question for some multicast request such as bandwidth, energy, delay, delay jitter constraints can be described in following way: to find a multicast tree $T(s,M)$ in network $G=(V,E)$ which can meet $bandwidth(P_T(s,t)) \geq B_{\min} \wedge energy(P_T(s,t)) \geq E_{\min}$ and at the same time to make the value of $delay(P_T(s,t))$ and $delay\_jitter(P_T(s,t))$ as little as possible.

2.3 Hierarchy of Networks

Owing to the movable characteristics of nodes in the mobile Ad Hoc networks and network information spending of network management, Ad Hoc networks normally adopt a multilevel hierarchy manage architecture[9] to clustering manage nodes in the network. Every node belongs to one type of cluster node, cluster bridge node and cluster inside node[6]. Since the Ad Hoc networks is divided into clusters and all the nodes are classified into these different cluster, a backbone link can be formed of cluster nodes, cluster bridge nodes and the links between them. Seeing it in Fig.1:

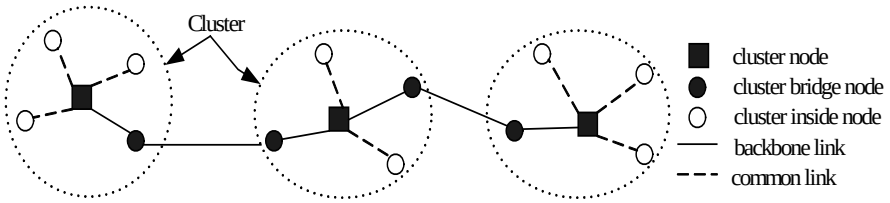


Fig. 1. Hierarchy of Networks

3. DESCRIPTION AND REALIZATION OF QMALS

QMALS in this paper consists of three parts: algorithm of cluster partition, looping test algorithm and redundant pruning algorithm. Firstly, the algorithm of cluster partition is applied to Ad Hoc mobile networks. Then the virtual backbone network is formed by cluster nodes, cluster bridge nodes and multicast member nodes. At lastly needed multicast tree can be created by dealing with the virtual backbone network with looping test algorithm and redundant pruning algorithm.

3.1 Algorithm of Cluster Partition

1) Initialization

Definition 4 In an existing Ad Hoc network $G=(V,E)$, for any node $n \in V$, the weight coefficient is defined as follows:

$$w = \alpha \cdot energy(n) + \beta \cdot (D_{\max} - delay(n))/D_{\max} + \gamma \cdot (J_{\max} - delay\_jitter(n))/J_{\max}$$

α, β, γ are weight coefficients of node surplus energy, delay and delay jitter, and $D_{\max}$, $J_{\max}$ are the max value of acceptable delay value and delay jitter value in one multicast request.

Considering the multicast requirements in definition 3, before the QMALS is applied, the initialization for idiographic network model is needed: ①Getting rid of links which can not meet bandwidth constraint is necessary. If the multicast resource node is not in the same graph with a destination node, negotiation should be made to reduce bandwidth constraint; ②A impermissible mark (*no*=1) must be signed on the node which can not meet the energy constraint, and that is $energy(n) < E_{\min}$; ③Every node of network $G=(V,E)$ should be given unique *ip* number, and a mark(*id*=1) should be signed on the node which belong to multicast node set; ④The weight coefficient *w* of each node is calculated according to definition 4. After this initialization procedure, every node in the network $G=(V,E)$ can be denoted with four-element-group (*ip*,*id*,*w*,*no*).

Definition 5 In an existing Ad Hoc network $G=(V,E)$, for any network node $i \in V$, $N(i)$ is defined as a set of adjacent, $NBList(i)$ as a set of adjacent property and *NBTable* as adjacent table for the whole network.

At the beginning of the algorithm of cluster partition, every node in the network send *Hello* message to its adjacent periodically. *Hello* message includes node's *ip* number、*id* and node weight coefficient *w*. Every node can form self adjacent property set *NbList* by receiving *Hello* message of its adjacent. For any node $i \in V$, the adjacent property set can be denoted as:

$$NBList(i) = \bigcup_{j \in N(i)} (ip_j, id_j, w_j, no_j)$$

All the adjacent property sets of all nodes form adjacent table of the whole network *NBTable*.

$$NBTable = \bigcup_{i \in V} NBList(i).$$

2) Cluster Partition

Cluster is formed based on the principle that when the weight coefficient of the node *w* is the largest and

$no \neq 1$,the node has the priority of becoming the cluster head node. According to *NBTable*, a cluster head node will be the node with largest *w* and $no \neq 1$ in the *NbList* which is the node self property starting from the node *ip*=1. The other nodes are cluster inside nodes. To every cluster, form cluster member node set *CNList*. To the nodes which have joined a cluster, related records in *NBTable* are deleted.

If the value *w* of several nodes inside a node adjacent property set are equal, the node in which *id*=1and $no \neq 1$ is chosen as the cluster head node. If the value *w* of several nodes are equal, and $id = 1 \wedge no \neq 1$, the node in which *ip* is the lowest is cluster head node. By doing this, all of the nodes can become member nodes of a cluster.

3) Marking Cluster Bridge Node

Definition 6 For a partitioned cluster, the cluster member node set is CNList; for any node *i* which belongs to this cluster, $N(i)$ is the set of adjacent, if $N(i) \subseteq CNList$, *i* is cluster inside node or it is the cluster bridge node.

After cluster partition, all of the cluster bridge nodes can be marked according to definition 6 to form virtual backbone network.

3.2 The Formation of Multicast Tree

Definition 7 For an existing Ad Hoc network $G = (V, E)$, after the cluster partition only cluster head node, cluster bridge node and cluster inside node that *id*=1 and links among them are reserved to form a new network $G' = (V', E')$. G' is called virtual backbone network.

The G' that is formed according to definition 7 is a foundation to the formation of multicast tree.

1) Looping Test Algorithm

The virtual backbone network G' mentioned above which is a foundation to the formation of multicast tree, is not a tree structure because there are loops in it. So it is necessary to deal with the question of loops. The looping test algorithm is:

From the multicast resource node a testing message *Test*(*ip*,*i*) is sent. *i* is the number of testing message, and *ip* is the node number which transfers the message. The message is transferred by multicast to the bigger *ip* number adjacent node between each cluster node and cluster bridge node. If any nodes receive the message twice, the following rules will be carried out.

①If the nodes which send the message don't have cluster inside member nodes, reserve the link to the nodes which send message with smaller *ip* number and cut other links.

②If the nodes which send the message include some cluster inside member nodes, reserve the link to the cluster inside member nodes with smaller *ip* number and cut other links.

After application of the algorithm mentioned above to the looping test, the virtual backbone network G' can form a multicast tree with tree structure.

2) Redundant Pruning Algorithm

Through the looping test the multicast tree also has some redundant nodes that are leaf nodes of non-multicast destination nodes which needs handing by the method of pruning. The process of redundant pruning algorithm is: Making use of the postorder traversing algorithm to traverse the above multicast tree in which the multicast resource node is defined as root node. When the nodes are visited, if the node and all of its son nodes are non-multicast nodes, the node should be deleted. The algorithm will not stop operating until all the leaf nodes in the multicast tree belong to the multicast nodes set *M*.

After the redundant pruning algorithm is operated, a multicast tree $T(s, M)$ which consist of multicast resource node *s* and multicast destination nodes set $M \subseteq \{V - \{s\}\}$ can be formed. Since weight coefficient method is used in cluster partition, this multicast tree has QoS constraints of bandwidth, energy, delay and delay jitter and meets multicast technology requirements with QoS constraint.

4. EXPERIMENT SIMULATION AND PERFORMANCE ANALYSIS

4.1 Experiment Environment

Simulation experiment environment is a rectangle area of 2200×1000m^2 and the launch arrange for each node is 80m. If two nodes are located within their mutual launch arrange, there exists a link between them; the bandwidth value of link is formed by (1-6)Mbps random number. The delay value and delay jitter value for each node are equally distributing on [3ms,10ms], and the initial energy value is 5J(joule); Each node chooses its direction and speed of motion randomly and the max speed is 30m/s.

Before the simulation starts, 100 nodes with ordinal number (*ip*) which are produced randomly come into the experiment environment and multicast members are appointed (one of them is the resource node and it sends data to other multicast members). During the process of simulation, multicast members do not change. Every node sends a *Hello* message every t=5s; each group of experiment can produce 500 requests, and the simulation environment maintains 600s. The result of simulation experiments is the average value of the results of many experiments.

4.2 Contrast Analysis of Results

In the simulation experiments, multicast performance analysis of QMALS should be done for different account of multicast members, and the performance parameters are:

①The success rate of sending data package: the data package received by the destination nodes to the data package sent by the resource node.

②Waiting time for transferring: he time needed to form multicast tree during the multicast request period.

The success rate of sending data package mainly reflects the reliability of the algorithm during implement process. Fig.2 shows the success rate varies with the change of node moving speed. When the number of multicast members is small, the success rate is high while the rate falls when the number of multicast members rises with the limited bandwidth. Even if the number of multicast members comes to 60, the success rate is higher than 50%. The graph shows that the change of node motion speed has little effects on the success rate of sending data package.

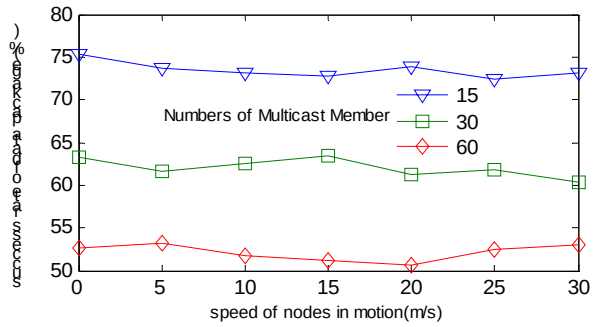


Fig. 2. Success rate of sending data package

Waiting time for transferring includes the spending time of algorithm of cluster partition, looping test algorithm and redundant pruning algorithm, especially the time spend by the algorithm of cluster partition. Fig.3 describes the change of the number of multicast member nodes and Transfer waiting time. It shows that when there are few multicast members, cluster partition should also be done, thus more time is needed. However, with the increase of multicast members, only dynamic rebuilt of cluster should be done, thus little time is needed. When the percentage of the multicast member nodes to all nodes is above 35%, waiting time for transferring has a stable value that is about 40ms.

To estimate and verify that QMALS supports QoS, the paper compares the QMALS with the LBMR algorithm mentioned by [7] and the Flooding algorithm mentioned by [10] on the aspect of routing success rate and multicast tree average transfer delay. Routing success rate is the rate of successful connection time to the whole request connection time; transfer delay is the multicast data transfer delay from resource node to every destination node.

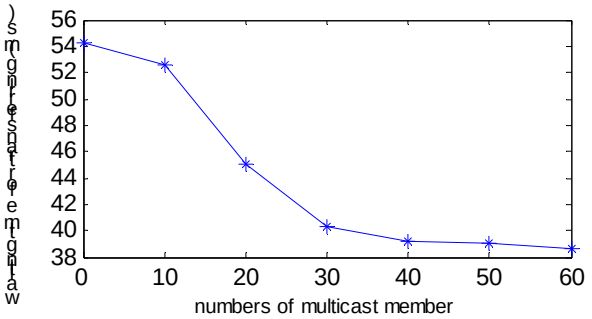


Fig. 3. Waiting time for transferring

In the experiments, the energy consumption of nodes adopt following node energy consumption calculation formula mentioned to by [11], which is shown in (1). It depends on the power of receiving or sending data and time needed to transact data package. P_{tx} means sending power and P_{rx} means receiving power.

$$E = (P_{tx} + P_{rx}) \times (8 \times PacketSize / Bandwidth) \quad (1)$$

When there are 15 multicast member nodes, not considering surplus energy constraint, the routing success rate of these three algorithms with different bandwidth constraint can be shown in Fig.4. Fig.5 shows the routing success rate simulation result of these three algorithms with different surplus energy constraint when there are 15 multicast member nodes and bandwidth constraint is 2Mbps.

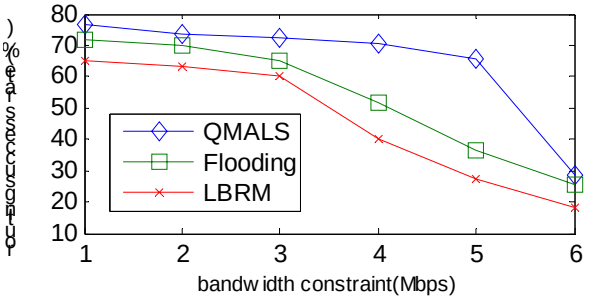


Fig. 4. Routing success rate based on bandwidth constraint

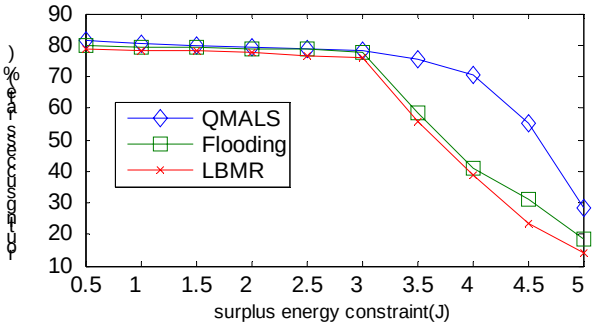


Fig. 5. Routing success rate based on energy constraint

Graphs show that with different bandwidth constraint, the routing success rate of QMALS is higher than Flooding algorithm and LBMR algorithm. This is because QMALS, which deals with bandwidth constraint at, first directly promotes the routing success rate of different bandwidth constraints. When it comes to different surplus energy constraint, results show that when the surplus energy constraint is not high, there is little difference among the routing success rate of these three algorithms. When the surplus energy constraint becomes higher (every node surplus energy value is more than 3J), the routing success rate of QMALS is obviously higher than others. Experiments show that surplus energy is considered as node weight key coefficient and improving the adaptability for different surplus energy constraint can promote the multicast routing success rate.

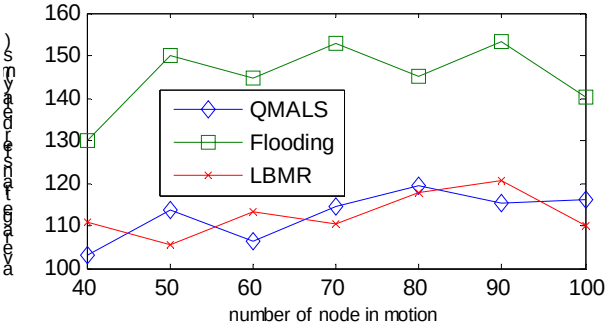


Fig. 6. Comparison of average transfer delay

Fig.6 shows the comparison of the result of the average multicast transfer delay value in different network scale of these three algorithms when there are 15 multicast member nodes. Though QMALS does not define delay constraint as necessary precondition to produce multicast tree, it has the same effect as the LBMR, which uses delay constraint, and the delay value is far higher than that of Flooding. The reason is QMALS uses delay and delay jitter as weight coefficients in node weight calculation process to form multicast tree with little transfer delay. Since delay jitter has the same characteristic with delay, there are no special experiments aiming at delay jitter.

5. CONCLUSION

This paper puts forward the QMALS algorithm based on the thought of multilevel hierarchy manage architecture which defines the weight of nodes by nodes' surplus energy, delay and delay jitter and produces multicast tree by cluster partition with weight coefficient to meet the multicast QoS constraints. Simulation experiments show that QMALS algorithm can meet some of the QoS constraints with good performance. However, we still need to do further research on how to produce the lowest cost multicast tree while the algorithm should also meet the constraints such as bandwidth, surplus energy, delay and delay jitter.

REFERENCES

[1] N.S. Chen,L.Y. Li, et al, " Research on QoS routing based on MANET", Computer Engineering and Applications, 2004, 40(30):120-123.
[2] C.R. Perkins, "Ad_hoc on demand distance vector routing",In Proceedings WMCSA'99-Scond IEEE Workstop on Mobile Computing Systems and Applications, 1999,90-100.
[3] Z.P. Hass, "The pertormance of query control scheme for the zone routing protocol",ACM SIGCOMM'98, Vancouver,Canoda,1998,167-177.
[4] J.M. Johnson, "Dynamic source routing in ad hoc wireless networks.:Mobile Computing",Kluwer Academic 1996,153-181.
[5] S.G. Chen, N. Klara, "Distributed quality-of-service routing in Ad hoc networks", IEEE Journal on Selected Areas in Communications, 1999, 17(8): 561-571.
[6] N.S. Chen, L.Y. Li, et al, "Multicast routing algorithm of Ad Hoc networks based on layered structure",Computer Engineering, 2005, 25(16):110-112.
[7] J. Shi, L. Zou, "A QoS based distributed multicast routing algorithm in Ad Hoc networks", Journal of China Institute of Communications, 2003, 24(6): 60-68.
[8]B.L. Sun, L.Y. Li, "A QoS based multicast routing protocol in Ad Hoc networks", Chinese Journal OF Computers, 2004,27(10):1402-1407.
[9]W. Chen, N. Jain, et al, "ANMP:Ad hoc networks management protocol", IEEE Journal on Selected Areas in Communications, 1999,17(9): 1506~1531.
[10] C.K. Toh, "Assiciatibity_based routing for ad hoc mobile networks[J]", Wireless Personal Communications 1997, 4(2): 103-109.
[11]M.Yuan,L.F. Zhang, et al, "An analysis of energy consumption for Ad Hoc networks routing protocols", Computer Engineering and Applications, 2003, 39(15):146-149.

Performance Comparison of Genetic Algorithm and Particle Swarm Optimization on QoS Multicast Routing Problem

Jie Qin, Jing Liu, Wenbo Xu
School of Information Technology, Southern Yangtze University
No. 1800, Lihudadao Road, Wuxi 214122, Jiangsu, China
E-mail: qinjie173@163.com

Abstract:

QoS multicast routing in networks is a very important research issues in the areas of networks and distributed systems. This is also a challenging and hard problem for the next generation Internet a high-performance networks. Because of its NP-completeness, many heuristics such as Genetic Algorithms are employed solve the problem. This paper proposes an Improved PSO (IPSO) and present how to solve QoS multicast routing problem by IPSO. The proposed algorithm converts the QoS multicast routing problem into an integer-programming problem. We test IPSO-base routing algorithm on a network model. For performance comparison, we also test Particle Swarm Optimization (PSO) algorithm and GA. The experiment results show the availability and efficiency of IPSO on the problem and its superiority to PSO and GA.

Keywords: Multicast routing, QoS, routing optimization, Particle swarm optimization algorithms, PSO.

1 Introduction

Multicast routing including QoS (Quality-of-Service) multicast routing has continued to be a very important research issue in the areas of networks and distributed systems [1, 3, 5, 7]. It has been demonstrated that it is NP-Complete to find a feasible multicast tree with two independent additive path constraints. The QoS requirements can be classified into link constraints (e.g., bandwidth), path constraints (e.g., end to end delay) and tree constraints (e.g., delay-jitter). Generally, heuristics are employed to solve this NP-complete problem. Some GAs has been used to solve the problem form different aspects. GA reassures a higher chance of reaching a global optimum by starting with multiple random search points and considering several candidate solutions simultaneously. Particle Swarm Optimization (PSO) method is one of its member. It was originally proposed by J. Kennedy as a simulation of social behavior of bird flock, and it was initially introduced as an optimization method in 1995 ([4]). PSO has been proved to be an efficient method for many global optimization problems and in some cases it does not suffer the difficulties encountered by GA.

PSO was originally proposed to solve the continuous functions optimization. So far, it has rarely been used to discrete Combinatory Optimization Problems (COP). The goal of this paper is to solving QoS multicast routing problem via PSO. The rest of the paper is organized as follows. In Section 2, the network model of QoS multicast routing problem is introduced. The concept of PSO is described in Section 3. A revised PSO is introduced to solve the QoS routing problem. The experiment results are given in Section 5 and the paper is concluded in Section 6.

2 Network Model

A network is usually represented as a weighted digraph $G = (V, E)$, where V denotes the set of nodes and E denotes the set of communication links connecting the nodes. |V| and |E| denote the number of nodes and links in the network, respectively, without loss of generality, only digraphs are considered in which there exists at most one link between a pair of ordered nodes [13].

Let $s \in V$ be source node of a multicast tree, and $M \subseteq \{V - \{s\}\}$ be a set of end nodes of the multicast tree. Let R be the positive weight and R+ be the nonnegative weight. For any link $e \in |E|$, we can define the some QoS metrics: delay function delay (e): $E \to R$, cost function cost (e): $E \to R$, bandwidth function bandwidth (e): $E \to R$; and delay jitter function delay-jitter (e): $E \to R^+$. Similarly, for any node $n \in V$, one can also define some metrics: delay function delay (n): $V \to R$, cost function cost (n): $V \to R$, delay jitter function delay-jitter (n): $V \to R^+$ and packet loss function packet-loss (n): $V \to R^+$. We also use $T(s, M)$ to denote a multicast tree, which has the following relations:

$$delay\left(p\left(s,T\right)\right) = \sum_{e \in p\left(s,T\right)} delay\left(e\right) + \sum_{n \in p\left(s,T\right)} delay\left(n\right)$$

$$\cos t\left(T\left(s,M\right)\right) = \sum_{e \in p\left(s,M\right)} \cos t\left(e\right) + \sum_{n \in p\left(s,M\right)} \cos t\left(n\right)$$

$$bandwidth\left(p\left(s,T\right)\right) = min\left(bandwidth\left(e\right)\right), e \in p\left(s,T\right)$$

$$delay-jitter\left(p\left(s,T\right)\right) = \sum_{e \in p\left(s,T\right)} delay-jitter\left(e\right) + \sum_{n \in p\left(s,T\right)} delay-jitter\left(n\right)$$

$$packet-loss\left(p\left(s,T\right)\right) = 1 - \prod_{n \in p\left(s,T\right)} \left(1 - packet-loss\left(n\right)\right)$$

Where $p\left(s,T\right)$ denotes the path from source s to end node t to T(s, M). With QoS requirements, the problem can be represented as finding a multicast tree T(s, M) satisfying the following constraints

1. Delay Constraint: $delay(p(s,T)) \le D$;
2. Bandwidth Constraint: $bandwidth(p(s,T)) \ge B$;
3. Delay-jitter Constraint: $delay-jitter(p(s,T)) \le J$;
4. Packet-loss Constraint: $packet-loss(p(s,T)) \le L$;

QoS multicast routing problem is a NP-complete hard problem, which is also a challenging problem for high-performance networks.

3 Particle Swarm Optimization

Particle Swarm Optimization (PSO), originally proposed by J. Kennedy and R. Eberhart [2, 4, 6], has become a most fascinating branch of evolutionary computation. The underlying motivation for the development of PSO algorithm was social behavior of animals such as bird flocking, fish schooling, and swarm theory. Like genetic algorithm (GA), PSO is a population-based random search technique but that outperforms GA in many practical applications, particularly in nonlinear optimization problems. In the Standard PSO model, each individual is treated as a volume-less particle in the D-dimensional space, with the position and velocity of *i*th particle represented as $X_i = (X_{i1}, X_{i2}, \cdots, X_{iD})$ and $V_i = (V_{i1}, V_{i2}, \cdots, V_{iD})$. The particles move according to the following equation:

$$V_{id} = w \cdot V_{id} + c_1 \cdot rand(\cdot) \cdot (P_{id} - X_{id}) + c_2 \cdot Rand(\cdot) \cdot (P_g - X_{id})$$

$$X_{id} = X_{id} + V_{id}$$

Where c_1 and c_2 are positive constant and rand () and Rand () are two random functions within [0,1]. Parameter *w* is the inertia weight introduced to accelerate the convergence speed of the PSO. Vector $P_i = (P_{i1}, P_{i2}, L, P_{iD})$ is the best previous position (the position giving the best fitness value) of particle *i* called ***pbest***, and vector $P_g = (P_{g1}, P_{g2}, L, P_{gD})$ is the position of the best particle among all the particles in the population and called ***gbest***.

4 The Proposed Algorithm

1) The Improved PSO

As we know, the original PSO is not a global convergence guaranteed algorithm and may encounter premature convergence. Some researchers test on benchmark functions by several widely known heuristics such as Genetic Algorithms, Simulated Annealing Algorithms, Particle Swarm Optimization and Differential Evolution algorithm and made a conclude that PSO may find out the optima of easy problems more efficiently than other algorithms, but its performance is inferior to others on complex problem due to its poor global search ability.

In this paper, we proposed an improved PSO with mutation operator on the velocity of each particle. The mutation operation is exerted on the particle when its velocity component on each dimension is less than a threshold value. The procedure of the operation is described as follows.

$$V_{id} = V_{id} + \sigma \cdot N(0,1) \quad if \quad V_{id} < \varepsilon$$

where N (0,1) is a random number with standard normal distribution.

2) Coding

The coding involves encoding a path serial into a feasible solution (or a position) in search space of the particle. In our proposed method, we design a new integral coding scheme for PSO and Improved PSO (IPSO) to be employed to solve the discrete combinatory optimization problem. In the coding scheme, the numbers of paths (no loop) reaching each end node $t \in M$ are worked out first. With the number of end nodes denoted by $|M|$, the number of paths to end node *i* is represented as an integer $n_i (1 \le i \le |M|)$. Thus each path to end node *i* can be numbered by an integer variable $t_i (1 \le i \le |M|)$, where $t_i \in [1, n_i] (1 \le i \le |M|)$. Consequently, we can obtain a $|M|$-dimensional integral vector $(t_1, t_2, \cdots, t_{|M|})$ denoting a possible path serial with each component t_i varying with $[1, n_i]$. In the IPSO for QoS Multicasting routing problem, such an integral vector represents the position of the particle and the combinatory optimization problem is reduced to a $|M|$-dimensional integral programming.

The initial population is a matrix with row vectors representing particles' positions. The dimension of a row vector is the number of end nodes $|M|$. The value of the *i*th component of a certain row vector represents the path number of a path from the source node to end node *i*. It can be initialized by randomly select an integer number in the interval $[1, n_i]$.

3) Fitness Function

In the proposed algorithm, the fitness unction is defined as:

$$f(x) = \frac{\omega_1}{\cos t(T(s,M))}\left(\omega_2 \cdot f(d) + \omega_3 \cdot f(j) + \omega_4 \cdot f(p)\right)$$

Where ω_1, ω_2, ω_3 and ω_4 is the weight of cost, delay, delay-jitter and packet loss, respectively; f(d), f(j) and f(p) are defined as

$$f(d) = \prod_{t \in M} F_d\left(delay(p(s,t)) - D\right),$$

$$F_d\left(delay(p(s,t)) - D\right) = \begin{cases} 1, delay(p(s,t)) < D \\ \alpha, delay(p(s,t)) \ge D \end{cases}$$

$$f(j) = \prod_{t \in M} F_j\left(delay\_jitter(p(s,t)) - J\right),$$

$$F_j\left(delay\_jitter(p(s,t)) - J\right) = \begin{cases} 1, delay\_jitter(p(s,t)) < J \\ \beta, delay\_jitter(p(s,t)) \ge J \end{cases}$$

$$f(p) = \prod_{t \in M} F_p\left(packet\_loss(p(s,t)) - L\right),$$

$$F_p\left(packet\_loss(p(s,t)) - L\right) = \begin{cases} 1, packet\_loss(p(s,t)) < L \\ \sigma, packet\_loss(p(s,t)) \ge L \end{cases}$$

where F_d (x), F_j (x) and F_p (x) are penalty functions for delay, delay-jitter and packet loss, respectively, and α, β and σ are positive numbers smaller than 1.

4) IPSO on QoS Routing Problem

IPSO-based QoS Multicast Routing Algorithm

Input: The dimension of the particles' positions (equal to the number of end nodes); Population size; Parameters of the network model.

Output: The best fitness value after IPSO executes for MAXITER iterations; optimal multicast tree.

Procedure:

1. *Initialize the population;*
2. ***for*** *t=1 to MAXITER*
3. *Compute the fitness value of each particle according to (17);*
4. *Update the personal best position* P_i *;*
5. *Update the global best position* P_g *;*
6. *Compute the mbest of the population by (8) and the value of . by (24);*
7. ***for*** *each particle in the population*

8. Update each component of the particle's position by (9) and (10) and adjust the component t_i *as an integer in* $[1,n_i]$*;*

9. ***endfor***

10.endfor

5) Loop-deletion Operation

Implementation of IPSO on the problem yields an optimal multicast tree denoted by a path serial. The path serial is a $|M|$-dimensional integral vector with each component being the path number of a path from the source code to corresponding end node. To make the multicast tree a feasible solution of the problem, we must delete loops existing in it. The operation of loop deletion is as follows.

Loop-deletion Operation:

1. ***for*** *I=1 to* $|M|$

2. ***for*** *j=1 to i+1*

3. ***if*** *there exists loop between ith and jth path of the path serial;*

4. Compare the costs of the routs that constitute the loop and delete the more expensive rout.

5. ***endif***

6. ***endfor***

7. ***endfor***

By the above operation, we can obtain a non-loop optimal multicast tree with lowest cost.

5 Experiment

To test the performance of the IPSO-based Multicast Routing Algorithm, we use the network model in Figure 1 as our tested problem. In the experiments, it is assumed that all the end nodes of multicast satisfy the same set of QoS constraints without regard to the characteristics of the nodes. The characteristics of the edges described by a quaternion (d, j, b, c) with the components representing delay, delay-jitter, bandwidth and cost, respectively. For performance comparison, we also used Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) to test the problem. The experiments were realized with Visual C++6.0 on Windows XP and executed on a PC with 2.10GHz-CPU and 256MB-RAM.

The experiment configuration is as follows. The population size for PSO and IPSO is 50 and maximum number of iterations is for all three algorithms and the number of the end nodes is 5. The fitness function is formula (17) with ω_1=1, ω_2=0.5, ω_3=0.5, ω_4=0.3, α=0.5, β=0.5, σ=0.5. There are 23 nodes in the network model (Figure 1), and we assume node 0 to be the source node; the set of end nodes to be M= {4,9,14,19,22}. For IPSO, the value of varies linearly from 1.0 to 0.5 over the course of running. The inertia weight *w* in PSO decreases linearly from 0.9 to 0.4 over a running and acceleration coefficients c_1 and c_2 are fixed at 2.0. For GA, the population size is 100 and binary tournament selection is used. The probability of crossover operation is 0.2 and that of mutation operation is 0.002.

We adopt two sets of constraints in the experiments:

1) When delay constraint D=20, delay-jitter constraint J=30, bandwidth constraint B=40 and packet loss constraint L=0.002, the multicast trees generated by the three algorithms are shown in Figure 2(a), Figure 3(a) and Figure 4(a), respectively.

2) When delay constraint D=25, delay-jitter constraint J=35 and bandwidth constraint B=40 and packet loss constraint L=0.002, the multicast trees generated by the three algorithms are shown in Figure 2(b), Figure 3(b) and Figure 4(b), respectively.

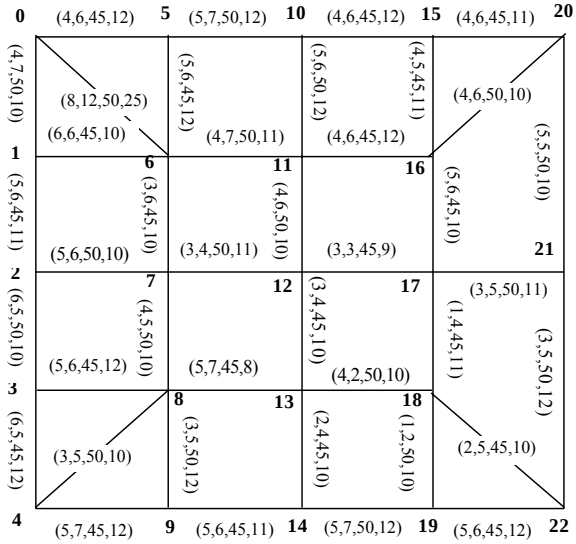


Fig. 1. A network model as the testing paradigm in our experiments

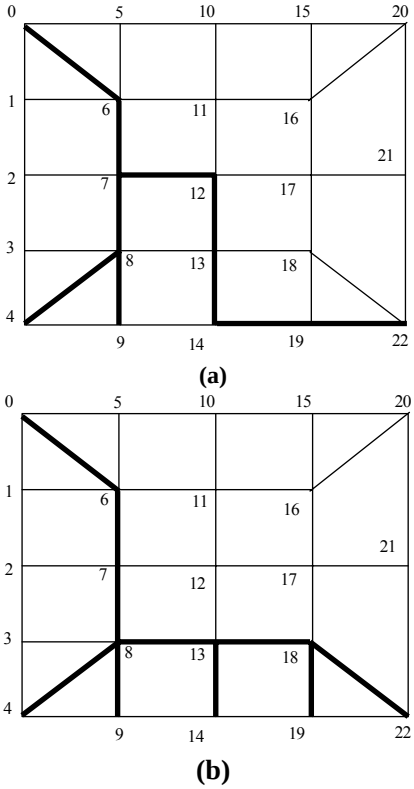


Fig. 2. Multicast trees (broad-brush) generated by Genetic Algorithm. (a). D=20, J=30, B=40 and L=0.002; (b). D=25, J=35, B=40 and L=0.002;

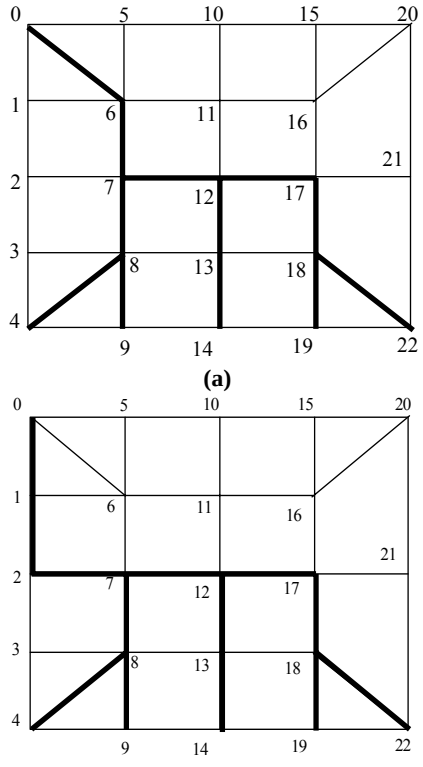


Fig. 3. Multicast trees (broad-brush) generated by PSO Algorithm. (a). D=20, J=30, B=40 and L=0.002; (b). D=25, J=35, B=40 and L=0.002;

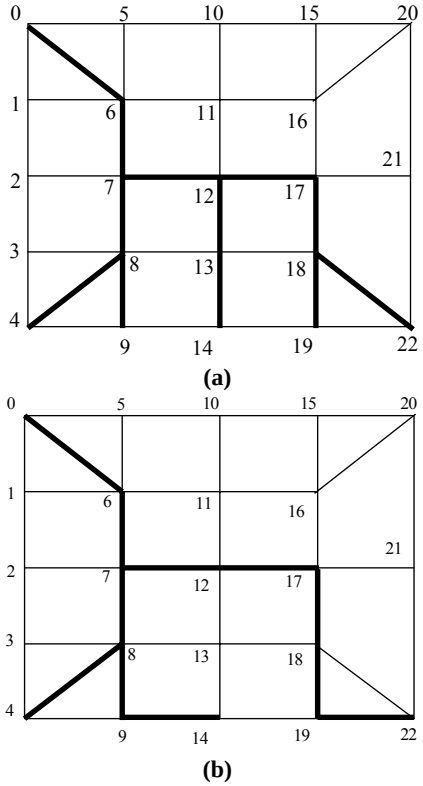


Fig. 4. Multicast trees (broad-brush) generated by IPSO Algorithm. (a). D=20, J=30, B=40 and L=0.002; (b). D=25, J=35, B=40 and L=0.002;

For constraints that D=25, J=35, B=40 and L=0.002, we recorded in Table 1 the dynamic changes of best fitness values as the algorithms are executing. The best fitness values generated by IPSO, PSO and GA after 200 iterations are 0.226449, 0.223214 and 0.116460. Thus we can conclude that IPSO has the best performance and could yield the better-multicast tree than two other algorithms. Figure 5(a) is the visualization of Table 1 and Figure (b) to (d) shows the dynamic changes of cost, delay and delay-jitter with the development of iteration for three algorithms. It can be seen that convergence speed of IPSO is most rapid, while Genetic Algorithm may encounter premature convergence. It means that among the three algorithms, IPSO has the stronger global search ability than PSO and GA for the tested problem.

6 Conclusion

The paper has presented a IPSO-based multicast routing policy for Internet, mobile network or other high-performance networks. This algorithm provides QoS-sensitive paths in a scalable and flexible way in the networks environment. It can also optimize the network resources such as bandwidth and delay, and can converge to the optimal on near-optimal solution within less iteration. The incremental rate of computational cost can close to polynomial and is less than exponential rate. The availability and efficiency of IPSO on the problem have been verified by experiments. We also test the other two heuristics, PSO and GA, for performance comparison, and the experiment results show that IPSO outperforms PSO and GA on QoS the tested multicast routing problem. Our future work will focus on using IPSO to solve QoS multicast routing in network environment with uncertain parameter.

Reference

[1] C. Moses, N. Joseph, et al, “Resource optimization in QoS multicast routing of real-time multimedia,” Proc of IEEE INFOCOM. 2000, pp. 1518-1527.

[2] M. Clerc, “The Swarm and Queen: Towards a Deterministic and Adaptive Particle Swarm Optimization,” Proc. Congress on Evolutionary Computation 1999, pp. 1951-1957.

[3] R.A. Guerin, A. Orda. “QoS routing in networks with inaccurate information: Theory and algorithms,” *IEEE/ACM. Trans. On Networking*, No.3, Vol.7. June. 1999, pp. 350-363.

[4] J. Kennedy, R.C. Eberhart, “Particle Swarm Optimization,” Proc. IEEE Int’l Conference on Neural Networks, IV. Piscataway, NJ: IEEE Service Center, 1995, pp. 1942-1948.

[5] L.Y. Li, “The routing protocol for dynamic and large computer networks”. Journal of computers, No.2, Vol.11, 1998, pp. 137-144.

[6] Y. Shi, R. C. Eberhart, “A Modified Particle Swarm,” Proc. 1998 IEEE International Conference on Evolutionary Computation, pp. 1945-1950.

[7] B.M. Waxman. “Routing of multipoint Connections” *IEEE Journal of Selected Area in Communications*, Dec.1998, pp. 1617-1622.

Study of one-way hash function to Digital Signature Technology

Chuanhua Zhou [1,2], Baohua Zhao [2]
[1] **School of Management Science and Engineering, Anhui University of Technology**,
Ma'anshan city, Anhui Province 243002, P.R.China
[2] **Department of Computer Science and Technology, University of Science and Technology of China**
Hefei city, Anhui Province 230027, P.R.China
e-mail: chzhou@ahut.edu.cn , bhzhao@ustc.edu.cn
And
Gong Li [1]
[1] **School of Management Science and Engineering,Anhui University of Technology**
Ma'anshan city, Anhui Province 243002 P.R.China
e-mail : powerlee@ahut.edu.cn

ABSTRACT

Digital signature technology is very important in today's e-commercial environment. To ensure the security of Digital signatures, One-way hash function has recently been widely used. In this paper, a new method of constructing the one-way hash function is proposed on the basis of chaotic theory. Firstly, digital signature technology is introduced. Secondly, after a simply explanation of one-way Hash function, the process and algorithm of applying the Hè non-like map to the construction of a one-way hash function is depicted. And then, the Security of the Hash function obtained is analyzed. Finally, this method is exemplified by an application case, which indicates that this novel function is sensitive to initial conditions and irreversible, simple to construct, and easy to practice, thus applicable to digital signature field of E-commerce security.

Keywords: Digital signature, E-commerce, Chaos theory, Hash function.

1. INTRODUCTION

With the rapid development and wide application of Internet technology in recent years, E-commerce is coming to us as a technology bringing Internet into full play. E-commerce is the process of electronic transactions of goods service and information, which take place over Internet.

Though its development is very astonishing, it takes a small part in the global total volume of trade. A major barrier is how to ensure the security of transmitting data and confirm the identities of two transaction parties. Security is the key factor to ensure E-commerce to develop safely and orderly. It is also paid close attention to people. In traditional trade, we use personal signature or seal to denote one's identity and duty. In E-commerce, we use digital signature to verify identity and truth of data.

2. DIGITAL SIGNATURE TECHNOLOGY

The digital signature is analogous to the written signature. It must have the following properties:
The sender can verify the recipient's signature in the message.
The sender can't deny his signature afterwards.
The recipient can't forge the sender's signature.
The digital signature is based on public key encryption technology. Its basic principle is very simple and it works in the following steps: The sender creates a fixed length digital digest from the message and encrypts it with his private key to form his digital signature; The digital signature is then appended to the message and sent to the recipient with the message; While the receipt calculates the original message by Hash function and get digital digest H1， and then decrypts the signature phrase by the sender's public key and gains H2. If H1 is the same as H2, the recipient knows that it is the holder of the sender's private key who sent the message. By digital signature, the authentication and preventing denial of original message can be practiced

Digital digest is a method, which uses the one-way Hash function to avoid the message is changed, this function can turn a large of messages into a small information range and send it combined with the message to the receipt, so it is more adapted to E-commerce. Using one-way Hash function to extract plaintext, which needs encryption to a fixed length cipher (often 128-bit), it is also called digital fingerprint. It is uncorrelated with the length of the plaintext. The digital digest has the property of being formed completely different with only a little change of the content of the message.

3. ONE-WAY HASH FUNCTION

A Hash value is generated by a function H of the form
$h= H (M)$
Where M is a variable-length message and $H (M)$ is the fixed-length hash value. There are a lot of functions whose inputs are variable-length and outputs are fixed-length. But the one-way hash function has the following special properties:
h is relatively easy to compute for given M.
For any given code h, it is computationally infeasible to find M such that $H (M)$, it is called irreversibility.

The major purpose of one-way hash function is to produce a "fingerprint" of a file that others cannot forge. In some application, only the "one-way "property is not enough, it also needs "collision resistance", they are:
For any given block M' , it is computationally infeasible to find $M' \neq M$ with $H(M)= H（M'）$.This is referred to as weak collision resistance.It is computationally infeasible to find any pair (M, M') such that $H(M)= H(M')$.This is referred to as strong collision resistance.

4. ONE-WAY HASH FUNCTION BASED ON CHAOS

4.1 Hènon-like Map

It is very difficult to construct the function receiving any input size, especially one-way function. In practice, one-way hash function is generated by compression function. For a

given m-length input, one-way function produces a hash value of n-length output. The input function compressed is the group of messages and the output of the former group. The output is the hash value of all groups to the point, such that hash value Mi is *hi=f (Mi, hi-1)*

This hash value combined with the next group of message is considered as the next input of compression function, the hash value of last group is the hash value of the entire message. The algorithm of the most typical one-way hash functions are MD5 ， SHA. The reference [4] also presents a kind of method constructing one-way hash function based on chaos.

The chaotic phenomenon is a way of motion of non-linear dynamical system that shows in some conditions, it is a random action of the system in the unbalance of process. Chaos is generated by non-linear simple systems, and then it is a fixed rule without any random factor. In the ideal state without noise, if a chaotic system moves precisely in an unstable cyclical orbit, it will move in this unstable cyclical orbit forever. But, an unstable cyclical orbit means that any small disturbance caused by internal or external reasons will make the system to deviate from the cyclical orbit very fast, the size of deviation will partial increase exponent with time, if the initial conditions of a chaotic system have a finite precision, with time passing, the precision of its state will be worse, at last, it cannot be accepted, its long run action will show obvious random and unpredictability. It is the sensitivity to the initial conditions, unpredictability and ergodic property that make chaos have all properties required by one-way hash function, and thus we can construct one-way hash function based on chaos.

A dynamical equation of Hènon-like attractor model is

$$\begin{aligned} x_{n+1} &= 1 - \mathrm{a}x_n^2 + y_n \\ y_{n+1} &= \mathrm{b}\sin(wx_n) \end{aligned} \qquad (1)$$

This is a two-dimension non-linear mapping, its one-dimension form is:

$$x_{n+1} = 1 - ax_n^{\ 2} + b\sin(wx_{n-1}) \qquad (2)$$

Suppose an initial values (0.54,0.24,0.69,10000) in given, the chaos graph can be obtained as shown in figure 1.

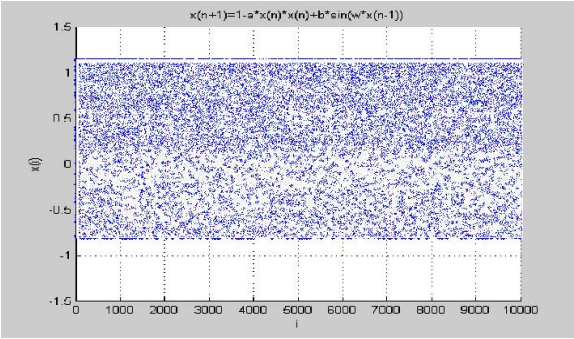


Fig. 1. The chaos graph

Where

$x \in [-1.5,1.5]$, $a = 1.4, b = 0.3, w \in [0.2,1.3]$.

Any point in the figure is distributed randomly, non-periodically, which means it is neither convergent nor divergent. Therefore, the iterative result of this function is chaotic.

The values of parameters of system (2) are a=1.4,b=0.3, 0.2 ⩽ *Ω*⩽1.3，the system has a stable chaotic attractor[3] when its initial value is between ± 1.5 in the equation (2) .It is very difficult to infer to the n-1_th and n-2_th terms if only the *n_th* term is known, so it ensures that the system has irreversibility and the property of preventing forgery. The system can be well distributed after a number of fixed iterations to arbitrary initial values. For (2) I choose the following different initial condition (0.42,0.24,0.69,5000), (0.42,0.2400001,0.69,5000), (0.42,0.24,0.6900001,5000) to run this function. For better observation, choose $n \in$ [4950,5000], I found that slightly change of initial condition can change the iterative result dramatically. So this chaotic system has a significant sensitivity to the initial conditions.

Because of the characters of the chaos of the system (2), which is sensitive dependent to initial conditions and irreversible because of the property of preventing forgery. It can be used to construct one-way hash function. Input a message, through iteration acts, and then to produce a fixed length hash value.

4.2 Constructing One-way Hash Function

Form equation (2) I construct a Hash function with key (*w, k1, k2)*, the key is to choose initial condition (*w, x0, x1*) for the first cycle. From figure 1, we can see that the system can has a better chaotic effect when the initial condition is between [-0.8，1.2], therefore I define $ki \in$[-0.8，1.2] in this algorithm. The construction process can be carried out through the following three modules: mapping 1, repeated iteration, mapping two.

1) Analysis of Construction Process

Mapping 1

Because all the condition selected and calculated of chaotic system are operating in floating point type data field. It has to map the digested characters on the data domain of the chaotic system. Here I will complete this using the mapping of common used characters and code of ASCII. To obtain a better mapping effect I improve the mapping as following:

$C(i)$=0.225+0.004×ch.

Ch is the *ith* character where $i \in$(0,n); 0.225 here is to prevent the condition of 0 in ASCII code still equal to 0 after mapping on the chaotic system, this will affect the chaotic effect of chaotic system; while the coefficient 0.004 is to increase the differentials of different conditions when mapping onto the chaotic system. Thus the mapping of two adjacent conditions in ASCII code on data domain of chaotic system will discrepant for 0.004. This will magnify the effect of sensitivity to the initial conditions in the chaotic system.

Iteration

First step, I substitute (*w, k1, k2*) into Hènon-like map (x0= *k1,x1= k2*) and start iteration. After n times iteration ,we get *Xn+1*; second, substitute *Xn+1* and the sequence *C(i)* (derived from character mapping of mapping 1) into Hè non-like map, let X0=C(i), *X1=Xn+1* and we get another Xn+1 after n times of iteration. Then repeat the second step for i times where $i \in$(0,n). Finally we get a chaotic sequence *X (i)*.

Mapping 2.

Since we have got the chaotic sequence *X (i)* from the iteration of the second module. We choose 8 values from it to ha(i)*(*$i \in$[0,7]), and map each of them to a 32 digit data

string which composed only of 0 and 1 hash[i][j]($i \in [0,7]$,$j \in [0,31]$).

2) Algorithm Design

The algorithm to construct one-way hash function is as follows:

Step1: input the key (*w, k1, k2*) of the hash function.
Step2: read plaintext *M* one by one and give value to *ch: ch* ← getchar().
Step3: calculate c[i],and repeat step2, step 3 until all the plaintext M have been inputted.
Step4: substitute *w←w, x0←k1, x1←k2* into Hènon-like mapping
Step5: calculate *Xn+1* and iterate for 500 times.
Step6: let *x0←c [i], x1←x500*; repeat step5 and 6 until all C (i) have been dealt with.
Step7: choose 8 values from *X (i)* we finally evaluate to *ha (i)*.
Step8: operate each ha[i] ($i \in [0,7]$) as follows:
if *ha[i]* $\geqslant$0, hash[i][0]=1; Otherwise: hash [i][0]=0
Let absolute value of ha [i], ha [i]=| ha [i]|.
if *ha[i]* $\geqslant$ 1, hash[i][0]=1 and ha[i]=ha[i]-1, otherwise, ha[i][j]=0,$j \in [0,31]$.
Repeat 3,4 until j=31 which means the cycle is ended until ha [i] gets a value.
Step9: output hash[i][j]($i \in [0,7]$,$j \in [0,31]$) .
From above, we can know that the one-way hash function based on chaos has the characters of simple algorithm and easy to practice. Thus it has big potential in the digital signature field of E-commerce security.
The algorithm designed by this paper can be implemented in any program language. However, due to the different characters of languages, there will be difference in operation speed to get the results across those languages, but the results remain the same. In this paper we use Visual C# to get the results. For research convenience, the program save the hash value of each calculation using a chain, it also include a function of comparing two different hash values and calculating the total number of different bits in two hash values.

4.3 Analyzing the Security of the Hash Function

Though the use of One-way hash functions Based on chaos in the provision of digital signatures field of E-commerce security. But the method has some problems. We cannot make clear some properties of chaos, such as collision and so on. Even if the scatter ability of chaotic function is very strong, it has probability of collision.

1) Birthday Attack

A birthday attack is a name used from the surprising result that the probability that two or more people in a group of 23 share the same birthday is greater than 1/2; such a result is called a birthday paradox. Birthday attacks are often used to find collisions of hash functions. For a code of length n if we were trying to find a collision, then by the birthday paradox we would expect that after trying *2n/2* possible input values we would have some collision. Because the message digest from the above algorithm is 256-bit, it needs 2128 hash computations to find a collision, which is safer.

2) Reverted Plaintext Attack

From the Hènon-like map (2) and the algorithm of one-way hash function, we know that we can get *ha(i)* by iterating the mapping value of key and plaintext. However, due to the extreme initial condition sensitivity and the non-periodicity of the chaotic system, we cannot derive the key and plaintext from *ha (i)*, And since we can dispose the missing value, the attacker cannot carry out attack with reverted plaintext for a minimum value because of the extreme sensitivity of the chaotic system.

5. APPLICATION CASE OF DIGITAL SIGNATURE

Here, an application case of the method above-mentioned is given to verify its effectiveness. Based on RSA of public key algorithm and chaotic one-way hash function, the process of application case of digital signature scheme is shown as figure 2.

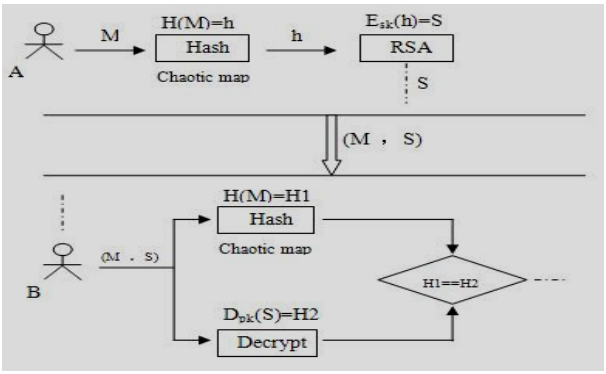


Fig. 2. The process of digital signature scheme

The sender A turns message M into short digital digest h through one-way hash function based on chaotic map, such that Hash *H(M)=h*， h can be produced very fast by *M*, but M is computationally infeasible to find M such that *H(M)=h*, then the sender use his private key SK to do RSA algorithm for digital digest, to produce signature S, denoted as *ESK（h）=S*. This digital digest is then appended to the message and combined with the message such that （*M*，*S*）, and sent to recipient B. After B receiving （*M*，*S*）,he will verify that S is the signature of A, B uses the same one-way hash function to produce digital digest H1 by computing to the message *M*, and then decrypts signature *S* using the sender's public key PK.H2 will be produced such that DPK （*ESK（h）*）*= H2*. If H1=H2，we can say S is A's signature. Thus, the recipient knows that the message could have only been sent by the holder of the sender's private key therefore, the message couldn't be modified in the transmission.

6. CONCLUSIONS AND FUTURE WORK

As an important component of E-commerce security, Digital signature will imposes higher requirements for security technology. Thus, we should continue to explore new methods so that we can find a more safe and simple signature method. This paper suggests us to construct a potential one-way hash function based on chaos theory, further more to form a newer digital signature scheme.

7. REFERENCES

[1] M. Henon, “A two-dimensional mapping with a strange attractor”, Comm Math Phys, 1976,Vol.50: 69-70.
[2] J.N. Liu, J.C. Xie, et al, “One way Hash function construction based on chaotic mapping”, Tsinghua University, 2000, 40(7): 55-58.
[3] S. William, Cryptography and Network Security: Principles and Practice, Tsinghua press, 2002, 253-299.
[4] Bruce Schneier, Applied Cryptography, China Machine Press, 2000, 307-319.
[5] F. Warwick, S.B. Michael, Secure Electronic Commerce, Posts & Telecommunications Press, 2002, 65-69.
[6] O. Goldreich, S. Goldwasser, et al, “How to construct random functions”, Journal of the ACM, Lecture Notes in Computer Science 19, October 1986, 33(4): 792-807.
[7] R. Abraham, L. Gardini, et al, Chaos in Discrete Dynamical Systems, New York: Telos, 1997.
[8] J. Dumortier, D.E.S. Van, “Electronic signatures and trusted archival services”, Proceedings of the DLMForum 2002, Barcelona 6-8 May 2002, Luxembourg, Office for Official Publications of the European Communities, 2002, 520-526.
[9] D.C. Lou, J.L. Liu, “Fault resilient and compression tolerant digital signature for image authentication”, IEEE Transactions on Consumer Electronics, 2000, 46(1): 31-39.

A Novel method to Evaluate the robustness of Network

Xinhua Wang,Liancheng Xu, Fangai Liu and Naiwen Liu
College of Information Science and Engineering，Shandong Normal University
Jinan, 250014, China
Email: wangxhe@sdnu.edu.cn

ABSTRACT

Network robustness is an active research area of network reliability and invulnerability is the main research content. A new model evaluating network invulnerability is presented based on routing-topology. It integrates communication traffic, routing algorithm and physical topology. This method gives more accurate evaluation of invulnerability compared with previous solutions. A heuristic algorithm accounting for selecting vital link set is also given by means of constructing routing-topology and establishing vital link set evaluation model. Simulation results demonstrate the effectiveness of this algorithm.

Keywords: invulnerability; routing-topology; vital link set.

1. INTRODUCTION

Network robustness describes the reliability level of communication network under malice destroying (such as terror affairs) [1,4,5,11]. Generally network robustness is related to the invulnerability，which can be evaluated to judge the reliability of communication networks. Some appropriate fault tolerance strategies can be adopted to improve robustness according to the analysis of network invulnerability. On the other hand it is possible that the capability of communication can be destroyed by attackers deploying the minimal cost solution such as military application field.

Generally speaking, network invulnerability refers to the reliability of network topology, which reflects the detailed distribution of network nodes (such as switch devices or hosts) and links (physical lines) .In [1] authors propose a network invulnerability model, which can be evaluated by evaluation rules of node-similar and optimal connectivity. Another evaluated model based on the degree and between ness of nodes and links is given in the literature [2,3] and some metrics of invulnerability are investigated under various network modes. Literatures [3,8,9,10] discuss the properties of invulnerability in complex network. Much previous work of network invulnerability research mainly depends on the characteristics of network physical topology and commonly graph theory is introduced to settle this problem. However, network topology is not the only factor that impacts the communication capability of network but others such as traffic distribution and routing algorithm are also included [6,7]. In this paper a new model evaluating network invulnerability is given by considering three important factors: traffic distribution, topology and routing strategies. We also propose a heuristic algorithm to evaluate vital link set and simulation results show that this algorithm performs effectively.

The rest of this paper is organized as follows: A new definition of invulnerability is given in section 2. Section 3 presents the evaluation model of establishing vital link set depending on routing-topology. A heuristic algorithm regarding vital links set and related instances are also provided. Section 4 concludes this paper.

2. NEW INVULNERABILITY DEFINITION

Network invulnerability commonly measures the most extent of destroying a network-connected topology under a limited number of nodes or links removed. It is assumed that the full information of network topology is exposed clearly to attackers and the solution of destroying is made explicitly. A definition of invulnerability is followed:

[Definition 1]: Invulnerability: It indicates the maximum tolerance ratio of network which satisfies the constraint of network-connected against removing nodes or links. Generally arbitrary node pair is guaranteed to exist at least one path.

The reliability of network topology against attack is measured according to this definition of invulnerability. However, another evaluating model depending on other factors perhaps would be more valuable versus network topology. In fact, traffic distribution and routing algorithm also impact the communication capability of network.

1) Traffic distribution is related to the communication capacity of network in practice. It impacts the distribution of payload in network. In this paper network traffic refers to the connections between nodes pair.

2) Routing algorithm determines whether one link is chosen by a special routing path. Results of invulnerability evaluation are different by implementing different routing algorithms. For example, distinct results exist between minimum link cost routing algorithm and load balance routing algorithm.

After synthesizing above factors we propose a novel evaluating model, which can represent the payload of communication traffic in practice. The definition of routing topology is given versus network topology.

[Definition 2] Routing topology: The reduced topology achieved from original network topology is based on traffic distribution and routing algorithm. The weight of link is the traffic payload which link carries. In this paper traffic payload is defined as the number of traffic connections traversing a physical link.

Routing topology is a subset of network topology mapping practical communication traffic. For example, we get routing topology (fig 1(3)) according to network traffic (table 1) and the minimum cost routing algorithm. There are some differences between routing topology and network topology:

1) Network topology is not equivalent to the routing topology. Network topology is not connected whereas *r*outing topology is connected fully once link e_7、e_8 and e_9 are disrupted. Apparently traffic connections are not impacted at all in the routing topology.

2) The weight of links in routing topology represents traffic payloads whereas one in network topology indicates distance or hop number between adjacent nodes.

According to analysis above, we give a novel revised definition of invulnerability.

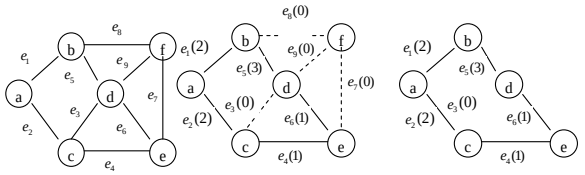
[Definition 3] Invulnerability: It indicates the maximum tolerance ratio of network which satisfies the constraint of network-connected based on routing topology against removing nodes or links .we denote it as symbol I .

3. THE APPROACH OF CONSTRUCTING A ROUTING TOPOLOGY

The capability of invulnerability for links is less than nodes. So in this paper we only discuss invulnerability model based on links. Firstly we investigate the approach of constructing a routing topology, which is the precondition to settle this problem.

Table 1. an example of traffic and routing information

traffic demands	connection	routing path
a-d		e_1 e_5
a-e		e_2 e_4
b-c		e_1 e_2
b-d		e_5
b-e		e_5 e_6



(1)Network topology (2)Routing topology(real line) and not selected links(dash line) (3)Routing topology

Fig.1. Network topology and routing topology

3.1 denotation definition

Network topology can be formulated as an undirected graph $G_P(V_P, E_P, W_P)$, where V_P to be a node set comprised of IP routers or hosts and E_P is a link set comprised of physical links. W_P is the distance of physical link. In contrast, routing topology is formulated as an undirected graph $G_R(V_R, E_R, W_R)$, where V_R to be a node set comprised of traffic demand node pairs (source and destination nodes) and E_R is a link set comprised of routing links. W_R is the traffic payload which the links carry.

Notation definition:

N : the number of network nodes.

L : the number of network links.

P : the set of serial number comprised of all the node pairs, notating $P = \{1,2,..., j..., N(N-1)/2\}$

C : the matrix of traffic demands, $C = \{c_j\}$,

$$c_j = \begin{cases} 1 & \text{If a connection is exists between the jth node pair} \\ 0 & \text{Else} \end{cases}$$

R : the matrix of routing path $R = \{r_j^i\}$, $1 \le i \le L$

$$r_j^i = \begin{cases} 1 & \text{If link } e_i \text{ is selected by routing path between the jth node pair} \\ 0 & \text{Else} \end{cases}$$

$P(e_i)$: the serial number set of all node pairs traversing link e_i .

$Load(e_i)$: the payload of network traffic over link i , which is equal to the number of traffic connections traversing link i . $Load(e_i) = |P(e_i)|$ .To discuss this problem simply, all of the traffic demands between the same node pair are regarded as one connection.

3.2 The Algorithm of routing topology

In this section we discuss the routing algorithm based on the minimum distance path, and an algorithm for constructing routing topology is proposed below:

1) Constructing matrix R of routing path. According to Dijkstra algorithm and network topology $G_P(V_P, E_P, W_P)$, all the short paths between node pairs of set P can be achieved .We denote $r_j^i = 1$ if connection between node pair j traverses link e_i , else $r_j^i = 0$ 。

2) The node set V_R of routing topology is achieved according to matrix C of network traffic demands.

3) The link set E_R of routing topology is achieved according to matrix C of network traffic demand and routing path matrix R , $E_R = \{e_i \mid (e_i \in E_P) and (c_j \times r_j^i = 1)\}$.

4) The weight set W_R of routing topology is achieved according to matrix C of network traffic demand and routing path matrix R $Load(e_i) = |P(e_i)|$, $P(e_i) = \{j \mid (j \in P) and (c_j \times r_j^i = 1)\}$.

So a routing topology G_R is constructed and an example is illustrated in Fig.1 (3) according to traffic demands in table 1.

4. INVULNERABILITY MODEL

4.1 The evaluating model of link effecting- value.

A concept of link effect value and the evaluating model of vital links set are given below:

[Definition 4] Link effect value: the extent to which network communication may be affected upon any link or links set failure for a given matrix of traffic.

We denote the i th link effecting value as $effect(e_i)$, which is the ratio of payload between the i th link and whole network.

$$effect(e_i) = |P(e_i)| / \sum_{j=1}^{\frac{N(N-1)}{2}} c_j , \quad 1 \le i \le L \qquad (1)$$

A measure expression of links set effect value is proposed below:

$$effect(e_{x_1}, e_{x_2}, ..., e_{x_i}, ..e_{x_h}) = |\bigcup_{i=1}^{h} P(e_{x_i})| / \sum_{j=1}^{\frac{N(N-1)}{2}} c_j$$

$$, \{x_1, x_2, ..., x_i, ...x_h\} \subseteq E_R, \ 1 \le h \le L \quad (2)$$

Now we give a definition of vital link set in order to achieve the link set which has the maximum link effect value.

[Definition 5] The vital link set includes some links, which have the minimum links and arrive at the maximum link effect value allowed by the capability of network communication. We denote K to represent this set.

So the value of network invulnerability is represented as:

$$I = |K| / L \quad (3)$$

4.2 The algorithm of selecting vital link set:

According to definition 4, $K \subseteq E_R$ so the problem of getting an exact vital link set is an NP-hard combinatorial optimization problem. We propose a heuristic algorithm, which can achieve a vital link set quickly.

The algorithm of getting vital link set is below:

1) Initialization. $K = \{\}$； $sele\_links = \{\}$

// K is vital link set， $sele\_links$ is a candidate vital link set.

2) Constructing routing topology $G_R(V_R, E_R, W_R)$ and nodes pair set $P(e_i)$ according to matrix of traffic demands and routing algorithm, $E_{sele} = E_R$.

3) While $(E_{sele} != \{\} \ and \ effect(K) < upper)$

// $upper$ ： The constraint threshold value of network communication, $0 \le upper \le 1$

{ $sele\_links = \{e_i \mid (e_i \in E_{sele}) and (\max(|P(e_i)|))\}$

$e_{sele} = \{e_i \mid (e_i \in sele\_links) and (\min(i))\}$

$K \Leftarrow K + e_{sele}$； $K \Leftarrow K + e_{sele}$; $sele\_links = \{\}$

$P(e_i) = P(e_i) - P(e_{sele})$ }

4) End

We give an instance to select a vital link set according to above algorithm. A NSFNET with 14 nodes is given[7] (Fig. 2) and a traffic demand exists between every node pair. Some vital link sets are achieved in table 2 according to algorithm.

Table 2. An example of vital link set & link effect value、invulnerability value

Vital link set	Link effect value	Invulnerability value
$\{e_{16}\}$	26.4%	0．048
$\{e_{16},e_5\}$	42.8%	0．095
$\{e_{16},e5,e_{12}\}$	56.0%	0．143
$\{e_{16},e_5,e_{12},e_3\}$	62.6%	0．191
$\{e_{16},e_5,e_{12},e_3,e_{13}\}$	69.2%	0．238
$\{e_{16},e_5,e_{12},e_3,e_{13},e_7\}$	73.6%	0．286
$\{e_{16},e_5,e_{12},e_3,e_{13},e_7,e_8\}$	78.0%	0．333
$\{e_{16},e_5,e_{12},e_3,e_{13},e_7,e_8,e_{15}\}$	81.3%	0．381

5. CONCLUSION

In this paper, we propose a novel evaluating model of network invulnerability considering routing topology. An approach of constructing routing topology is given under special traffic distribution and routing algorithm. The evaluated value of network invulnerability can be represented according to expression of link set effect value. This model could evaluate the capability of network invulnerability efficiently and also provide accurate reference knowledge to construct robustness network.

6. ACKNOWLEDGMENTS:

This research is supported by National Natural Science Foundation of China under grant NO.69933020.

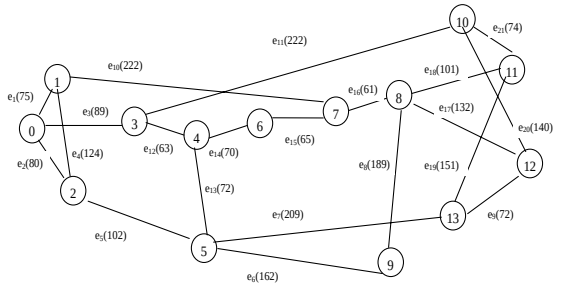


Fig. 2. the Topology of NSFNET

7. REFERENCES:

[1] A.H. Dekker, B.D. Colbert, “Network robustness and Graph Topology”, Proceedings of the 27th conference on Australasian computer science, Dunedin, New Zealand, Estivill Castro, V., ed., Conferences in Research and Practice in Information Technology, 26, 2004, 359-368.

[2] P. Holme, B. J. Kim, “attack vulnerability of complex networks”, Physical review E, 2002, 65(5): 56-109.

[3] P. Crucitti, V. Latora, “Error and attack tolerance of complex networks”, Physica review a, 2004, 388-394.

[4] R. Wilkov, “Analysis and Design of Reliable computer Networks”, IEEE Transactions On communication, 1972, 20(3): 660-678.

[5] L.B Page, J.E. Perry, “Reliability polynomials and link importance in networks”, IEEE Transactions on Reliability, 1994, 43(1): 51-58.

[6] B. Sanso, F. Soumis, “Communication and transportation network reliability using routing models”, IEEE Transactions on Reliability, 1991, 40(1): 29-38.

[7] B. Wang, “Towards Cost-Effective Provisioning and Survivability in Ultra High Speed Networks,” DoE Chicago Fermi Lab, Chicago, IL, September 15, 2004.

[8]http://computing.fnal.gov/cd/DOE/ultranet/talk-bin-wang.htm.

[9] Y. Jiang, B.X. Fang, et al, “An example of analyzing the characteristics of a large scale ISP topology measured from multiple vantage points”, Journal of software, 2005, 16(5): 846-856.

[10] J. Wu, Y.J. Tan, et al, “Study on measure of complex network invulnerability”, Journal of systems engineering, 2005, 20(2): 128-131.

[11] P.C. Luo, G. Jin, et al, “A review of study on reliability of communication network”, MINI-MICRO system, 2000, 21(10):1073-1077.

An Elliptic Curve Proxy Signature Scheme

Youan Xiao, Zude Zhou, Layuan Li
School of Information Engineering, Wuhan University of Technology
Wuhan, Hubei 430070, P.R.China
Email: youan@21cn.com

ABSTRACT

In this paper, a new controllable proxy digital signature scheme based on the elliptic curve cryptosystem is proposed to solve the problems that most of the known proxy digital signature schemes cannot provide the overall, complete, reliable control ability on the stretch of proxy digital signature power. It has the ability to provide the overall, reliable control on the using of proxy digital signing power. In addition, it has been applied for the Chinese patent of invention (200510019214.4). Furthermore, we also analyzed the computational complexity and security about the proposed scheme. The proposed scheme, which is an extension of elliptic curve cryptosystems, is beneficial for the study on the domain for the proxy digital signature. It can be applied in many domains, such as Electronic Commerce, Distributed Network Computing, etc.

Keywords: digital signature, proxy digital signature, elliptic curve, elliptic curve discrete logarithmic problems, security.

1. INTRODUCTION

Elliptic Curve Cryptosystem was first introduced by N. Koblitz and V. Miller in 1985 independently, which is a kind of public key cryptosystem based on the elliptic curve discrete logarithmic problems [1]. This cryptosystem has relatively shorter key length, lower bandwidth requirements for data, and faster encryption and decryption process when compared with other public key cryptosystems with the same security level [2]. And the implementation and application of elliptic curve cryptography has become hot topic in the field of cryptography and information security in the recent years.

Digital signature, like physical signature, can be verified while a specific user affixed their signature to a document and it can be verified that the document is the same as when the user affixed the digital signature. In addition, the digital signature scheme has the ability of identity authentication, source differentiate, anti-denial, anti-forge, to ensure the document is real, safety, integrity and reliability [3]. Today, the digital signature has the recognition from law, and becomes one of the basis technology for the activities based on Internet such as electronic commercial, electronic affairs, electronic publication, etc [4].

In real life, we often use the seal in order to transfer our signing power to the other authorized agent. The delegated agent can sign some documents for us when needed. The proxy digital signature scheme that was firstly introduced by Mambo in 1996 is one of the important digital signature schemes that can simulate the functions of seal of out real life [5]. A proxy signature scheme is a method that allows an original signer to delegate his signing power to a proxy signer, in order to let the proxy signer act as agent via proxy signature scheme without leaking the private key of the original signer. It has the full function of a common digital signature scheme, that anyone who knows the public key of the signers can verify a proxy signature is legal or not.

Today, there are many proxy digital signature schemes presented by many scholars over the world [6~11]. Most of the schemes are based on the problems of factoring a large integer or the discrete logarithmic problems. In addition, there are some of the proxy digital signature schemes based on the elliptic curve discrete logarithmic problems.

Most of the known proxy digital signature schemes have only solved the problem of the basis signing power delegation and transfer. However, we often need to set some limit for the proxy signing power in order to prevent the abuse of it by the proxy digital signer. Sometimes we need to set one or more delegation periods to the proxy signer. For example, when a boss is on a business trip, he may needs to delegate his digital signing power to his assistant only in the travel period or only in the office hours even. Sometimes we need to set the scope of proxy signing authority order to limit the proxy signer can only sign the special documents. For example, an enterprise product solicitor can sign the special business sales contract on behalf of his boss only in the scope of authority. Sometimes we need to set the maximum signing times that the proxy signer can exercise his proxy signing power. For example, the maximum signing times should be set to 150 when the college of a university is delegated to sign the electronic diplomas of 150 students by the committee for the conferment of academic degrees. To meet all these requirements, the proxy digital signature scheme must have the ability to provide the overall, reliable control to the using of proxy digital signing power based on the need of original signer, including the delegation periods, the scope of authority, and the maximum signing times.

Many known proxy digital signature schemes don't have the ability that setting a limit to the proxy digital signer. Some of the proxy digital signature schemes use a warrant appearing in the signature verification equation to declare the valid authorized scope and delegation period. However, the declaration in the warrant is useless because no one can know the exact time when the proxy signer signed a message. To avoid the proxy signer abusing the signing power, a new proxy digital signature scheme based on elliptic curve discrete algorithmic problems is presented to solve this problem in this paper. In the later sections of the paper, Section 2 illustrates the new proxy digital signature scheme and Section 3 emphasizes on the analyses of computational complexity and security. Finally, Section 4 concludes the research in various points.

* This work was supported by the Sunshine Young Project in Wuhan City of China under Grant No.20055003059-5, and it has been applied for the Chinese patent of invention (200510019214.4).

2. THE CONTROLLABLE PROXY DIGITAL SIGNATURE SCHEME BASED ON ELLIPTIC CURVE

The proposed proxy digital signature scheme is based on the discrete logarithm problem over elliptic curve groups. For more background information on the elliptic curve cryptosystem, please refer to the reference [12]. In this paper, the following parameters and arrangement about the proposed proxy digital signature scheme over the elliptic curve domain are required:

Suppose E is a security elliptic curve defined over the finite field F_p. The field size p is a large odd prime. Parameter $n = \#E$ is the order of the elliptic curve, which is equal to the number of points on the elliptic curve. G is a randomly selected element of the elliptic curve E called as the base point, whose order r is a large prime divisor of n.
The private key SK of a signer is a random integer in the range $[1, r - 1]$. And the corresponding public key PK, which is a point on E, is stored on the certification centre CA, where $PK=SK \times G$.

The structure of the proposed proxy digital signature scheme can be divided into three phases, including the digital signing power delegation phase, the proxy digital signature generation phase, and the proxy digital signature verification phase. The purpose of the digital signing power delegation phase is transferring the signing power from the original signer to the proxy signer without leaking the private key of the original signer. Then at the proxy digital signature generation phase, the proxy signer can sign the special message within the scope of authority. The function of the proxy digital signature verification phase is to let anyone who knows the public key of the signers can verify a proxy signature affixed with a message is legal or not.

2.1. Digital signing power delegation phase

If the original signer A wants to transfer his signing power to the proxy signer without leaking his private key, he executes the follow signing operations. The progress is described as Fig. 1.

Step 1: The original signer A creates a warrant named A_p containing some information about the limit of the proxy signing power, such as the appointed CA, the identity of the original and proxy signer, the delegation periods, the scope of authority, and the maximum signing times, etc.

Step 2: The original signer A signs A_p with his private key SK_A, affixes the signature to A_p, and gets the authority certificate C_p.

Step 3: A sends C_p to CA in a public channel.

Step 4: CA verifies C_p firstly. If it is legal, the maximum sign times should be set in its internal database based on C_p.

Step 5: The original signer A randomly selects an integer $k \in [1, r\text{-}1]$ in secure, and computes $Q_p = k \times G$.

Step 6: The original signer A computes S_p as following:

$$S_p = Hash(C_p, Q_p) \times SK_A + k \bmod r,$$

Where *Hash* () is a public collision resistant hash function such as SHA-1. Then the delegation parameter (C_p, S_p, Q_p) is constructed, and should be sent to the proxy signer B in a public channel.

Step 7: Once the proxy signer B receivers the delegation parameter (C_p, S_p, Q_p), he verifies C_p is legal or not firstly. If it is legal, the proxy signer B verifies whether $S_p \times G$ is equal to $Q_p + Hash(C_p, Q_p) \times PK_A$ or not.

Step 8: If the proxy signer B confirms the validity of the delegation parameter (C_p, S_p, Q_p), and accepts the delegation, he computes the proxy signature private key, which is known by the proxy signer B only, as follows:

$$SK_p = S_p + SK_B \times Hash(C_p, Q_p)$$

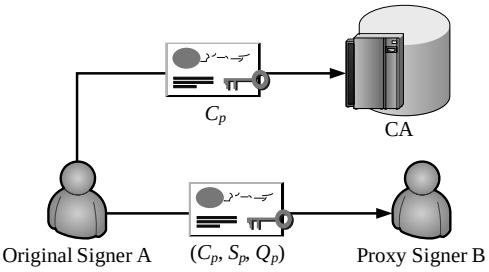


Fig. 1. Progress of the signing power delegation phase

2.2. Proxy digital signature generation phase

When the proxy signer B needs to sign a message m for the original signer A in the scope of authority, he executes the follow signing operations with CA. The progress is described as figure 2.

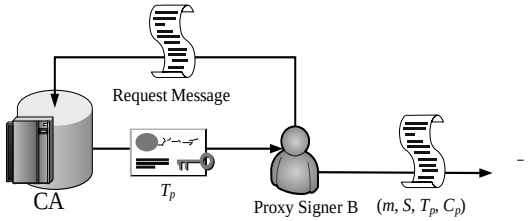


Fig. 2. Progress of the proxy signature generation phase

Step 1: The proxy signer B signs the special message m by executing the ordinary elliptic curve signature scheme with the proxy signature private key SK_p. Assume that the resulting signature is $S' = Sign(SK_p, m)$. Then the proxy signature is $S = (S', Q_p)$.

Step 2: The proxy signer B generates a request message which containing the information such as the authority certificate C_p, the proxy signature S, then signs the request message and sends it to CA for certification.

Step 3: Once CA receives the request message from B, it verifies the message firstly. If the message is illegal, reject it.

Step 4: CA searches its internal database for the maximum signing times defined by the original signer A, and the signing times that the proxy signer B has executed the proxy signing power, then checks the current time of CA.

Step 5: If the signature time is in the delegation periods, and the signing times is not larger than the maximum

signing times, CA signs the proxy digital signature S with current time of CA as a time-stamp certification T_p, and changes the internal database for the signing times. Otherwise, CA should refuse the request.

Step 6: CA sends the certification T_p to the proxy signer B in a public channel.

Step 7: The proxy signer B affix the proxy digital signature S, and its time-stamp certification T_p , the authority certificate C_p to the special message m, as the entire message $M = (m, S, T_p, C_p)$.

2.3. Proxy digital signature verification phase

Step 1: The verifier C confirms the validity of authority certificate C_p and time-stamp certification T_p .

Step 2: The verifier C checks if the proxy signing action is in the scope of authority or not.

Step 3: The verifier C computes the corresponding proxy signature public key by the equation:

$$PK_p = Hash(C_p, Q_p) \times (PK_A + PK_B) + Q_p$$

Where *Hash* () is the same public collision resistant hash function as the one in section 2.1.

Step 4: In the ordinary digital signature scheme, with the proxy signature public key PK_p , the verifier C confirms the validity of ordinary signature S', which is separated from the proxy signature, by verifying the accuracy of the verification equation as:

$$Ver(S', PK_p, m) = True.$$

3. SCHEME ANALYSES

In this section, we would emphasize on the analyses of correct proving, computational complexity and security about the proposed proxy signature scheme.

3.1. Correct Proving

This scheme can be proved as follows:
The proxy digital signature public key computed by the verifier C is:

$$\begin{aligned} PK_p' &= Hash(C_P, Q_P) \times (PK_A + PK_B) + Q_P \\ &= Hash(C_P, Q_P) \times (SK_A + SK_B) \times G + k \times G \\ &= (Hash(C_P, Q_P) \times SK_B + (k + Hash(C_P, Q_P) \times SK_A)) \times G \\ &= (SK_B \times Hash(C_P, Q_P) + S_p) \times G \\ &= SK_p \times G \end{aligned}$$

While the proxy digital signature public key generated by the proxy signer B should be $PK_p = SK_p \times G$. It is obvious that the PK_p is the same with PK_p'.

From here we see that this scheme is correct.

3.2. Performance Analyses

Because the time spent in the digital signing power delegation phase has small effect to the use of the new scheme, it can be ignored.

In the proxy signature generation phase, it is clear that the new scheme needs a new signature generation phase in order to sign the request message sent to CA when compared with the ordinary elliptic curve signature schemes. So the workload is twice as much as the workload of the ordinary elliptic curve signature schemes.

At the Proxy digital signature verification phase, the new scheme needs one extra scalar multiplication of elliptic curve point by integers and two extra addition of elliptic curve points than the ordinary elliptic curve signature schemes.

In a word, the new scheme needs few extra workloads than the ordinary schemes. However, we can have the overall, reliable control to the using of proxy digital signing power based on the need of original signer, so the extra workload is accepted.

3.3. Security considerations

Basically, the security of the proposed schemes is based on the difficulty of breaking the one-way hash function and the elliptic curve discrete logarithm problem (ECDLP). That is say; the difficulty for any attackers to forge another public key from the above equation is equivalent to the solution complicated by a one-way hash function and the ECDLP problem at the same time. Its difficulty is even than the ECDLP itself. There are several security properties such as unforgotten-ability, verifiability, un-deniability, distinguishing ability, identify-ability, etc. In this section, we shall consider some possible attacks against the proposed scheme. We shall prove that the proposed scheme can withstand these possible attacks.

Attack 1: Someone, including the proxy signer B, tries to derive the private key of original signer from all the available public information.

Analysis of Attack 1: The public information includes the public key of the signers, the delegation parameter (C_p, S_p, Q_p), and the proxy digital signature. Assume the adversary wants to derive the private key of original signer from the public information Q_p. It is as difficult as breaking the ECDLP to obtain the private key of original signer. Therefore, any other aggressors are none can forge any common digital signature too. Either the original signer A or the proxy signer B cannot a valid common digital signature after he signed a message. Therefore, the original cannot deny the authority certificate C_p.

Attack 2: Someone, including the original signer A, tries to forge a valid proxy digital signature of the proxy signer B for the message m to pass the proxy digital signature verification phase.

Analysis of Attack 2: The key to forge a valid proxy digital signature of the legal proxy signer B is deriving the proxy digital signature private key SK_p. From the equation $SK_p = S_p + SK_B \times Hash(C_p, Q_p)$,we know that is impossible to computes SK_p without SK_B. So any other one (including the original signer A) can not forge a valid proxy digital signature of the proxy signer B for the message m to pass the proxy digital signature verification phase. Therefore the proxy signer B cannot deny a valid proxy digital signature.

Attack 3: The proxy signer B tries to abuse his proxy signing power outside the scope of authority.

Analysis of Attack 3: The proposed scheme can set a limit to the proxy signer B, such as defining the delegation periods, declaring the scope of authority, and setting the

maximum signing times. If the proxy B tries to sign a message not in the delegation periods or exceed the maximum signing times, the CA will not subscribe a time-stamp certification T_p for the proxy digital signature *S*, which is generated by the proxy signer *B*. Assume the proxy *B* tries to forge time-stamp certification T_p, he will encounter the same difficulty as he tries to obtain the private key of CA or forge a proxy digital signature that he has signed before. If the proxy B tries to generate a proxy digital signature for the message which is gone beyond the limits of the scope of authority, the signature can be confirm illegal because of the authority certificate C_p. If the proxy *B* tries to forge the authority certificate C_p, he has to face the difficult problem of forge a valid digital signature of the original signer A. Because each proxy signature is affixed with the authority certificate C_p, so everyone can inspect who sign the message as an agent.

Therefore, the proxy signer cannot exceed the scope of delegation because of the authority certificate C_p and the certification center CA.

Attack 4: The proxy signer B tries to transfer directly his proxy signing power to another party D.

Analysis of Attack 4: Because the identity information of the proxy signer B is included in the authority certificate C_p, there for, it is impossible for D to generate a proxy digital signature with his own private key.

4. CONCLUSIONS

In this paper, we have proposed a new controllable proxy digital signature scheme based on the elliptic curve discrete logarithm problems. This scheme is securer than the typical ones. In this proposed scheme, the original signer can control reliably the scope of delegation, including the delegation periods, the scope of authority, and the maximum signing times.

For avoiding the abuse of the proxy signing power, the proposed scheme provides the ability to identify the actual original signer and the actual proxy signer, and control the scope of delegation, including the scope of delegation includes the delegation periods, the scope of authority, and the maximum signing times, with the authority certificate Cp and the assist from the certification center CA. The actual original signer cannot deny delegating the authority certificate, and the actual proxy signer cannot deny signing the proxy digital signature either. Some possible attacks have been considered. None of them can successfully break this proposed scheme.

5. REFERENCES

[1] Y.A. Xiao, “Researches on elliptic curve public key cryptosystems in network information security [PhD Thesis”. Wuhan University of Technology, 2003.10.

[2] Y.A. Xiao, L.Y. Li, “Research on the Xiao's key agreement scheme [J]”. Journal of Wuhan University of Technology (Transportation Science and Engineering), 2003, 27(6):758-762

[3] Y.A. Xiao, L.Y. Li, “Study of digital signature techniques”. Journal of Wuhan University of Technology (Transportation Science and Engineering), 2002, 26(6):737-740

[4] N.Zhang, Q.Shiet et al, “Efficient fair digital-signature exchange based on misbehavior penalization”,. IEEE Proceedings-Communications, 2005,152(3): 257-261

[5] M. Mambo, K. Usuda, et al, “Proxy signatures: Delegation of the power to sign messages” [J]. IEICE Trans. Fundamentals, 1996, E792A (9): 1338-1354.

[6] L.J. Yi, Study on Proxy Signature Schemes and their applications [PhD Thesis]. Xidian University, 2000.10.

[7] T.Okamoto, A.Inomata et al, “A proposal of short proxy signature using pairing”. International Conference on Information Technology: Coding and Computing, 2005. ITCC 2005. 2005. pp. 631-635

[8] H.Elkamchouchi, K.Elshenawy et al, “New digital signature schemes in the real and Gaussian integers' domains”. Proceedings of the Twenty-First National Radio Science Conference, 2004. NRSC 2004. 2004. C2, pp. 1-8.

[9] X.G. Cheng, J.M. Liu, et al, “An identity-based signature and its threshold version”. 19th International Conference on Advanced Information Networking and Applications, 2005. AINA 2005. 2005. pp. 973-977

[10] B. Lee, H. Kim, et al, “Strong proxy signature and its applications” [A]. International Conference on Information and Communication Security, Proc. of SCIS[C].Japan, 2001. pp. 603-608.

[11] C. Boyd, “Comment: new digital signature scheme based discrete logarithm” [J]. Electronics Letters, 2004, 30(6):480-481.

[12] IEEE 1363-2000, “Standard Specifications for Public-Key Cryptography”
Available at: http://www.ieee.org/grouper/1363

Control of dynamic correspondence network's simulation system

Jiabao Hu
Department of Computer Science, Wuhan University of Technology
Wuhan, Hubei 430063, P.R. China
Email: hjby@public.wh.hb.cn

ABSTRACT:

Network simulation is an important method in the network research. However, existing network systems have some deficiency in route simulation, which can't simulate dynamic network and imprecise state information. This article explores the method of researching dynamic network simulation system and schemes out a suit of network simulation software based on NS2, which can effectively dynamic network and imprecise state information. It provides a tool to research dynamic network and Qos route.

Keywords: simulation, dynamic network, imprecise state information, Qos route, and topology flow.

1 INTRODUCTION

Network simulation is a network virtual technology. It can build net topology, get flow load, accomplish net protocol and evaluate its performance through correlative software technology. Now there are several kinds of correlative software, such as NS, OPNET, OMNET++, COMNET etc [1]; but they have some deficiency on simulating dynamic network and imprecise state information. The text explores the method of researching dynamic network simulation system and think out a network simulation software DRS （Dynamic Routing Simulator）based on NS2 which can simulate effectively dynamic network and imprecise state information[2]. It provides a tool to research dynamic network and QoS routing for network researchers.

2 STIMULANT ENGINE

By and large, actual network simulation software adopts the architecture like chart 1. Imitating such software, the DRS adopt the frame, too. Stimulant engine is the foundational component [3]. That its design conception is general usually adopt the mode of affairs drive, which controls the whole movement of simulative process according to set temporal event queue. The time cell of event scheduling is virtual clock second. Usually, that stimulant engine's line scheduling and processor allocation take on the method of single line. Concurrent event be processed in the mode of first schedule [4].

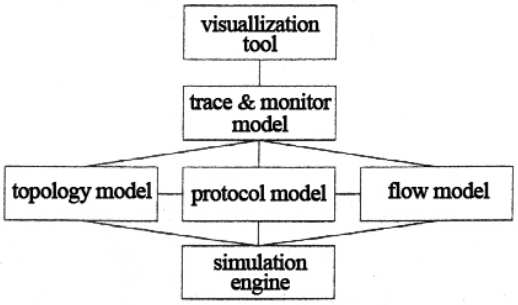


Chart 1 the architecture of network simulation software

For the favorable universality, openness and expansibility of NS2 network simulation software, we make use of the stimulant engine, network components, route management module, part flow generator and model for tracing and surveillance of NS2 in designing DRS. We can get the details of stimulant engine from document.

3 TOPOLOGY MODEL

To accomplish network simulation, we must abstract from real net element, reserve the underlying property and build model of network simulation through equivalent description. DRS' simulation model includes topology model, protocol model, flow model, in which topology model actualized by To-Builder (DRS' subset).

3.1 Random topology models

Random topology model can be classified into three types: rule model, for example: topology structure such as annulus, tree and astral model; well-known model, like trunk net of ARPA net and NSF net; random model. Different topology models have effect on route performance differently. We usually token diverse topology models with degree of node and net diameter. Random topology models are broadly used for research of routing algorithm, and Top-Builder is exactly the generator of random network topology. The fundamental principle to build random network topology is all network nodes are put in one surface randomly (or by heavy-tailed distribution), and add an edge (link) between nodes couple (u, v) with some probability, considering over each couple of nodes. There are various methods to commutate P (u, v), in which Waxman algorithm is most commonly used:

$$P(u,v) = a e^{-d(\beta L)}$$

Where $\alpha > 0$, $\beta \leq 1$ are model parameters, d is the Euclid distance from node u to v, and L is the maximum distance between any two nodes in the whole surface.

3.2 Design and realization of Top-Builder

The core pack unit of Top-Builder is composed of each model type that can generate different network topology. From received topology configuration files, these model types analyze parameters, which are needed in generating some kind of topology, and then accomplish the modeling process of this topology. In Top-Builder, the modeling process of random topology may split into four steps on the whole: place nodes in a surface; joint these nodes with Waxman algorithm; assign attribute to topology elements, such as id code of nodes, the bandwidth, delay and cost of link, and so on; output topology documents in TB (top builder) style. Now let me introduce some concrete modeling methods of network topology.

3.2.1 Topology of Router Level

Router Model class produces topology of Router level. Taking

on random distribution or heavy-tailed distribution, in the class function of Place Nodes achieve the placement of nodes in plane; Assign Bandwidth function can allocate bandwidth value of link by constant distribution, uniform distribution, exponential distribution or heavy-tailed distribution; link's delay is decided for Euclid length of itself; cost of link is its Euclid length.

3.2.2 Topology of AS (autonomic system) Level

AS Model class AS level produces topology of. The modeling method of such topology is very similar with one of Router style, all the difference between the two is just that every AS node has the ability to hold topology of Router level.

3.2.3 Topology of Hierarchical Structure

Modeling method of hierarchical topology: produce topology of AS level; beget a topology of Router level for every AS node; connect each one with junction enginery of some edge, according to communication of topology model of AS level.

Hypothesizing *RT（i）* and *RT（j）* denotes two facultative topology of Router level, and that u and v is node in *RT（i）* and *RT(j)*, we connect *RT(i）* and *RT(j)* with several edging-joint enginery of GT-ITM. random link: random picking u and v in *RT（i）* and *RT(j)*; link of least degree: that u and v is node of least degree in *RT（i）* and *RT(j)*; not leaf link of least degree: that u and v is node of least degree in *RT（i）* and *RT（j）*, but not leaf node; link of least *K*-degree: that u and v is node whose degree greater than or equal to *K* in *RT（i）* and *RT（j）*.[4]

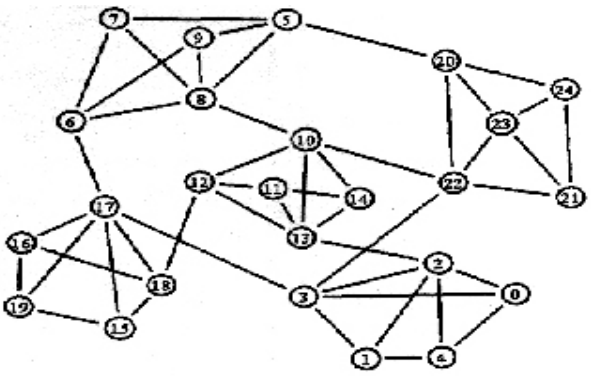


Chart 2 produced hierarchical topology

4 PROTOCOL MODEL

DRS is simulative software whose key function is to simulate dynamic network. Its protocol model mainly includes three function modules: QoS route computing, resource provision and refreshing of state information. Chart 3 displays their relation. Existing software doesn't have competence in emulating effectively load wave of real net and answerable imprecise status information, which can simulate dynamic of network topology. The QoS route algorism designed by us is able to provide QoS assurance for multimedia and real-time business, balance network load and accommodate imprecision of status information. Simulating to flow load and status information actualizes that QoS route algorism is not included in emulator design and the emulation of network dynamic and imprecise status information.

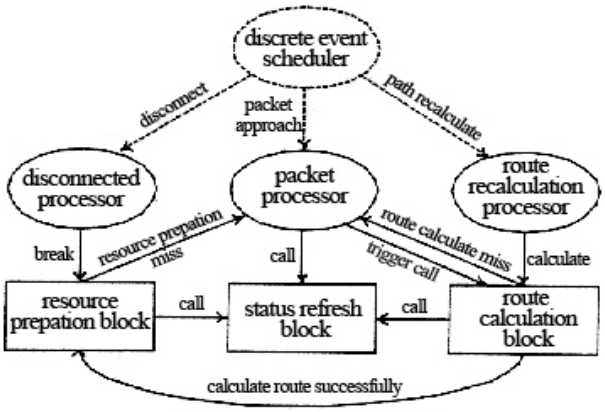


Chart 3 protocol models of DRS

4.1 Refresh Strategy of Status Information

Network status information includes the connectivity of topology structure, the remainder bandwidth, delay and cost of link and the queue delay and cost of node. In such status information, for that the connectivity of topology, transport delays and cost of link and cost of node have relative stability, the status information which has the most effect on route dynamic is remainder bandwidth of link and queue delay of node.

Traditional route protocol just broadcasts message of connective topology, while in the procedure of dynamic route, some dynamic status information (remainder bandwidth of link and queue delay of node) requires to be broadcasted to other nodes. By reasons of transport delay of link and high frequency fluctuation of status information, the status information each node concerned with other nodes and link is usually dated. We denote impreciseness rate of status information with δ.

$$\delta = \sup_{remum}\{\frac{|b_{act}(i,j) - b_{act}(i,j)|}{b_{act}(i,j)}\}$$

In the formula, b_{act} (i, j) is actual remainder bandwidth of link while computing route, and b_{act} (i,j) represents remainder bandwidth for route calculation. If the status information is postpone and imprecise, it will badly influence the calculation of Qos route. It can but renew status to guarantee the timeliness and preciseness of such information.

Policy of renewing status information mainly includes cyclic renew and trigger one. The former broadcasts renewed information cyclically to other nodes of network per fixed time interval. The latter triggers transmit of renewed information through detecting Δ_b, the alternation of link's remained bandwidth.

$$\Delta_b = \frac{|b_{newt}(i,j) - b_{old}(i,j)|}{b_{newt}(i,j)}$$

That b_{old} (i,j) and b_{newt} (i, j) are two remainder bandwidth of link (i, j); the first is the value before status change, and the second is the one after change. Cyclic renew introduces an interchange mode of status information, which can forecast load; trigger one provides more precise message. Especially in the case of heavy-load link, trigger renew have the ability to answer for tiny status change; its message load is greater

accordingly. So we usually combines the two renew strategy with hold-down timer, which can hold down over quick state refresh.

4.2 Selection of Renew Parameter

Here we choose a mixed renew model, which has three refresh elements: the one includes a refresh trigger (trigger off status renew through changing remainder bandwidth of link), a hold-down timer (be used for ensure least time interval of status renew) and a refresh circle (provide upper range of refresh time interval). In addition, while in the procedure for cyclic refresh, we allocate a renewed random number for each link in order to generate consecutive and smooth status renew, for the purpose of preventing producing load peak triggered by update synchronization.

Frequent state renew ensures timeliness and accuracy of status information, but the performance of QoS route will decline badly because of abundant update message generated. We denote refresh frequency of each link in unit time with $\theta = 2\lambda hN/L$ (please refer to next section for the meaning of parameters in the formula) and in this way the load of message updated. It is decided by route blocking rate, resource preparing miss ratio, load of renewed information, and load type of flow and network topological structure how to select parameter of status update.

4.2.1 Refresh Circle

The longer the refresh circle, the less the load of renewing message and the consumption for data processing, and yet the higher the imprecision of status information. Status refresh of large grain (long updated circle) will enlarge blocking rate of (including route blocking rate and miss ratio of resource preparation), and especially increase resource preparing miss ratio, which has apparent adverse effect on performance of route. Especially in the case of the duration of data flow is short and the topological structure is lank, the adverse effect appears more evidently than before. The refresh circle is usually thirty times of interval for generating connective quest.

4.2.2 Trigger Grain

Trigger of large grain can reduce the cost of processing and the load of updating message. It is almost unrelated between the blocking ratio of connection and the size of trigger grain, however, the trigger of large grain may minify route blocking rate and at the same time increase miss ratio of resource preparation. We can overlook the effect while choosing trigger grain. According to the above analysis, we should balk over-frequent trigger updating with a short-timed hold-down timer.

4.2.3 Hold-down Time

It will increase miss ratio of resource preparation while enlarging the time of hold-down timer. So usually this time is set as short as 1~2 times of the requested one.

5 FLOW MODEL

While designing the DRS flow model, we exploit some flow generators of NS2 as well as retain. The dynamic flow generators being adopted mainly include EXPOO, POO and Trace generators, whose function is to generate data flow with different characters, simulating application programs. CBR flow generator of NS2 mainly makes up static flow model, where nodes send out data in constant speed in order to beget constant flow.

5.1 Flow Model

In whole network typology, we choose source node and destination node couples. The connection-requesting numbers from each source node satisfy Poisson distribution in one time interval, that is 15 seconds, and arrival rate is λ. The connecting request of source node may also engender in the manner of unitary bursting, the time interval of its reaches conform to Weibull distribution whose shape parameter is a.

5.2 Character of Flow

The character of flow influences much over QoS route's calculation, status renew and the imprecision of status information. To be convenient for comparing performance in experiment, we set flow's time duration meeting two distributions: exponential distribution and Pareto one, among whom the average time of the former is τ , the shape parameter of the latter $\alpha = 2.5$. The bandwidth magnitudes of all flows distribute evenly on interval[$b - s / 2$, $b + s/2$], where the medium of bandwidth b is usually 2～5% of link capacity while the variation of bandwidth s is usually 5～10% of link capacity.

Finally, we import network load factor $\rho = \lambda\tau bhN/L$, where N and L refer to respectively the node count and the link mount of net, h is the average distance (being signified by hop count) of all the node couples between source to target.

6 CONCLUSION

This article has introduced the design method of dynamic network facsimile system; moreover, we have developed a suit of simulated software DRS that can simulate network dynamic route effectively. But being developed on the basic of NS2, DRS adopt a large number of network components and functional modules of NS2. How to upgrade DRS to a suit of independent route facsimile software will be our next target.

REFERENCES

[1] O. John, Scripting, higher-level programming for The 21st century, IEEE Computer, 1998,31 (3): 23~30.
[2] L.M Xu, B Pang, et al, NS and Network Simulation. Beijing, Posts & Telecom Press, 2003: 1~121.
[3] O. John, Tcl and Tk toolkit. MA, Addison Wesley, 1994.437.
[4] R. Brown, Calendar queues, a fast O (1) priority queue implementation for the simulation event set problem, Communication of the ACM 1998,31 (10): 1220~1227.

A VRF/RT-based Algorithm for Discovering VPN Topologies in BGP/MPLS IP VPNs

Haiying Liang [1,2], Wenxiu Bai [2], Peng Zhao [2], Jiansheng Yao [2], Yuan Gao [1]
[1] School of Information Science and Engineering, Northeastern University
Shenyang, Liaoning, 110004,China
[2] School of Computer, Jilin Normal University
Siping, Jilin, 136000,China
Email: haiying@mail.neuq.edu.cn

ABSTRACT

BGP/MPLS IP VPNs defined in RFC4364 allow service providers to use their IP backbone to provide VPN (Virtual Private Network) services. Currently most of service providers track 4364 VPN topologies either manually or by using a provisioning database. The goal of our algorithm is to discover VPN topologies as a composition of atomic and molecular components. Our algorithm uses a matrix model to represent a complex VPN. The algorithm decomposes this matrix to identify all the atomic components, and finally to construct molecular components from the atomic components. By using our algorithms service providers can automatically discover VPN topologies that have already been configured using the current network configuration information about VRF (VPN Routing and Forwarding) and RT (Route Target).

Keywords: VPN Topology, Discovery, RT, VRF.

1. INTRODUCTION

A VPN (Virtual Private Network) is a network where customers have connectivity across a shared infrastructure using the same access and security policies as a private network. One of methods to build VPN has been defined in RFC 4346[1] (i.e. BGP/MPLS IP VPNs). In BGP/MPLS IP VPNs, BGP (Border Gateway Protocol) is used to distribute VPN routing information and MPLS (Multiprotocol Label Switching) is used to forward VPN data traffic across SP (Service Provider) backbone network (i.e. Internet). BGP/MPLS IP VPNs allow service providers to define any arbitrary topology in a VPN. The service provider can create multiple VPNs using the same core network.

Currently most of the service providers track 4346 VPN topologies either manually or by using a provisioning database. Algorithms described in this paper aims at automating this VPN topologies' discovery procedure. Using our algorithms service providers can automatically discover VPNs that have already been configured using the current network configuration information.

2. OVERVIEW OF BGP/MPLS IP VPNS

Fig. 1. shows an example of the BGP/MPLS IP VPNs where there are two VPNs: VPN A and VPN B.

This work was supported by the National Science Foundation of China (NSFC) under the grant No.60073059 and No.60273078.

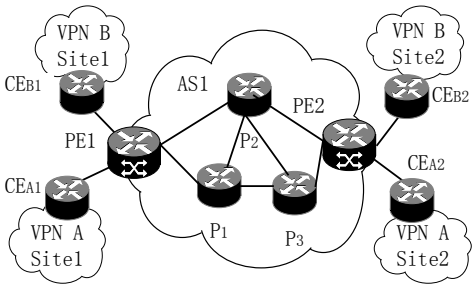


Fig. 1. BGP/MPLS IP VPNs

2.1 Costumer Edge (CE)

A CE router is a customer border device that connects the customer site via a data link, such as FR, ATM or leased line, to one or more PE routers. The CE device is a host or a router. Typically, the CE device is a router that establishes an adjacency with its directly connected PE routers. In this case, CE routers exchange routing information for network reachability of a customer VPN with PE routers using static routing, Interior Gateway Protocol (IGP) or External BGP (EBGP) between Autonomous Systems (ASes).

2.2 Provider Edge (PE)

PE routers are border routers in the SP backbone network that attach directly to CE devices. The PE router first learns local VPN routes from CE devices, and then exchanges VPN routing information with other PE routers using Internal BGP (IBGP) within AS.

2.3 VPN Routing and Forwarding (VRF)

A VRF table defines the VPN topology of a customer site attached to a PE router. Each PE router maintains per-VPN VRF tables and the IP addresses only need to be unique within each VRF table. The PE router is only required to maintain VPN routes for those VPNs to which it is directly attached. Only sites belonging to a VPN should be able to communicate with each other. The service provider associates each of the incoming ports at a PE router to a VPN routing and forwarding (VRF) table. This table contains VPN routing information exchanged by the PE router with the CE router connected to that port.

2.4 Route Target (RT)

Associating a particular Route Target attribute with a route, allows that route to be placed in all VRF tables that

are used for routing traffic received from the corresponding sites. The Route Target attribute is BGP extended community attribute [2]. RT attribute identifies a collection of VRFs to which a PE router distributes routes. A PE router uses this attribute to export local routes to other VRFs and to constrain the import of remote routes into its own VRFs.

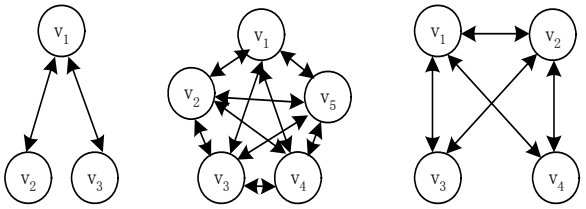
3. VPN TOPOLOGY

A VPN topology can be provisioned using RTs and the export and import of these RTs by the VRFs. Different VPN topologies can be provisioned [3]. Some of the topologies that are illustrated in Fig. 2 include:

(1) Single-hub-and-spoke (SHS): In this topology, a single hub VRF can send and receive VPN traffic to a set of spoke VRFs who are not capable of exchanging VPN traffic with each other. Fig. 2 (a) shows the SHS where $v_1$as a hub and v_2, v_3 as spokes.

(2) Full-mesh (FM): In this topology, a set of VRFs can all exchange VPN traffic with each other. That is, the VRFs are completely connected. Fig. 2 (b) shows the FM where v_1,v_2,v_3, v_4 and v_5 exchange VPN traffic with each other.

(3) Multi-hub-and-spoke (MHS): In this topology, a set of hub VRFs can exchange VPN traffic among each other as well exchange VPN traffic with a set of spoke VRFs. The spoke VRFs cannot exchange VPN traffic with each other. Fig. 2 (c) shows the MHS where $v_1$and v_2 as two hubs that form a FM and v_3 and v_4 as three spokes.



(a)SHS (b)FM (c) MHS

Fig. 2. VPN Topology

4. ALGORITHMS FOR TOPOLOGY DISCOVERY

4.1 Relative Definitions

When the VRFs are provisioned, they may be provisioned using a minimum number of RTs. For example, to provision a FM topology, only one RT is needed. As long as a single RT is defined on all the VRFs and is exported and imported by all the VRFs, VPN connectivity is established between every pair of VRFs thus leading to a FM topology. Similarly, to provision a SHS topology or a MHS topology, only two RTs are needed. One RT will be exported by the (multi) hub which will be imported by all the spokes and all the spokes will import a single RT which will be imported by the (multi) hub. We refer to the largest such components provisioned using the minimum number of RTs as atomic and molecular components as defined below.

1) Atomic Component Definition: The largest SHS topology with two RTs and the largest FM topology with one RT are atomic components. Denote the set of atomic FM components as F, atomic SHS components as HS. Fig. 2 .(a) and (b) are examples of atomic components.

2) Molecular Component Definition: The largest MHS topology with two RTs without any restriction on overlapping links and nodes with atomic components is called a molecular component. Denote the molecular MHS components as M. Fig. 2.(c) is an example of molecular component. Note that it is composed of four atomic components, one atomic FM component and three atomic SHS components.

4.2 The Discovery Algorithm

Given a description of a VPN using RTs, it can be decomposed into different sets of components. We use (f_1, f_2, …, f_x) to denote a FM topology created using nodes f_t, t = 1, …, x. We use (h → s_1, s_2, …, s_x) to denote a SHS topology with h as the hub and s_t, t = 1, …, x as the spokes. Similarly we use (h_1, h_2, …, h_y → s_1, s_2, …, s_x) to denote a MHS topology with h_t, t = 1, …,y as the hubs and s_t, t = 1, …, x as the spokes. Note that (h_1, h_2, …, h_y) is a FM topology. For the given network, let the number of VRFs be n and the number of RTs be m. Number the VRFs as v_1, v_2, …, v_n and number the RTs as r_1, r_2, …. r_m. The steps of our algorithm are enumerated below.

Step 1: Construct the VRF-RT Table (VR)

```
for k=1 to m
  { for t=1 to n
    { if r_k ∈ InputRT(v_t) then
        { if r_k ∈ OutputRT(v_t) then
            VR(r_k,v_t) =B
          else
            VR(r_k,v_t) =I }
      else
        { if r_k ∈ OutputRT(v_t) then
            VR(r_k,v_t) =E
          Else
            VR(r_k,v_t) =Null } } }
```

The preparation of this table would take O(nm) time.

Step 2: Construct the Relative Matrix (RM)

```
for k=1 to m
  { for t=1 to n
    { for j=1 to n
      { if t=j then
          RM (v_t,v_j)=Null
        else
          {if ((VR(r_k,v_t)=E ∨ VR(r_k,v_t)=B)∧
          (VR(r_k,v_t)=E∨VR(r_k,v_j)=I)) then
            RM (v_t,v_j)= r_k
          else
            RM (v_t,v_j)= Null }}}}
```

This step would take $O(n^2m)$ time.

Step 3: Determine the Set of Atomic FM Components (H)

```
y=1
for k=1 to m
  {x=o
   for t=1 to n
     {b_y=Φ
      if VR(r_k,v_t) =B then
        {x=x+1
         b_y= b_y ∪ {v_t} }}
   if x>1 then
```

{F=F ∪ b_y
d_y= b_y
y=y+1} }
for k=1 to m
{for t=1 to n
{for j=1 to n
{if (v_t,v_j ∈ F ∧ VR(r_k,v_t) =B ∧ VR(r_k,v_j) =B)
then RM(v_i,v_j)=RM(v_j,v_i)=Null } } }
It would require O(n^2m) time.

Step 4: Create the Set of Candidate Hubs (CH)
CH=Φ
for t=1 to n
{for j=1 to n
if RM(v_t,v_j)≠Φ then CH=CH ∪ {v_t} } }
It would require O(n^2) time.

Step 5: Create the Set of Preferred Hubs (PH)
PH=Φ
for k=1 to n
{if v_k ∈ F then
{for j=1 to m
{for t=1 to n
{if (VR(r_j,v_k) =B
∧ AM(v_k,v_t) = r_j ∧ VR(r_j,v_t)=I) then
PH=PH ∪ {v_t} } } }
It would require O(n^2m) time.

Step 6: Determine the Set of Atomic SHSs (HS) Component

We determine how many of hubs become part of an atomic SHS component. In order to qualify, there must be two distinct RTs, one where the candidate hub exports to a set of entries and the other where the candidate hub imports from the same set of entries. Denote the set of spokes to which the hub v_h uses r_k to export as SO(v_h, r_k).

While CH≠Φ do

(1) For each entry v_h ∈ CH, find all the distinct RTs r_k, 1 ≤ k ≤ m, used to export from v_h. Form the set of spokes SO(v_h, r_k) for each distinct RT r_k. SO(v_h, r_k) = {s | VR(v_h, s) = r_k}.

(2) For each of the set of spokes S(v_h, r_k), 1 ≤ k ≤ m, find the largest subset of nodes that uses the same RT to export to the hub. The cardinality of the largest subset is the in-degree of the hub.

In-Degree(v_h) = max1 ≤ k ≤ m | {s | VR(s, v_h) = rj, s ∈ S(v_h, r_k), r_j ≠ r_k, 1 ≤ j, k ≤ m}|.

(3) Find the hub v_h ∈ CH with the largest in-degree. If multiple hubs qualify, then select a hub v_h ∈ PH. Make (v_h → {s_i | VR(s_i, v_h) = r_j, s_i ∈ S(v_h, r_k), r_j ≠ r_k, 1 ≤ j,k ≤ m}). Include this SHS, (v_h → s_1, …, s_x) in HS. Therefore, HS = HS U {(v_h → s_1, …, s_x)}.

(4) Remove all links associated with this SHS component (v_h → s_1, …, s_x) from the graph. That is assign RM(v_h, s_i) = RM(s_i, v_h) =Null.

(5) Remove singleton nodes (i.e., nodes with no incoming and outgoing links) from CH.

It would take O(mn + n^2) time to execute the above sub steps.

Step 7: Determine Set of Molecular MHSs

The following steps prepare the set M.

(1) From the set F, take a new FM component all whose entries are members of PH. That is, F={b_k| b_k ∈ F ∧ b_k ∉ PH,1≤k≤m}. Stop if there are no new FM components for consideration.

(2) Check if each of the entries of the FM component b_k is a hub in the set of SHS.

(3) Check for each atomic SHS component (where the hub ∈ b_k) the RT exported by the hubs is the same and the RT exported is the same one for creating FM b_k. Also check that the RT imported by the hubs is the same.

(4) If the test passes, put the FM and the associated SHS components into the MHS set M and remove them from F and HS, respectively. Otherwise go back to step (1).

The time required to execute the above sub-steps would be O (nm).

Steps 6 and 7 are the cores of the algorithm where the complex VPN is discovered and represented as a composition of atomic and molecular components. This is the main goal of the algorithm.

5. EXPERIMENTS AND ANALYSIS

According to the configured information in a BGP/MPLS IP VPN, we explain steps of the algorithm to discover the relative VPN topology in the BGP/MPLS IP VPN.

Step 1: Construct the VRF-RT table of this network, shown table 1, where n=8 and m=6.

Table 1. VRF-RT Table (VR)

RT \ VRF	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
r_1	B	B	I	I				
r_2				B	B	B		
r_3	I	I	E	E				
r_4		E					I	I
r_5		I					E	E
r_6	E	B						

Step 2: Construct the relative matrix (RM) illustrated in table 2.

Table 2. Relative Matrix (RM)

VRF \ VRF	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
v_1		r_1,r_6	r_1	r_1				
v_2	r_1		r_1	r_1			r_4	r_4
v_3	r_3	r_3						
v_4	r_3	r_3			r_2	r_2		
v_5				r_2		r_2		
v_6				r_2	r_2			
v_7		r_5						
v_8		r_5						

Step 3: Determine the set of atomic FM component F={(v_1,v_2),(v_4,v_5,v_6)},the result relative matrix shown in table 3.

Table 3. The RM after deleting the atomic component

VRF \ VRF	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
v_1			r_1	r_1				
v_2			r_1	r_1			r_4	r_4
v_3	r_3	r_3						
v_4	r_3	r_3						
v_5								

v_6		
v_7	r_5	
v_8	r_5	

Step 4: Determine the set of candidate hubs CH={v_1,v_2,v_3,v_4,v_7,v_8}.

Step 5: Create the set of prefect hubs PH={v_1,v_2}.

Step 6: Determine the set of atomic SHS component. The analysis procedure is shown in table 4..

Table 4. The procedure of determining the set of atomic SHS

Spoke set	Elements	Export to hub	In-Degree
S(v_1,r_1)	{v_3,v_4}	{v_3,v_4}using $r_3$to v_1	Max|{v_3,v_4}|=2
S(v_2,r_1)	{v_3,v_4}	{v_3,v_4}using r_3 to v_2	Max|{v_3,v_4}|=2
S(v_3,r_3)	{v_1,v_2}	{v_1,v_2}using r_1 to v_3	Max|{v_1,v_2}|=2
S(v_4,r_3)	{v_1,v_2}	{v_1,v_2}using r_1 to v_4	Max|{v_1,v_2}|=2
S(v_2,r_4)	{v_7,v_8}	{v_7,v_8}using r_5 to v_2	Max|{v_7,v_8}|=2
S(v_7,r_5)	{v_2}	{v_2}using r_4 to v_7	Max|{v_2}|=1
S(v_8,r_5)	{v_2}	{v_2}using r_4 to v_8	Max|{v_2}|=1

From table 4, we can see that the maximum in-degree is 2, so v_1,v_2 and v_3 met this condition. We choice (v_1->v_3,v_4), since $v_1 \in$ PH. Hence, HS={(v_1->v_3,v_4)}, CH={v_2,v_3,v_4,v_7,v_8}. The relative matrix after deleting this atomic component is shown table 5.

Table 5. The RM after deleting this atomic component

VRF \ VRF	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
v_1								
v_2			r_1	r_1			r_4	r_4
v_3		r_3						
v_4		r_3						
v_5								
v_6								
v_7		r_5						
v_8		r_5						

Similarly we find atomic components (v_2->v_3,v_4) and (v_2->v_7,v_8).

So HS={(v_1->v_3,v_4), (v_2->v_3,v_4),(v_2->v_7,v_8)} and CH={}.

Step 7: Determine the set molecule MHS component. We find (v_1,v_2)$\in$F, and v_1 and v_2 are the hubs in HS, (v_1->v_3,v_4) and (v_2->v_3,v_4) are in HS. The RT outputted by v_3 and v_4 is r_3, and the RT inputted by v_1 and v_2 are the same one. So F={v_4,v_5, v_6), HS={ (v_2->v_7,v_8)}, M=(v_1,v_2->v_3,v_4)}.

By above 7 steps, we find the relative VPN topology shown in Fig. 3.

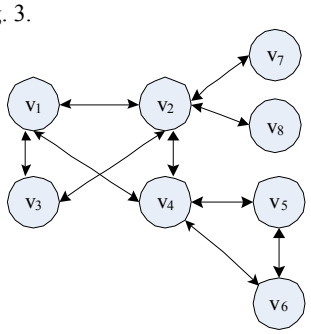


Fig. 3. VPN Topology Discovery

6. CONCLUSION

This paper uses the configuration information about VRF and RT stored in the BGP/ MPLS IP VPNs, to describe the algorithm-automating discovery of VPN topology procedure. The goal of our algorithm is to discover VPN topologies as a composition of atomic and molecular components. We first identify all the atomic components, and then construct molecular components from the atomic components. From the VRF-RT tables it can be seen that they do represent atomic and molecular components. An important problem to be solved is the discovery of different components that a VPN is comprised of. Using our algorithms service providers can automatically discover VPN topologies that have already been configured using the current network configuration information.

REFERENCES

[1] E. Rosen, Y. Rekhter, "RFC4364:BGP/MPLS IP VPNs", IETF, February 2006.

[2] R.S. Srihari, T. Daniel, "RFC4360:BGP Extended Communities Attribute", IETF, February 2006.

[3] W. Russ, M. Danny, et al, "Practical BGP", Addison Wesley Professional, July 2004.

Application of Independent Component Analysis Approach for Multi-variant Neurobiological Signals*

Jing Hu
College of Information Engineering, Zhejiang University of Technology
Hangzhou, 310032, China
Email: hj@zjut.edu.cn
Jie Hu
Dept. of Equipment, First Affiliated Hospital, Zhejiang University
Hangzhou, 310003, China
Email: jhuhz@yahoo.com

ABSTRACT

Electroencephalograms (EEG) can provide a unique window on the human brain. Contamination of EEG recording with artifacts, caused by muscular activity, eye movements, cardiac rhythm and power noise etc., can decrease the efficiency of diagnosis procedure. A kind of fast independent component analysis (ICA) approach present here is applied to analyze the real multi-variant EEG recorded signals. The removal of cardiac rhythm, comparison of before and after artifact removal, and comparison between ICA and a second order statistic (SOS) algorithm are analyzed respectively. The experiments based on the real measurement data demonstrate ICA is a powerful tool in neurophysiological interpretation.

Keywords: Independent Component Analysis, EEG, Multi-variant, Second Order Statistic, Artifact Removal.

1. INTRODUCTION

A great challenge in neurophysiology is to asses non-invasively the physiological changes occurring in different parts of the brain [1]. EEG, measuring electric field of signals emerging from neural currents within the brain, is a functional brain imaging technique with millisecond temporal resolution and millimeter spatial sensitivity.

To understand how the brain function works, recent advances statistical signal processing theory has provided new and powerful algorithms for the EEG recording analysis. Unfortunately, the EEG is frequently contaminated by artifacts originating from various sources such as scalp muscles, eye blinks, eye movements, or patient movement [2]. Although there are couples of techniques to deal with that, including adaptive noise canceling and singular value decomposition etc., they cannot provide detailed spatio-temporal information specific to the multiple brain sources that simultaneously contribute to the total signal [3]. Since it can be assumed that most of the artifacts are independent of the brain activity, it is hoped that ICA can find the artifacts.

ICA is a recently developed method in which the goal is to find a linear representation of non-gaussian data so that the components are statistically independent, or as independent as possible [4]. The problem of ICA has received wide attention in various fields such as data mining, wireless communication and biomedical signal processing etc.

Here, we present the application of this relatively novel technique to the real EEG recorded data. We focus on the removal of cardiac rhythm from the measured EEG signal in this paper due to the limitation of our experimental data. Moreover, comparison between ICA (based on high order statistic) and a second order statistic (SOS) method is also analyzed. Finally, results before and after artifact removal for measured EEG recording are obtained.

2. METHOD

As the ICA method interprets the measured data as a linear superposition of statistically independent processes. With this assumption, the data can be written as

$$X = AS \tag{1}$$

Where X is the observed data matrix with m (the number of channels, e.g., electrodes for EEG) by n (sampling time points), A is the so-called mixing matrix. S consists of the original source signals. The goal is to estimate both unknown A and S from X. With appropriate assumptions on the statistical properties of the source distributions, the solution is sought in the form,

$$S_{est} = WX \tag{2}$$

where W is called the separating or de-mixed matrix. In the ideal case, W should be the inverse matrix of A, thus S_{est} is just the expected source signal S.

The measurements contain signals resulting in the electrical activity of the brain but also signals, which can be considered artifacts. The key assumption used to distinguish sources from mixtures is that all the signal components in S, are statistically independent, i.e., the joint probability density is the product of the marginal densities. Meanwhile they should have non-Gaussian probability density functions, except at most one, which may be a Gaussian.

To Solve the problem is to find W iteratively. In our ICA-based method, firstly, the centering and whitening procedure is implemented, in order to get a new measured data $\tilde{X}$ with zero mean and unit variance. Furthermore, run the following steps, which is a kind of fast ICA algorithm,

Step1: initialize the de-mixed vector w_i ,

Step2: the iteration function is:

$$w_i = E\{\tilde{X} g(w_i^T \tilde{X})\} - E\{g'(w_i^T \tilde{X})\} w_i \tag{3}$$

let

$$g(u) = \tanh(u) \tag{4}$$

Step 3: Decorrelation: $w_i = w_i - \sum_{j=1}^{i-1} w_i^T w_j w_j$ (5)

* This work was supported by China postdoc science fund(2003033530) and Zhejiang Provincial 14th startup fund for studying abroad and returning personnel.

Step 4: Normalization, that is, $w_i = w_i / \|w_i\|$ (6)

Go to step 2 if not converged, otherwise, output the vector w.

To prevent different vectors from converging to the same maxima, step 3 is necessary. Thus, we can estimate several independent components with weight matrix W [4].

Additionally, for comparing with ICA algorithm, here we adopt a kind of second order statistics (we call SOS here) algorithm based on eigenvalue decomposition (EVD) to obtain the estimated signals. This SOS method belongs to the decorrelation technique, which ensure only that output pairs are uncorrelated, i.e., $< s_i s_j >= 0, \forall ij$, whereas ICA impose a much stronger criterion, statistical independence, which occurs when the multivariate probability density function factorizes, e.g., $\rho_S(S) = \prod_{i=1}^{N} \rho_{s_i}(s_i)$. Statistical independence requires that all second-order and higher-order correlations of the s_i are zero, while decorrelation only seeks to minimize second-order statistics (covariance or correlation)[5].

3. EXPERIMENTAL RESULTS

All signals used in our experiments were obtained from online polysomnographic database. There are seven-channel signals recorded synchronously, including one EEG (electrode placement recordings at positions C3), one Electrocardiography ECG) signal, one electrooculogram(EOG), one electromyography EMG), the rest signals are caused by abdominal movement, nasal air flow and blood pressure.

In consideration of data matrix dimensions, in experiments, we only select 4 signals out of 7 signals measured synchronously and meanwhile use a sampled interval of 5.996 seconds recording (a period from 1 to 1500 in the total 15000 sampling length). Therefore, the dimension of the experimental data is 4 by 1500. Fig.1 gives four selected measured signals in our experiments, which are also called the mixed signals, the only data we know before our experiments.

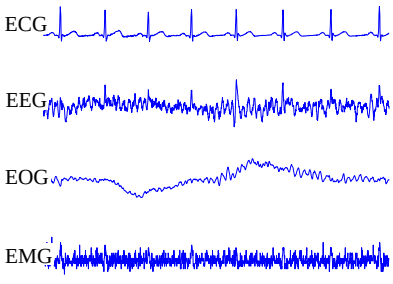


Fig. 1. Recorded signals

Then the fast ICA and SOS approaches are performed and then two de-mixed matrixes are computed respectively,

$$W_{ICA} = \begin{pmatrix} 3.4540 & 4.5591 & -0.7440 & 15.3987 \\ 0.6942 & -48.5505 & 27.9056 & 108.8050 \\ 1.3821 & -85.4983 & -19.3225 & 30.1904 \\ 1.7035 & -17.0226 & 1.8735 & -583.3683 \end{pmatrix}$$

and

$$W_{SOS} = \begin{pmatrix} 0.3603 & -4.3362 & 33.6008 & 17.3818 \\ -1.9786 & -60.1187 & 0.0019 & 66.4222 \\ -3.5488 & 79.4060 & 5.0823 & 103.2395 \\ -0.7451 & 2.3277 & 1.1076 & 333.9881 \end{pmatrix}$$

The mixture matrix A in Equation(1) is unknown since we cannot know the true original signals, therefore it is impossible to evaluate the separation capabilities of these two approaches quantitatively by two above yielded de-mixed matrixes W_{ICA} and W_{SOS}. We then turn to use the separated results shown in Fig.2 instead to see the result difference between two separation approachs. By comparison, we can know this SOS method is hardly capable of eliminating ECG from EEG signal, the component SOS3 still have the obvious SOS2 interferential effect, whereas by the fast ICA, component IC3 becomes a clean EEG signal.

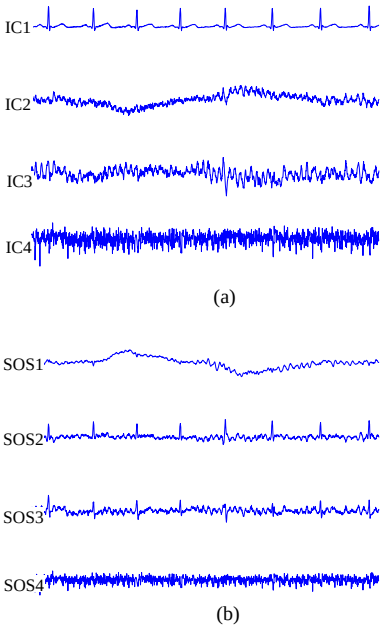


Fig. 2. Comparison of separated results by fast ICA and SOS algorithm: (a) fast ICA ;(b) SOS.

Furthermore, the results before and after the artifact removal are illustrated with Fig.3 and Fig.4. It is clear to see EEG is contaminated by ECG before the separation experiment, whereas this phenomenon disappears in Fig.4 after running ICA.

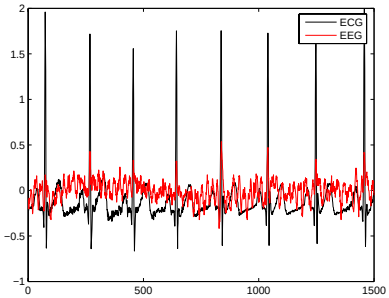


Fig. 3. Comparison between the measured EEG (red line) and ECG (black line) signals before using fast ICA algorithm

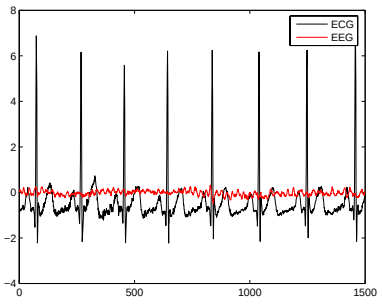


Fig. 4. The clean EEG (red line) after artifact removal and ECG signal (black line).

So, by applying ICA to mixtures of independent sources such as the real biological signals, we can reach the blind separation. We only focus on the removal of cardiac rhythm from the measured EEG signal due to the limitation of the experimental data here. As we know, there is a batch of various ICA algorithm developed so far. Due to the different algorithms suitable to the different applications, hence how to select the right ICA methods to remove affection of artifacts from EOG, EMG or power noise on EEG recording will be in our next work.

4. CONCLUSION

ICA is a powerful tool that aims to find the independent components from a set of signal. Neurophysiological interpretation of the ICA algorithm sources poses a further research challenge. We will develop this work in the near future. We can project the previously selected components of interest back to the recording space and by doing so we can obtain a set of cleaned measurements, and then we further want to know the relationship between the independent components and brain activities more directly. Meanwhile, we need to combine the prior and clinical neurophysiologic knowledge to analysis these separated components, to tell the true EEG source signals and artifacts among them, to know the information about dipole sources in brain (e.g., the amount of dipoles, and their location, magnitude and direction) which produce the measured electric field around the head and so on. We will try to make a more serious use of ICA into the EEG inverse problem since it can be an additional useful tool in preprocessing EEG data for source analysis. Moreover, due to some new variational ICA methods [6, 7] in the literature recently, this work is open and promising.

5. REFERENCES

[1] A. Cichock, “Blind Signal Processing Methods for Analyzing Multichannel Brain Signals,” International Journal of Bioelectromagnetism, 6(1), 2004.

[2] E. Niedermeyer, S.F. Lopes, “Electroencephalography: basic principles, clinical applications, and related fields. 5th ed. Philadelphia”, PA: Lippincott/Williams & Wilkins; 2005.

[3] C. Salustri, E. Kronberg, “Language-related brain activity revealed by independent component analysis,” Clinical Neurophysiology, Vol.115, 2004,385-395.

[4] A. Hyvarinen, E. Oja, “Independent component analysis: algorithms and applications”, Neural Networks, 13(4-5), 2000,411-430.

[5] T.P. Jung, S. Makeig, et al., “Removal of eye activity artifacts from visual event-related pententials in normal and clinical subjects”, Clinical Neurophysiology, Vol.111, 2000,1745-1758.

[6] S. Chiappa, D. Barber, “EEG classification using generative independent component analysis”, Neurocomputing, 2006,769–777.

[7] P.L. Van, E. Urrestarazu, et al, “A system for automatic artifact removal in ictal scalp EEG based on independent component analysis and Bayesian classification”, Clinical Neurophysiology, No.117, 2006,912–927.

An Assisting and Cooperative Localization Method in Wireless Sensor Networks

Xingyu Pi, Hongyi Yu
Communication Engineering Department, Information Science and Technology Institute
Zhengzhou City, Henan Province, P.R.China
Email: {psingyu@163.com, maxyucn@sohu.com}

ABSTRACT[1]

Wireless sensor networks (WSNs) are important for a number of strategic applications. Target localization is an important task in WSNs. In this paper, an assisting and cooperative localization method for the target in wireless sensor network systems is presented. This new approach uses energy measurements from the individual sensors in the sensor field to detect the targets first. Then the position of the target can be estimated using a distributed and cooperative localization model based on a simple geometry relation between the sensor nodes and the target. Extensive simulation results show that the new approach can locate the target effectively. Besides, it can reduce the complexity of the processing. Therefore, it is a feasible and attractive localization method for the wireless sensor network system.

Keywords: sensor networks; radio localization; cooperative processing; RSSI.

1. INTRODUCTION

The emergence of small, low-power devices that integrate micro-sensing and actuation with on-board processing and wireless communication capabilities stimulates great interests in wireless distributed sensor network. Such distributed sensor networks are important for a number of strategic applications such as coordinated target detection, surveillance, and localization [1, 2]. Source localization is an important task in wireless distributed sensor networking system.

This paper is organized as follows: Section 2 analyzes the localization model. Section 3 proposes some possible expand based on the analysis of the localization model. Section 4 describes the simulation experiment results. Section 5 concludes the paper.

2. LOCALIZATION MODEL

2.1 Problem Formulation

Considering a sensor network with a total of R sensors, which are randomly deployed in the region of interest (ROI), as shown in fig. 1.. Assume that the positions of sensors are all known. And (x_i, y_i) are the coordinates of sensor i, where $i = 1, \cdots, R$. The sensors use the energy decay model to pre-detect the target first, and then locate the target in the detected region.

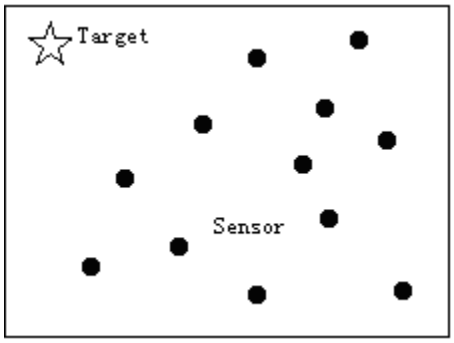


Fig. 1. Sensor network and target sketch map

2.2 Localization Method

Assume that there are K sensors (named as *detection sensor*) which have decided the presence of the target [4,5]. An isotropic signal power attenuation model is adopted here [3]: $P_i = P_0 / (1 + d_i)^{\alpha}$, where P_i is the received signal power of the *i*-th sensor, P_0 is the signal power emitted by the target at distance zero, d_i is the distance between the target and local sensor i, α is the signal decay exponent and takes values between 2 and 3. Denote the information set $\{(x_k, y_k), S_k\}$, where (x_k, y_k) is the coordinate of the *k*-th detection sensor, S_k is the received information containing both the target signal and background noise in the corresponding detection sensor, $k = 1, \cdots, K$. So the information set of all the detection sensor is $\bigcup_k \{(x_k, y_k), S_k\}$., where $k = 1, \cdots, K$.

Each detection sensor exchanges the received information each other and compare the difference P_{ij} of the received energy of the detection sensor i and j with the threshold- η_0, where $\eta_0 \to 0$. If $P_{ij} \to \eta_0$, select the two sensors and regard them as a group, otherwise they are rejected. Based on the model, the P_{ij} approximate to zero and the target must locate at the perpendicular bisector of the linked line of the two sensor nodes. As shown in Fig. 2.. Assume that there are M groups in all. Denote the position information of the two sensors in a group (x_i, y_i) and (x_j, y_j) respectively. Assume that

[1]This research work was supported in part by the National Nature Science Foundation program of China under contract number 60472064.

(x_s, y_s) is the coordinate of the target. Choose a group of sensors, the perpendicular bisector of the linked line of the two sensor nodes can be written as:

$$y_s - \frac{y_i + y_j}{2} = -\frac{x_i - x_j}{y_i - y_j}\left(x_s - \frac{x_i + x_j}{2}\right) \quad (1)$$

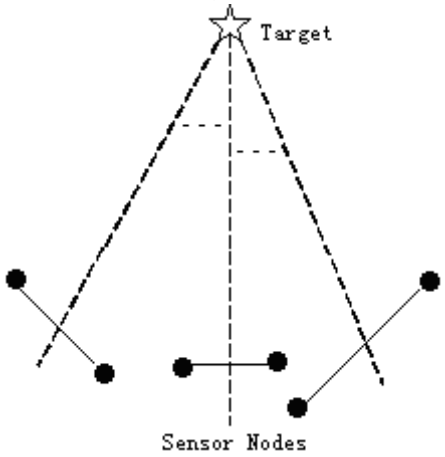


Fig. 2. Localization sketch map

For simplicity and ease of presentation, we analyze the proposed localization algorithm only considering two groups of the chosen sensor nodes. Without loss of generality, assume that the sensor position of group 1, 2 and M are at the coordinates $\{(x_1, y_1), (x_2, y_2)\}$, $\{(x_3, y_3), (x_4, y_4)\}$ and $\{(x_{M-1}, y_{M-1}), (x_M, y_M)\}$ respectively for a two dimensional system (although the extension of these results to 3-D is also possible).

So we can derive the following system of equations:

$$\left(y_s - \frac{y_1 + y_2}{2}\right) + \frac{x_1 - x_2}{y_1 - y_2}\left(x_s - \frac{x_1 + x_2}{2}\right) = 0 \quad (2)$$

$$\left(y_s - \frac{y_3 + y_4}{2}\right) + \frac{x_3 - x_4}{y_3 - y_4}\left(x_s - \frac{x_3 + x_4}{2}\right) = 0 \quad (3)$$

$$\vdots$$

$$\left(y_s - \frac{y_{M-1} + y_M}{2}\right) + \frac{x_{M-1} - x_M}{y_{M-1} - y_M}\left(x_s - \frac{x_{M-1} + x_M}{2}\right) = 0 \quad (4)$$

Thus the system equation can be given as follows:

$Ay = b$; Where: $y = \begin{bmatrix} x_s & y_s \end{bmatrix}^T$

$$A = \begin{bmatrix} (x_1 - x_2) & (y_1 - y_2) \\ (x_3 - x_4) & (y_3 - y_4) \\ \vdots & \vdots \\ (x_{M-1} - x_M) & (y_{M-1} - y_M) \end{bmatrix},$$

$$b = (1/2) \begin{bmatrix} \|r_1\|^2 - \|r_2\|^2 \\ \|r_3\|^2 - \|r_4\|^2 \\ \vdots \\ \|r_{M-1}\|^2 - \|r_M\|^2 \end{bmatrix}.$$

$\|\cdot\|$ is denoted by the Euclidean distance. r_i is the position of the *i*-th sensor, $i = 1, \cdots, M$.

The LS solution to estimate the target position has the form of $y_{est} = (A^T A)^{-1} A^T b$. (5)

3. POSSIBLE EXTENSIONS

3.1 TOA Model

If we could obtain the time of arrive (TOA) information of the target, then select the sensor groups that the TOA values of two detection sensors are approximately equal. That is, the target locates at the perpendicular bisector of the linked line of the two sensor nodes. Thus we can estimate the positions of the target with respect to multiple group sensors.

A drawback of the model is that the method needs to know the transmitted time of the target and the received time of the detection sensors accurately. Thus this is not practical.

3.2 TDOA Model

The time difference of arrival (TDOA) is obtained with respect to the received information of the detection sensors. Find the sensor groups ($TDOA \to 0$), then the target locates at the perpendicular bisector of the linked line of the two sensor nodes. Thus we can estimate the positions of the target with respect to multiple group sensors.

A drawback of the model is that the method needs to estimate the TDOA of two sensors, and this needs much computation and high complexity and will consume the energy of the nodes greatly.

4. SIMULATION AND ANALYSIS

In this section we present some computer simulation results to show the performance of the proposed technique under various environmental conditions such as deployment area, node position error, and number of sensor nodes. The data points represent averages over 50 trials in the networks.

We model position errors for the sensors deployment as Gaussian noises. With a deployment error-e_d, a random value drawing from a normal distribution $e_d \times N(0,1)$ is added to the sensor nodes' original grid position.

Simulation Results on Different Deployment Area - Fig.3. shows the localization performance based on difference topology and sensor number. From the graph, we can see that when the sensor number increases, the error rates of localization descend. But the topology has a greater effect on the localization error.

Simulation Results on Position Error - Fig.4. shows the localization estimation results for varying position errors of sensor nodes. This might be caused by uncorrected GPS position reports, incorrect background noise estimation or some individual sensor fault measurements. We can see that the position errors of sensor nodes have a greater effect on the localization performance relative to the sensor number by comparing the three graphs of Fig. 3. and 4..

Simulation Results on Both the Position Error and Density of Sensor Deployment- Fig.5. shows the

localization performance for both the position errors of sensor nodes and different sensor numbers (100, 200,300 and 400). From the graph we can see that sensor numbers have a less effect on the localization performance. But the topology has a greater effect relative to both of them.

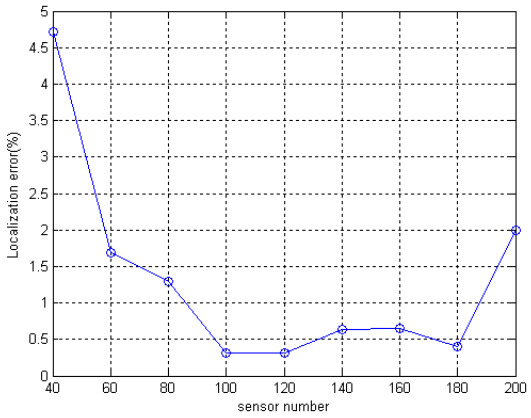


Fig. 3.

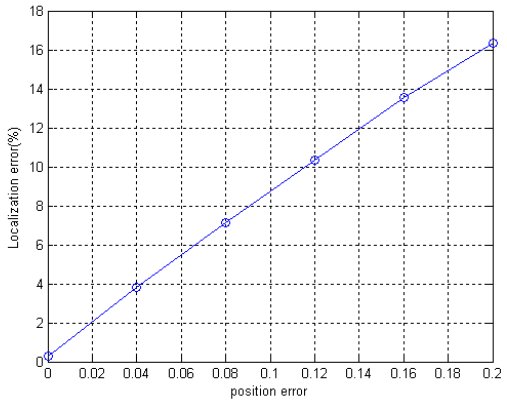


Fig. 4.

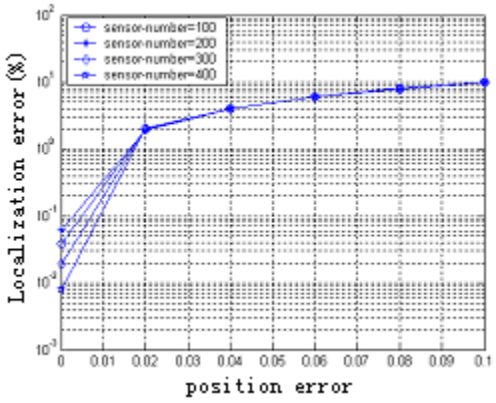


Fig. 5.

5. CONCLUSION

In this paper, a novel and cooperative localization method for the target in a wireless distributed sensor network system is presented. This new approach uses energy measurements from the individual sensors in the sensor field to detect the target first. Then the geometry shape of the target and the detection sensors is determined based on the energy measurements and an assisting and cooperative localization method is used to locate the target. Extensive simulation results show that this new approach can locate the target effectively. The proposed method only replies on the received energy of detection sensors and the position of the target can be determined based on simple geometry. So the method can reduce the computation complexity and save the energy of the nodes greatly. Thus it is a feasible and attractive localization method for sensor networks.

However, because of the complexity of the outside environment (like multipath attenuation, shadow effect and NLOS), these will affect the performance of our proposed method badly. The localization problem of considering other parameters synchronously will be studied further.

REFERENCES

[1] D. Estrin, D. Culler, et al, “Connecting the Physical World with Pervasive Networks”, IEEE Pervasive Computing, 1, Issue 1, 2002, 59-69.

[2] C. Savarese, J. M.Rabaey, et al, “Localization in distributed Ad-hoc wireless sensor networks”, Proc. ICASSP’01, Salt Lake City, UT, 2001, 2037-2040.

[3] D. Li, K. Wong, et al, “Detection, classification, tracking of targets in micro-sensor networks”, in IEEE Signal Processing Magazine, March 2002, 17-29.

[4] T. Clouqueur, P. Ramanathan, et al, “Value-Fusion versus Decision-Fusion for Fault-tolerance in Collaborative Target Detection in Sensor Networks”, Proceedings of 4th International Conference on Information Fusion, 2001.

[5] R.X. Niu, P.K. Varshney, et al, “Decision fusion in a wireless sensor network with a large number of sensors”, Proceedings of 7th International Conference on Information Fusion, June 2004.

The Information Service Alliance of Cross-regional Labor Market Based on Web Services Technologies

Dongjin Yu[1][2], Guangming Wang[1]
[1]Zhejiang Gongshang University Hangzhou, 310035, China
[2]Zhejiang Provincial Institute of Computing Technology
Hangzhou, 310006, China
Email: yudj@pop.zjgsu.edu.cn, gmwang@mail.zjgsu.edu.cn

ABSTRACT

Web Services technologies help the application coupling among the autonomous systems in the area of Electronic-Government (E-Government). This paper presents the Web-services-based architecture that supports the information service alliances for the Cross-Regional Labor Market (CRLM). Inside this architecture, the standard format for data exchanging is adopted and the application-level mechanism for distributed access control is introduced. The comparisons are also made between this Web-services-based architecture and that based on the Golden Security Project (GSP) Suggestion issued by the Ministry of Labor and Social Security, China. The information service alliance for the CRLM in the Big Hangzhou Area was successfully constructed using this Web-services-based architecture, which is proved to be more flexible than that based on GSP Suggestion.

Keywords: Web services, Cross-Regional Labor Market, Information Service Alliance.

1. INTRODUCTION[1]

Nowadays, the employment pressure in China is extremely severe. The statistics sample shows that a large number of rural labor forces pour into cities for jobs, with an increase of 9,560,000 rural migrants from 2001 to 2002 [1]. Besides, the number of the graduates from higher education almost tripled over 4 years from 1.04 million in 2001 to 3.07 million in 2005 [2,3]. Many of them find no suitable posts after graduation and eventually become the jobless. Due to the structural reforming, the laid-offs from state-running enterprises are also thrown to the labor markets. According to Chinese statistical yearbooks, the number of registered unemployed persons in urban areas rises steadily from 6.81 million in 2001, to 7.70 million in 2002, 8.00 million in 2003 and 8.27 million in 2004, with the registered unemployment rate of 3.6%, 4.0%, 4.3%, and 4.2% respectively [2,4,5,6].

However, through the analysis of the characteristics of labor markets, we can find that the problem is not due to the shortage of job opportunities, but to the information asymmetry for the posts available. How to share the employment information across the geographical boundaries in time and freely is becoming urgent especially when the labor shift is prevailing.

The chief aim of China's Electronic-Government (E-Government) project is based on the data integrations and sharing [7]. Nevertheless, various information systems are usually constructed separately, only for the specific needs locally. They often have different architectures, platforms or data formats, which bring about many so-called information silos. These separate silos of information mean management and policy decisions are often based on incomplete or even erroneous data. Citizens navigate a bewildering maze of agencies and departments to find information or submit applications [8].

Web services technologies allow applications to communicate with each other in a platform- and programming language-independent manner. A Web service is a software interface that describes a collection of operations that can be accessed over the network through standardized XML messaging. It uses protocols based on XML language to describe an operation to execute or data to exchange with another Web service [9].

The Web Services architecture is based upon the interactions among three roles: service provider, service registry and service requestor. The interactions involve the publishing, finding and binding operations. Together, these roles and operations act upon the Web Services artifacts: the Web service software module and its description.

A Web Services stack that embraces standards at each level is used to perform the three operations in an interoperable manner. Inside the stack, Web Services Description Language (WSDL) defines the interface and mechanics of service interaction. Universal Description, Discovery and Integration (UDDI) provide a mechanism for holding descriptions of Web Services. Simple Object Access Protocol (SOAP) is a simple and lightweight XML-based mechanism for exchanging structured data between network applications.

As one of the key supporting technologies for the E-Government projects, Web services technologies are quite suitable for the application integration. It's not feasible for all the E-Government participants to adapt to the same language and model. Through the standard networking protocols and XML-based data communication formats, Web services combine the advantages from both the Object Oriented methods and Web technologies, and thus have the high flexibility. In other words, Web Services technologies help integrate the separately deployed labor market applications to establish the Cross-Regional Labor Market (CRLM) information service alliance.

The work is supported in part by Zhejiang Provincial Innovation Fund for Small and Medium Technology-based Firms with grant No. 2004D40034, Natural Science Foundation of Zhejiang Province with grant No. Z104550, and Zhejiang Provincial Key Projects with grant No. 2005C21024, of China.

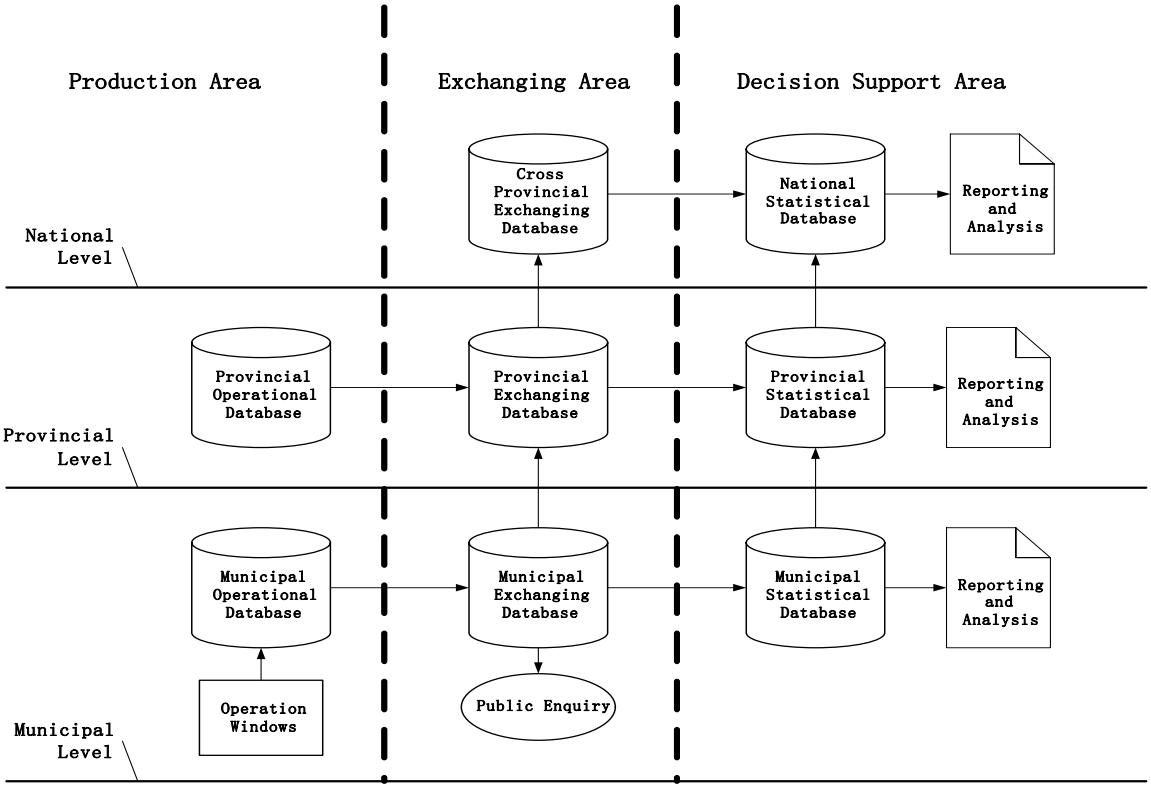


Fig. 1. The Architecture Suggested by National GSP

2. INFORMATION SERVICE ALLIANCE AMONG GOVERNMENT AGENCIES

For the service-oriented enterprise, building information service alliances is a novel way to promote its capabilities of customer service and market competence. Among some industries, information service alliances have already come into being, such as hotel alliances and travel agency alliances. Meanwhile, for the area of public administration, there also exists the demand of setting up new information service alliances. Many government agencies have already devoted large amount of resources to build their own service system for the whole society. Under such circumstances, like the service enterprises, the service agencies in government are now becoming anxious to set up their coalitions, by which the geographically separated government agencies could combine their service capacities and thus expand their service ranges. To some extent, this is a strategy decision, helping break through the limitation of time and space for their service activities.

According to the different organizational structures, the service alliance could be divided into the mutual alliance and the 3rd-party alliance [10]. Inside the mutual alliance, each member usually does not change their original structures, but promotes its existed service system to provide services for other members. On the other hand, its service could be supported and extended by other members within the same alliance by agreements. Thanks to the information technology, this kind of mutual alliance often takes the virtual form. The members in the 3rd-party alliance, however, set up the independent entity to integrate their service systems, and therefore share their service resource and profits.

3. THE ARCHITECTURE OF CRLM SERVICE ALLIANCE

Golden Security Project (GSP), approved by the State Council of China in Aug 2003, is one of 12 national E-Government projects [11], specifically for the information system construction of labor and social security management nationwide. Based on municipal networking and shareable data servers, it has developed Labor Resources Management System and Social Insurance Application System, to make the realization of data exchange with outer system, and finally march towards its grand goal of administrative centralization at provincial level with characteristic of data centralization, application centralization, physical centralization and security centralization. To guarantee the quality and its consistency in different places, the National Ministry of Labor and Social Security (MOLSS) issued the technological suggestion for the construction of GSP [12].

3.1 GSP Suggested Architecture

GSP suggests the nationwide information system be based on three tiers from the national center, towards provincial centers and municipal centers. The distribution strategy is as the following and illustrated in Fig. 1.

(1) To establish the centralized operational database server, namely operational data servers, in all municipal cities. Operation data servers store the data of persons and enterprises joining social security insurances funds and support local front-end operations. Municipal exchanging data servers are based on operational data servers, with the operational data summarized and filtered, and are used for data swapping for outer systems. Another data servers, called municipal statistical data servers, target to the support of local decision making, with the data extracted from

exchanging data servers.

(2) To establish the centralized provincial data exchanging server in each province. The data center in the provincial city filters and summarizes the data from their local operational servers, and merges the data from municipal exchanging data servers of each city in the same province. These data are used for data exchanging among cities inside the same province. The provincial statistical data server is constructed through its provincial exchanging database, to support the decision making provincial-wide.

(3) To establish the cross-province data exchanging server and national statistical data server in the Information Center of MOLSS. The former is used for data exchanging across all the provinces. The latter helps data analysis and supports decision-making nationwide.

This pattern of architecture was put forward when Web services technologies was still ambiguous. It is easy to implement, and helps the data summarization. However, the actual result is not as successful as expected due to the following reasons.

(1) The pattern has the top-down hierarchical structure, from the central government, towards provincial level and municipal level. Whenever the construction of upper site has not been finished, or fails to operate, all the following lower sites would fail to cooperate. In other words, the municipal sites are not capable of direct interaction without help of corresponding upper sites.

(2) Unreal-time interaction limits the usability of the whole system.

(3) Based only on data exchanging, but not applications coupling, the whole system lacks flexibility and extensibility.

Strictly speaking, this pattern of architecture is not for the information service alliance. It needs to build the specific service site on the upper level and thus has some features of 3rd party service alliance. It suits for the administrative operations by means of commands, such as individual security fund transferring, but not for the information service operations, which needs peer-to-peer collaborations.

Although the CRLM information services, such as job introducing, professional training and information searching, are largely demanded, there almost exist no successful service alliances based on this architecture. While Web services technologies are getting matured, this suggested the following one might substitute architecture.

3.2 Web Services Based CRLM Service Alliance Architecture

Thanks to Web Services technologies, the coupling of geographical separated applications can be easily achieved. Its main strategies include:

(1) The business services for outside usages should be provided based on legacy systems. These services could invoke or reuse the existed business logic, and make access to their local databases.

(2) The access control, including both the service control and the data control, should be provided to raise the security level. The outer users are allowed to invoke the services or access the data only when they are granted to do so.

(3) The clients search the services in the UDDI registry using the round-robin algorithm. After the successful searching, the clients bind and invoke the services. The server sides implement the Web services.

(4) The application couplings are made inside the government specific-purpose network, which targets for the information management of labor markets. If outward services are necessarily, the connection to the public network should be isolated with the inner specific-purpose network due to the security considerations by the Demilitarized Zone's setting. On the other hand, the virtual service alliance could be constructed under the public network. It provides the public services that couple with separated applications through its Web site.

This pattern of architecture has the following advantages over the suggested one by GSP.

(1) The application integration is achieved directly without the help of central control, which makes the construction more flexible.

(2) The server sides can define the services and restrict their access rights on the demand. For instance, the posts available for one region inside the service alliance can be selected one by one and released especially for job seekers outside the region.

(3) The operations are handled automatically and in real time. This kind of application coupling based on Web Services technologies is quite suitable for the integration of autonomous systems, and helps construct the virtual mutual service alliance. It focuses on the flexible and open information services with easy interaction. It neither changes the legacy systems, nor needs to establish the extra 3rd party alliance entity. It could be used to provide the information services such as cross-regional job introducing, cross-regional professional training, etc.

The software architecture is illustrated in Fig. 2. Any autonomous system in a region is regarded as both the requestor and provider of some certain services. In order to integrate the service requesting and responding, the architecture provides a uniform Web Service proxy. Because of possible heterogeneous databases, it includes a dynamic engine in charge of data transferring by which the data with different formats can be dynamically exchanged. The access to the Web services should pass the arbitrary procedure of identity authentication and rights control. Different meta-functions and data owners define rights. The rights based on meta-functions are divided till the atomic ones, such as buttons or URL links, while the rights based on data owners are divided till the database records.

The logic layer is further divided into the layer of data logic and the layer of business logic. In order to promote the system extensibility, the interaction with the legacy systems is restricted in the data logic layer, using the dynamic data-transforming engine.

4. DATA FORMAT AND DATA EXCHANGING

MOLSS has presented the specification for the index and code system of labor and social security information system. The specification gives the definition of 865 indexes, each having its name, type and length [13]. For instance, to describe the reemployment premium, there exist four indexes, namely the type of policy, the date of audit, the premium amount and the premium deadline. The index name consists of 6 digits, separated by 4 layers. The left most digit defines the index level, i.e. the national level, the provincial level, the municipal level or the county level. The second left most digits specifies the business types, such as type of labor market and employment, type of professional training and qualification, type of unemployment insurance. The following digit gives the index's entity type indicating

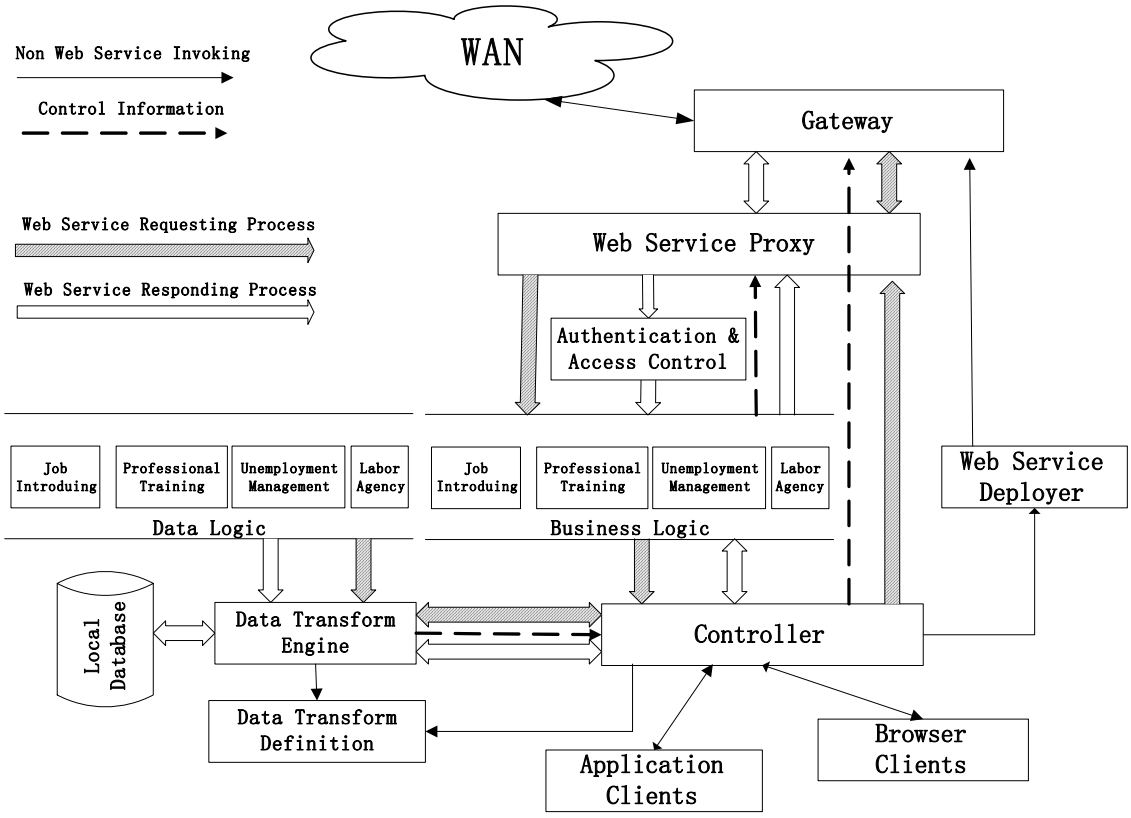


Fig. 2. The Software Architecture Based on Web Services

the individual one or that of the unit. The right most 3 digits are simply the serial number. Besides, the code system includes the code tables for 190 types of codes. For instance, the code of unemployment reason is discriminated among 8 types, including unemployment after graduation, unemployment because of labor contract termination with employed enterprise, unemployment because of dismissing by government agencies, unemployment because of land losing for farmers becoming citizens.

However, the local un-uniform index and code system challenges the system integrations. The national specification for the index and code system should be completely complied with when establishing the service alliance based Web Services technologies. Each autonomous system builds its mapping mechanism of data dictionary exchange for the mutual mapping of indexes and codes between the local data and remote data. The data of standard format is transferred among applications encapsulated by SOAP.

5. SYSTEM IMPLEMENTATION

Currently many independent labor market information systems, as parts of welfare facilities, have already been constructed, especially in the east and south urban areas of China. These systems typically have their data centers, and support operations of job introducing, employment and unemployment management, re-employment service, professional training, unemployment security fund collection and payment, etc, based on the network covering their local regions. However, these systems often have different technological architectures, such as J2EE and .NET, and probably have different data formats.

The authors and their research team have successfully implemented this kind of information service alliances. Using the new architecture discussed here, the labor market information service alliance for the Big Hangzhou Area, which locates on the low reaches of Qiantang River in southeast China, has been constructed. The alliance provides CRLM information service for more than 6 million people living in 6 districts, 2 counties and 5 county-level municipalities in the Yangzi Delta area. All these districts, counties and municipalities have their own independent information systems with different database management systems such as Oracle or Sybase, and different application software based on PowerBuilder or Java.

Each independent system in the Big Hangzhou Area connects with others through the government specific-purpose network. The special UDDI server is built for the registry of services provided by all the independent systems. If anyone inside the Big Hangzhou Area wants to seek job or query unemployment information, the alliance helps retrieve the registry information via the UDDI server, bind other remote systems if necessary, and finally invoke the remote service. All these processes undergo with no user interferences. In other words, there is no need for users to know where the service is invoked and how it is implemented, so long as the independent systems inside the alliance have registered their services needed to publish and the user request has passed the authentication checking.

Fig. 3. The Screen Shot of Cross-Regional Post Inquiry

The alliance platform is fulfilled using the Apache Axis, the open source framework for creating and deploying Web services applications [14]. It is browser enabled, and is compatible with Oracle Server, MS SQL Server, Sybase Adaptive Server, or any other relational databases providing the JDBC interfaces.

As Fig. 3 showed, upon request, the posts information wanted in different regions are listed in the same page, with their post names, age limitations, numbers wanted and also the deadlines, etc.

6. CONCLUSION

The fast growing Web services technologies are most suitable for application couplings among the autonomous systems. Due to vast land of the country and the hierarchical government structure, the construction of E-Government projects in China is usually distributed and independent with each other. Using the up-down administrative orders to force the interconnection of these systems has the practical difficulties. The direct horizontal interconnection of systems under different jurisdiction encounters even much more obstacles. Therefore, the application coupling without the vertical control has more realistic meanings.

The solution of building the CRLM information service alliance based on Web services technologies focuses on the standard of data formats, the distributed transactional control and also the security consideration. It can be adjusted to fulfill the implementation of whole GSP. One ideal scenario is as the following. The job seeker logins one Web site and submits his or her post requirements for just one time. The Web site automatically searches its local database and, if necessary, invokes other service providers in the service alliance the Web site belongs to. Finally, the results from local searching and remote service invoking are combined and presented to the requesting user. All these are done simultaneously and transparently. Thus the user no longer needs to try other sites one by one.

To construct the information service alliance is the key step for promoting the platform's effectiveness. The service alliance integrates the various resource advantages of each member and realizes the service sharing through distributing technologies such as Web services.

7. REFERENCES

[1] R. Mo, Employment Report of China: 2003-2004, China Labor & Social Security Publishing House, Beijing, 2004

[2] National Bureau of Statistics of China, China Statistical Yearbook: 2002, China Statistical Press, Beijing, 2002

[3] National Bureau of Statistics of China, Statistical Communiqué of the People's Republic of China on the 2005 National Economic and Social Development, http://www.stats.gov.cn/tjgb/ndtjgb/qgndtjgb/t20060227_402307796.htm, Feb 2006

[4] National Bureau of Statistics of China, China Statistical Yearbook: 2003, China Statistical Press, Beijing, 2003

[5] National Bureau of Statistics of China, China Statistical Yearbook: 2004, China Statistical Press, Beijing, 2004

[6] Ministry of Labor and Social Security of China, China Labor and Social Security Yearbook: 2005,

China Labor & Social Security Publishing House, Beijing, 2005
[7] National Research Center of Information Security Engineering Technology, The General Design and Technological Implementation of E-Government. Publishing House of Electronics Industry, Beijing, 2003: 23~28, 166~169
[8] R. Todd, On Demand Government: Continuing the e-government Journey, IBM Press, Sept. 2004
[9] H. Kreger, Web Services Conceptual Architecture, http://www-28.ibm.com/developerworks/webservices/newto/websvc.html. May 2001
[10] J.Y. Ji, W.J. Zhong, et al, "Service Alliance Strategies of Enterprises in Electronic Commerce", Journal of Industrial Engineering and Engineering Management, Publisher of Zhejiang University, 16(2), 2002, 79~82
[11] National Informationization Leading Group, The Guideline for National E-Government Construction suggested by National Informationization Leading Group, http://www.china.org.cn/chinese/zhuanti/283233.htm, 2002
[12] Ministry of Labor and Social Security of China, The Suggestion of Golden Security Project: the 1st Phase, May 2003
[13] Ministry of Labor and Social Security of China, The Index System of Labor Market Management Information System: the Business Part (LB102-2005), Departmental Information File No. 2005-10, 2005
[14] http://ws.apache.org/axis/

Research and Implementation of SNMP and AgentX based Network Management Scheme in ForCES Router

Youcheng Tong, Weiming Wang
College of Information and Electronic Engineering，Zhejiang Gongshang University
Hangzhou, P.R.China, 310035
Email: tongyc@pop.hzic.edu.cn , wmwang@mail.zjgsu.edu.cn
And
Fenggen Jia
National Digital Switching System Engineering & Technological R&D Center
Zhengzhou, Henan province, China
Email: fgjia@mail.zjgsu.edu.cn

ABSTRACT

An efficient, extensible and flexible network management scheme, which is based on SNMP and AgentX protocol, was proposed for next-generation ForCES based routers. With the analysis of two challenging problems of network management for the highly distributed ForCES based routers, their respective solutions were proposed. In the solutions, an AgentX based MIB variables distribution mechanism is applied, and a virtual-interface based SNMP packets transmission mechanism is used, which greatly ease the deployment of the ForCES application modules. The initial test has proven the validity and the feasibility of the scheme.

Keywords: SNMP, ForCES, AgentX.

1. INTRODUCTION

Analysis on trend of the demand for network devices reveals that open programmability would be an important character of next generation network devices. Owing to continuous emergence of new protocols, new algorisms, new services, a great deal of analysis on network tendency showed that hardware interoperability of next generation network devices could reach the degree of razor blades, software interoperability could reach the degree of APIs. The open programmable router system of ForCES[1] (Forwarding and Control Element Separation) caters to this new development tendency and it satisfies the open, extensible and programmable demand of next generation network. Therefore, it is an imminent task to find out an efficient and highly flexible network management scheme for ForCES based routers.

SNMP is a simple and widely supported network management protocol. However, with the increment in quantities of network, the enlargement of scale and diversities of devices, especially the flexibility of opened fabric poses great challenges and more demands. Extensible SNMP agent protocol-----AgentX broadens basic network devices, essential parts and application capability, with flexible expandability makes it possible to bring a fresh added network equipment or part into the scope of managed object in a modular way.

Project sponsored by National 863 Project (2005AA121310) , NSF China (60573116), and Zhejiang Key Sci&Tech Project (2005C21013)

The paper combines with National 863's theme research project of router programmability "The development prototype of ForCES router" and advantages of SNMP & AgentX, brought forward a network management frame based on SNMP and AgentX protocols. The author intended to discuss it from the point of SNMP management protocol and probed into the application of SNMP& AgentX protocols in ForCES router, and has made the beneficial exploration on the network management in the respect of key technology.

2. NETWORK MANAGEMENT IN FORCES ARCHITECTURE

ForCES are a working group under routing area of the Internet Engineering Task Force. It specializes in the IP router and protocol in open programmable networks, and now it becomes one of the most attractive workgroups on open programmable networks. The basic idea of ForCES is to separate the Control Elements and Forwarding Elements of an IP router. The ForCES protocol regulates the communication between CEs and FEs. An FE is combined by possibly multiple logical functional blocks (LFBs)[2]. A router employing ForCES protocol is called ForCES router. Figure 1 shows the architectural diagram of the ForCES router.

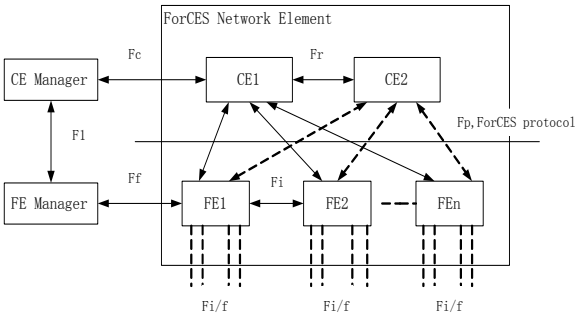


Fig. 1. ForCES Architecture Diagram

In the ForCES router, one CE can control one or more FEs. The interface between CE and FE is labeled as Fp ("p" meaning protocol) reference point. Fp reference points can be implemented ethor by Single-hop network like Ethernet or by Multi-hops network like Internet. The FE external interfaces are labeled as Fi/f reference points, which are the external interface to the outside world.

The core network device is router. ForCES router management includes configuration and performance management. The AgentX protocol combined with SNMP suit well the task of management of distributed system like the ForCES router. The SNMP protocol[3] is used to convey management information between devices manager and agent. In the ForCES Framework, two difficulties must be resolved. The first was how to correlate conceptual MIB and actual MIB, which was to say how to distribute MIB variables in order to improve efficiency, reliability and flexibility of the network management. The second was how to transmit the SNMP message and access the FE messages managed by CE. So, MIB distribution mechanism was put forward to solve the first difficulty and virtual interface based transmission mechanism of SNMP packet to settle the second one.

3. RESEARCH ON KEY TECHNOLOGIES OF NETWORK MANAGEMENT IN FORCES ROUTER

3.1 AgentX based MIB variables distribution mechanism

The key of understanding network management framework is to be familiar with the network management information model [4]. To form a better network of unified management is to build a better information model. It is divided into two aspects to describe the management information: the structure of management information (SMI) and management information base (MIB). SMI not only defines the structure of management information which is used to distinguish the different management information elements and different management information, but also defines the relationship between management targets. MIB which defines a specific management object is the logic integration of managed network resources and the conceptual base of management information. The network management activities are carried out through accessing and operating the management object of MIB. MIB is the heart of the management system, which provides interfaces for all functions of the management system. Therefore, the distribution of MIB in specific controlled equipment will affect the efficiency and flexibility of management system or agency system directly.

Considering ForCES router is made up by multiple CE and FE, their respective mandates are all very heavy. CE can be divided into controller and server in order to reduce the processing burden of server and the compatible of multi-CE and achieve load balance. The server is used to execute the routing protocol services and SNMP services. The network management agent based on AgentX protocol is exactly suitable for the ideal environment of CE for the functional distributed structure. CE server is the master agent and CE controller is the subagent, which are communicated through AgentX protocol. The master agent and the subagent of MIB safeguard themselves respectively for the nearest principles so as to reduce resources and raise the efficiency of management. So the agent procedures of CE is much easier to develop, its structure is clearer, also the management applications are easy to expand, and the management process would become much more normative. Thereby, it should work out the expandability and the flexibility of MIB dynamic loading or unloading better. When one of CE goes out of order, the CE redundancy can also be added to achieve the cooperation and interaction between different management systems.

3.2 Virtual interface based SNMP PACKET Transmission mechanism

Architecture of ForCES router, is showed one unitary function equipment externally, but is CE and FE separate each other and cooperative distributed architecture internally. So implementation network management of ForCES router is actually one of agent network devices. Following SNMP work model, manager and its request message is belong to external device and external message. With a view to the role of CE and PE, CE as a “manager” of control element, and FE as a “managed” of forward element, it is necessary via CE that the network manager is able to access to the information of FE. However, due to the interface of ForCES router is in the FE, external network data must pass FE interface and be transmitted accordingly, and the network manager only send or receive data from FE interface. The manger send a request message to a FE interface, it should return a response message from the original route and interface. In this way can capability estate of system be monitored and controlled.

By above analysis, virtual interface technology can solve the mapping between FE physical interface and CE logical interface. The virtual interface we supposed is that the component itself is not provided with this physical interface, it depends on software to build a logical interface. In order to let CE distinctly know the physical address, virtual interface technology (VIDD) was be adopted. Creating virtual interface in CE, building corresponding relation between virtual interface and FE real interface, adding virtual interface information on the top of every packet that pass in and out software of network management, packet can be sent from appointed FE interface and make software of network management know where the packet come from by information of virtual interface. The application of traditional centralized router may be utilized by ForCES distributed architecture in virtual interface technology, and it also provide enough flexible for implement of its network management.

ForCES protocol message are composed of control message and redirect message. The definition of redirect packet is other devices send packets to ForCES router itself, special treatment should be made by ForCES router, or ForCES router sends packets of other devices (generally occur in the application of functioning CE), under the system of distributed router like ForCES, these two types of data packets should transmitted from FE to CE or CE to FE by the way of ForCES protocol. ForCES router packet mainly includes router protocol and SNMP packet. When FE receives SNMP packet, through judging, it is required to contains two fields, processed by CE, then head of virtual interfaces (VIP) is added to IP data header, after that it is encapsulated in ForCES protocol of redirect message and sent to CE. VIP which is FEID and portID among which, arrival of FEID’s marked packet or destination is FE, arrival of portID’s marked packet and destination port. Main function of VIP header is to shield traits of distribution of hardware like ForCES, relation between virtual interface and real interface, enhance normal function of application in upper level of CE.

When CE controller receives redirect packet from FE, which is parsed by ForCES protocol, then SNMP packet is acquired, what is needed to be tackled is to take out the added packet of VIP header from ForCES redirect message, send to CE server through IP tunnel. Then after CE accepts the packet, according to FEID and portID in VIP header, virtual interface need to be delivered packet is found, then SNMP packet acts through TCP/IP protocol stack, in the end it is delivered to application program, after SNMP effects application management of this SNMP packet, certain reaction occur, which demanded to add certain VIP header to virtual network interfaces card, then is sent to CE controller by IP tunnel. When CE controller receives SNMP response packet from CE server, it is encapsulated to ForCES redirect message, after which, is sent to FE, after FE receives this message, VIP header is gotten rid of. Owing to VIP header information, port for sending out packet is found, then through this port, SNMP response packet should be sent to the network manager. The transmitting process of SNMP packet effecting on ForCES router is demonstrated in figure 2.

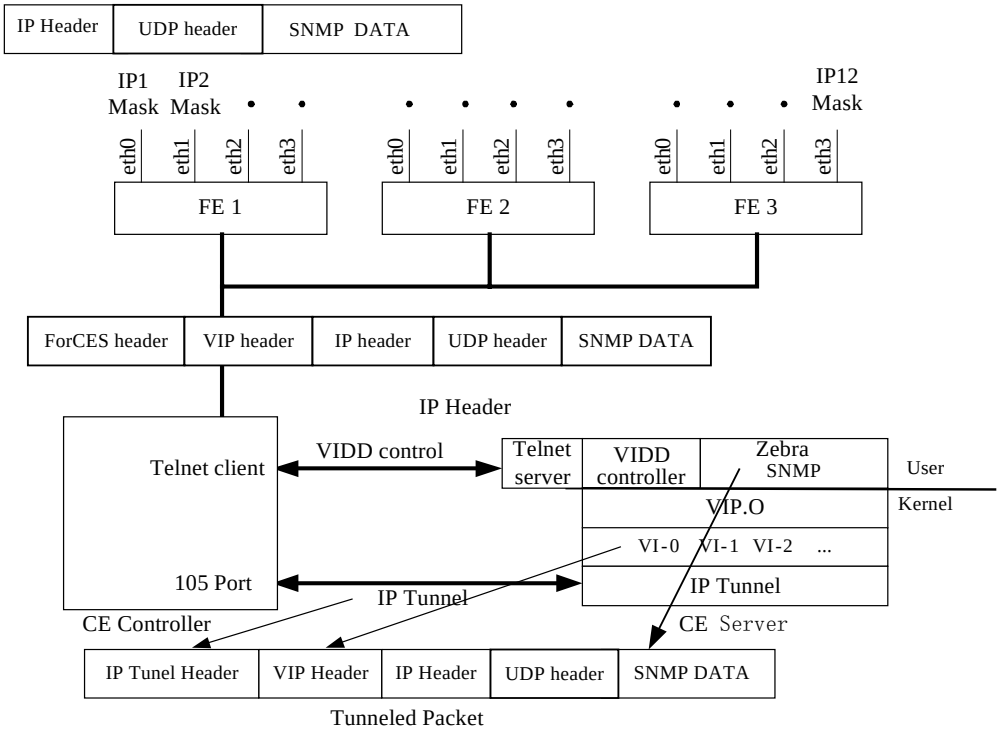


Fig. 2. SNMP packet encapsulation and transmission in ForCES router

4. THE SCHEME OF NETWORK MANAGEMENT IN FORCES ROUTER

4.1 The network management framework of ForCES router

Based on the above analysis, the ForCES router network management framework has been presented as shown in Figure 3. The entire structure is composed of three parts ----- Manager, FE and CE, the CE and FE constitute the ForCES router. The manager and ForCES router communicated with SNMP protocol, but the communication between CE and FE with ForCES protocol. Two PCs distributed Linux system and Windows system on CE, which was stationed between subagent and master-agent, communicated with AgentX protocol each other. The function of MIB interface module is the relation between conceptual MIB and real MIB for the implementation of MIB node query, configuration and trap message displaying in manager.

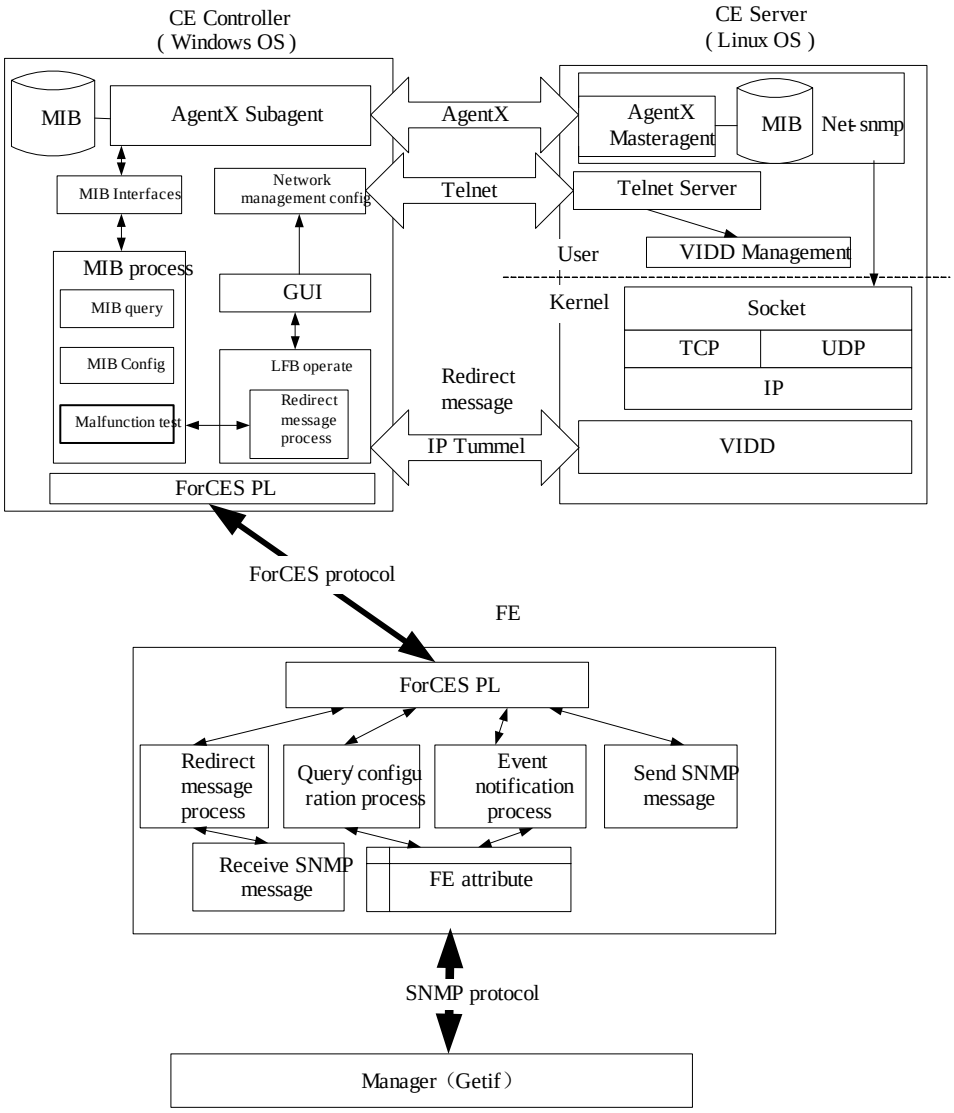


Fig. 3. Architecture of network management in ForCES router

4.2 The MIB variables operations Insight

CE acts as ForCES router's "manager", which manage attribute and LFB in FE by send and receive various ForCES message. These ForCES message is classified as follows[1]:

- ForCES Query and Query Response message;
- ForCES Configuration and Configuration Response message;
- ForCES Event notification and Event notification response message;
- ForCES Redirect message;
- Association Setup message, Association Setup Response message and Association Teardown message.

1) MIB query

When the manager user hopes to read ForCES variables, it would send a SNMP-Get-Request-PDU to the FE interface. If you want to get it response message, you must go as follow:

Firstly, FE sends the get-request message from manager through the SNMP packet transmission mechanism and ForCES protocol redirect mechanism to CE controller. CE controller receives the redirect message, judged the packet type by ForCES protocol parse as SNMP packet, sending to CE Server (master-agent) through IP tunnel technique.

Secondly, CE server receives the message; SNMP-Get-Request-PDU is mapped to AgentX-Get-Request-PDU owing to the function of administrational scheduler, and then sends the message to CE controller subagent mould through AgentX protocol[4].

Thirdly, CE controller subagent mould receives the packet, if the value of MIB is gotten in FE, CE sends ForCES query message to FE, FE carried the task of query out construct the query response message to CE controller. CE controller subagent constructs AgentX response message to CE server, then translates it to SNMP response message, send it to manager by the original way. So the whole process of query is finished. Figure 4 describes SNMP query process.

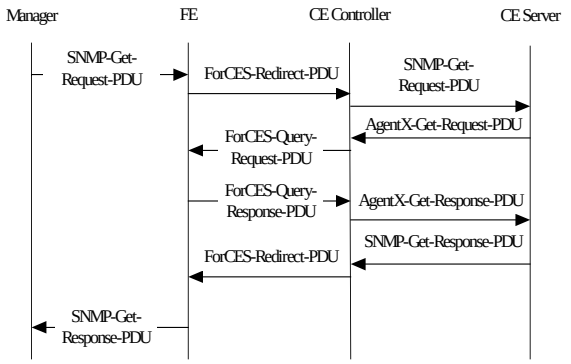


Fig. 4. Successful SNMP query sequence diagram

2) MIB configuration

If the user of management station wants to set the ForCES router variable, they can transmit SNMP-Set-Request -PDU to one of the FE interfaces, its treating process is similar to the MIB query. But the CE controller subagent module needs two stages to complete a set operation when receiving an AgentX-Set-Request-PDU:

- Preliminary stage: set operation test

After CE server master agent module receives SNMP-Set-PDU, it determines the respective index allocation MIB region and transmits one or many agentx-TestSet-PDUs to the corresponding region session. When the subagent receives the TestSet request, it abstracts the requesting establishment value, tests whether the Set operation can succeed or not. Possibly, there are many requesting establishment variables in a agentx-TestSet-PDU, then only all variables in PDU test successfully, the CE controller subagent module returns the success test response, otherwise it returns the error message.

- Second stage: Truly set operation execution

In case all transmitted TestSet-PDU return no-error response, the CE server master agent module transmits the CommitSet request to subagent to carry on the true establishment operation. Through the MIB interfaces function, CommitSet request transfers news dispose module of CE controller, lets FE carry out the CE controller Set order: If execution fails, it structures CommitSet operation failure response value SET_TO_FE_ERROR, to observe the disposition where the tie lies. And AgentX-UndoSet-PDU will be delivered to all the conversations which have transmitted AgentX-CommitSet-PDU in the preliminary stage by the agent, to inform the subagent to cancel the former CommitSet operation, AgentX-CleanupSet-PDU is transmitted to other involved sessions, meaning the end of Set operation; otherwise, it structures AgentX-Set-Response-PDU and returns to the CE server master agent.

3) Transmission of trap

When the event notification subscribed by FE happens, CE is notified by ForCES control message through the control path. CE controller analyzes the packet and gets the event notification, and sends AgentX-Notify-PDU constructed by subagent through the processing module connected to MIB's interface to master agent on CE server. The master agent receives the notification message and returns the AgentX response message to the subagent. At the same time, the master agent constructs a SNMP-Trap-PDU based on the Agentx-Notify-PDU it receives and sends it to CE controller's management process. Meanwhile, CE controller sends it to FE as a redirect packet and FE sends it to the manager. Thus, the manager can be aware of FE's exception situation.

5. TEST

The main frame of the testing system of the network management can be described in the picture of figure 5. The ForCES router is consisted of 4 parts, a host with OS of Windows CE, a CE server based on Linux, 3 FEs running on IXDP embedded platform with 2-3 of which running on IXMB2401 and the remainder running on IXMB2851, and a switch via which the CE and FE are connected each other. The Getif as MIB browse is appointed to display MIB node message.

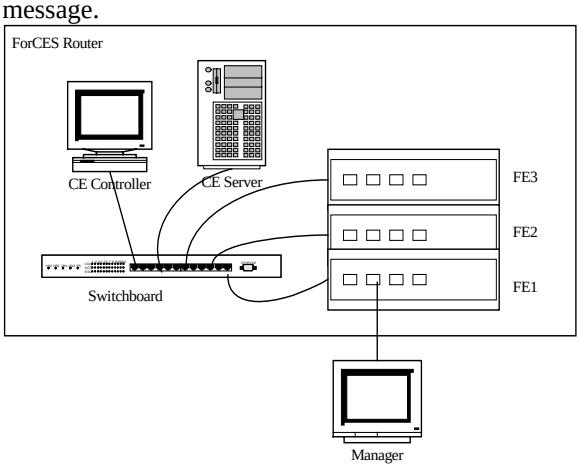


Fig. 5. Deploy diagram of the test bed for network management in ForCES router

It has been tested that the SNMP agent of ForCES router can configure and query the information on the FE, and it also can notify when exceptions happen. This shows that the management system is feasible and effective.

6. CONCLUSION

Currently, computer network evolve rapidly, router has become the core of the network devices. How to enhance highly efficient management of next generation network devices is a pressing topic. This paper bring forward a realizable, efficient, extensible and flexible network management scheme based on SNMP and AgentX for next-generation ForCES

based routers, this work is known to be the first in the field and has practical and academic significance. However, its theoretical meaning may get over realistic meaning, and it gives a kind of insight and reference to the management of distributed ForCES based systems.

7. REFERNCES

[1] A. Doria, ForCES Protocol Specification, ternet-Draft (IETF ForCES WG Document), work in progress, 2006.

[2] H. Khosravi, T. Anderson (editor), "Requirements for Separation of IP Control and Forwarding", RFC 3654, Nov. 2003, http://www.ietf.org/rfc/rfc3654.txt?number=3654.

[3] J.Case, M. Fedor, M. Schoffstall et al.“ A Simple Network Management Protocol (SNMP)” May, 1990, http://isc.faqs.org/rfcs/rfc1157.html.

[4] M. Daniele, M. Wijnen, et al, "Agent Extensibility (AgentX) Protocol Version 1", RFC 2741 [DB/OL], http://www.faqs.org/rfcs/rfc2741.html, Jan. 2000.

Reference Nodes Selection and Placement Mechanisms for Localization in Wireless Sensor Networks

Xingyu Pi, Hongyi Yu
Communication Engineering Department, Information Science and Technology Institute
Zhengzhou City, Henan Province, P.R.China
Email: {psingyu@163.com, maxyucn@sohu.com}

ABSTRACT

One of the important research issues in wireless sensor networks (WSNs) is the reference nodes selection and placement problem for localization. In this paper we analyze the localization error and propose two theorems with respect to the different geometric topologies of the terns of reference nodes (whose positions are known). Then the selection and placement mechanisms of reference nodes are proposed with respect to the two theorems. Extensive simulation experiments show that it is helpful to improve the localization performance of sensor nodes using the reference node selection and placement mechanisms, and the mechanisms do not need too much energy and computation complexity.

Keywords: sensor networks; radio localization; selection; placement; degree of collinearity.

1. INTRODUCTION

Many applications of sensor networks require that sensor nodes be aware of their absolute or relative (with respect to other nodes) location. The location information can be used to accomplish both application specific tasks and networking functions efficiently. For example, a sensor node operating in a monitoring or tracking system is typically required to not only report that an event of interest has occurred but is also responsible for reporting the location of the event. As such, the node must be capable of automatically estimating its current position. Localization in sensor networks is required to support location aware applications, object tracking, location based routing [1,2], coverage management [3] and collaborative signal processing.

Current research on sensor network localization focuses on two issues: accuracy and error analysis [7,8]. Many localization systems of sensor networks are proposed [4-6], but we see little of the research on the selection and placement of reference nodes for localization in sensor network. The main research issues of the paper are to analyze the localization performance with respect to the different geometry topology of the tern of reference nodes, and obtain the reference nodes selection and placement mechanisms for the localization application in sensor networks to make localization error minimum.

This paper is organized as follows: Section 2 analyzes the localization model and error. Section 3 describes the reference node selection and placement mechanisms. Section 4 describes the simulation experiment results. Section 5 concludes the paper.

2. LOCALIZATION MODEL

2.1 Problem Formulation

In generality, to perform the position estimation, a node assumes to locate at the circumference centered in reference coordinates with a radius equal to the estimated distance from a reference node. With this assumption, the located node needs at least three reference nodes (in a 2-D environment, 4 in 3-D) to obtain a unique position. But the errors occurring in the distance estimation result do not cross into a unique point. Considering a 2-D environment, if the tern of reference nodes is perfectly aligned the intersection between the three circumferences is double and the node can not estimate its position. Such the case is often remarked in many papers [4-7].

2.2 Parameter Definition

Assume that the three internal angles of a triangle consisted of three reference nodes are α_1 α_2 and α_3 respectively. Based on the cosine rule,

$$\alpha_1 = \cos^{-1}\left(\frac{d_{1,2}^{\ 2} + d_{1,3}^{\ 2} - d_{2,3}^{\ 2}}{2d_{1,2}d_{1,3}}\right),$$

$$\alpha_2 = \cos^{-1}\left(\frac{d_{2,1}^{\ 2} + d_{2,3}^{\ 2} - d_{1,3}^{\ 2}}{2d_{2,1}d_{2,3}}\right),$$

$$\alpha_3 = \cos^{-1}\left(\frac{d_{3,1}^{\ 2} + d_{3,2}^{\ 2} - d_{1,2}^{\ 2}}{2d_{3,1}d_{3,2}}\right),$$

where $d_{i,j}$ is the distance between the *i-th* and j-th reference node, $i, j = 1,2,3, i \neq j$.

Define the degree of collinearity (DC) of the tern of three reference nodes as $DC = \max_{i=1,2,3}\{\alpha_i\}$. It is easy to know that the range of DC is $60 \le DC \le 180$. At the value 180 corresponds a perfectly aligned tern; while at value 60 corresponds an equilateral tern.

2.3 Localization Method

Assume that the positions of the located node and reference nodes are (x_p, y_p) and (x_i, y_i), respectively. The estimated range between the located node and reference nodes is $d_{i,p}$. The system of equations can be written as $(x_p - x_i)^2 + (y_p - y_i)^2 = d_{i,p}^{\ 2}$, $i = 1,2,3$.

Finding the solution to the above equations set:

$$x_p = \frac{A \cdot B^T}{2D \cdot E^T}; y_p = -\frac{A \cdot C^T}{2D \cdot E^T} \quad (1)$$

Where $A = \begin{bmatrix} \|r_1\|^2 - d_{1,p}^2 & \|r_2\|^2 - d_{2,p}^2 & \|r_3\|^2 - d_{3,p}^2 \end{bmatrix}$,

$B=[y_3-y_2 \quad y_1-y_3 \quad y_2-y_1], C=[x_3-x_2 \quad x_1-x_3 \quad x_2-x_1],$

$D=[x_1-x_2 \quad x_2-x_3], E=[y_3-y_2 \quad y_1-y_2].$

Assume that the geometry of three reference nodes is shown as Fig. 1.. Without loss of generality, assume that the angle $\angle 123$ is the maximum internal angle, $x_1-x_2=-d_{12}\sin(\theta_1)$, $x_2-x_3=-d_{23}\sin(\theta_2)$, $y_1-y_2=-d_{12}\cos(\theta_1)$, $y_3-y_2=-d_{23}\cos(\theta_2)$.

Then the equations set (1) can be rewritten as follows:

$$\begin{cases} x_p=\dfrac{A\cdot B^T}{2d_{12}d_{23}\sin(\theta_1+\theta_2)}=\dfrac{A\cdot B^T}{2d_{12}d_{23}\sin(DC)} \\ y_p=-\dfrac{A\cdot C^T}{2d_{12}d_{23}\sin(\theta_1+\theta_2)}=-\dfrac{A\cdot C^T}{2d_{12}d_{23}\sin(DC)} \end{cases} \quad (2)$$

When the geometry of the three reference nodes is shown as Fig. 2., then

$$\begin{cases} x_p=-\dfrac{A\cdot B^T}{2d_{12}d_{23}\sin(DC)} \\ y_p=\dfrac{A\cdot C^T}{2d_{12}d_{23}\sin(DC)} \end{cases} \quad (3)$$

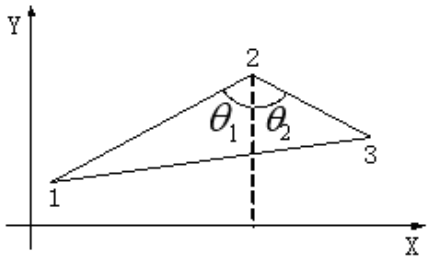


Fig. 1.

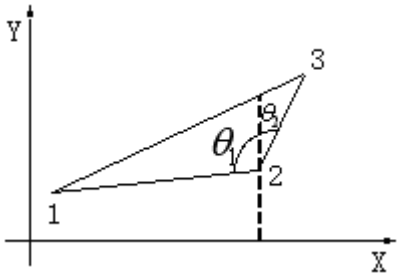


Fig. 2.

2.4 Performance Analysis

When the located node has collected all needed data from reference nodes, it uses equations (2) to estimate its positions. From the equations we can know that the solution do not exist if $DC=180$, that happens if terns of the reference nodes is perfectly aligned. Then we can not estimate the positions of the node to be located. When $DC\neq 180$, two theorems of the localization performance are derived as follows.

Theorem 1: Given certain range measurement error, the minimum localization error is obtained when the located node is at the centre of the circumcircle of the triangle consisted of tern of the reference nodes.

Proof: For convenience, we model the range measurement between the reference nodes and the located node as the real range added a range measurement error e_r. With respect to equations (2), the position estimation error of the located node can be written as equations (4), where d_{real} the real range value.

Without loss of generality, we mainly analyze the localization performance of X-coordination, and the analysis of Y-coordination can reference the case.

For a certain range error, to minimize the localization error, that is to minimize equations (4), where d_{12}, d_{23} and DC are constant. For the estimation error of X-coordination, the problem is formulated as follows:

$$\min\left\{\left|d_{(real)1,p}(y_3-y_2)+d_{(real)2,p}(y_1-y_3)+d_{(real)3,p}(y_2-y_1)\right|\right\}$$

Where y_1, y_2, y_3 are constant. It is easy to know the minimum value of the above objective function is zero when $d_{(real)1,p}=d_{(real)2,p}=d_{(real)3,p}$. So the position with minimum localization error is the centre of the circumcircle of the triangle. **End.**

Theorem 2: Given certain range error, the located node is within the triangle. The minimum localization error is obtained when $DC=60$.

Proof: The triangle topology is showed as Fig. 1.. Assume that d_{13} is fixed and $\angle 123=DC$. Then the denominator of equations (4) can be written as follows: $d_{12}d_{23}\sin(DC)=\dfrac{d_{13}\sin(\angle 132)}{\sin(DC)}\dfrac{d_{13}\sin(\angle 213)}{\sin(DC)}\sin(DC)$ That is:

$$d_{12}d_{23}\sin(DC)=\frac{{d_{13}}^2}{ctg\angle 132+ctg\angle 213} \quad (5)$$

To minimize equations (4), that is to maximize equation (5). The expression ($ctg\angle 132+ctg\angle 213$) achieves a minimum value when $\angle 132=\angle 213$. On the other hand, $\angle 123=DC$. So

$\angle 123=\angle 132=\angle 213=DC=60$. **End.**

3. REFERENCE NODE SELECTION AND PLACEMENT MECHANISMS

3.1 Node Selection for Static Networks

From the two theorems of the second section we can know that, to reduce the localization error, the located node should select the reference node sets that make it to locate at the centre of a circumcircle of the triangle or the triangle of the selected tern of the reference nodes is close to equilateral triangle. Assume that the nodes in sensor nodes are all static; the algorithm of node selection can be described as follows:

$$\begin{cases} |x_{error}| \approx \left| \dfrac{e_r \left[d_{(real)1,p}\left(y_3 - y_2\right) + d_{(real)2,p}\left(y_1 - y_3\right) + d_{(real)3,p}\left(y_2 - y_1\right) + e_r \right]}{d_{12} d_{23} \sin\left(DC\right)} \right| \\ |y_{error}| \approx \left| \dfrac{e_r \left[d_{(real)1,p}\left(x_3 - x_2\right) + d_{(real)2,p}\left(x_1 - x_3\right) + d_{(real)3,p}\left(x_2 - x_1\right) + e_r \right]}{d_{12} d_{23} \sin\left(DC\right)} \right| \end{cases} \quad (4)$$

i. Calculate the DC and the centre of a circumcircle of the triangle for the tern of the reference nodes in the sensor network, and flood the message into the whole network, the message includes the positions of each reference nod group, DC and the positions of the centre of a circumcircle of the triangle.

ii. The located nodes select the reference node sets of $DC \to 60$ or whose positions close to the centre of a circumcircle of the triangle.

iii. Estimate the node positions with respect to equations (2) using the selected reference node sets.

3.2 Reference Node Movement and Placement

When the network is static, the located nodes can only select the good sets of the reference nodes with respect to the fixed geometry topology of the reference nodes passively. To locate more accurately, we can assume that some reference nodes can move in the sensor network. Based on the two theorems of section 2, the movement of the reference nodes can describe as follows:

i. The reference nodes should move the bound of the network to make the located nodes to be inside the geometry topology of the reference nodes.

ii. The movement of the tern of the reference nodes should make their geometry topology to become equilateral triangle.

iii. Estimate the node positions with respect to equations (2) using the selected reference node sets.

On the other hand, to increase the density of reference nodes, we can also deploy additional reference nodes with respect to the proposed two theorems to reduce the localization error.

4. SIMULATION AND ANALYSIS

This section presents some simulation results in different configurations for parameter DC. Assume that the distance between the located node and reference nodes can be obtained by DV-distance or DV-hop algorithms, R is the communication range of nodes.

Fig. 3. shows the topology graph of both the reference nodes and located nodes.

Fig. 4. analyzes the localization performance at some range errors. Assume that the variance of the range measurement error is σ_r. Fig. 4.(a) and (b) denote the localization results at $\sigma_r = 0.001$ and $\sigma_r = 0.01$ respectively. As a whole, DC is very sensitive to range measurement error. When $DC \to 180$, a little range error can cause great localization error. From the figures we can also see that the localization performance is comparatively stabile when DC is less than 179.5, and when DC is larger than 179.5, the localization performance appears to be deteriorated.

Fig. 5. shows the localization performance when the located node is inside or outside of the triangle topology respectively. Assume that the range error $\sigma_r = 0.001$. From the graph we can see that the localization performance of $DC = 60$ is best when the located node is inside the triangle topology, here the three points make up of an equilateral triangle. But when the located node is outside of the triangle, the location performance of $DC = 60$ is not always best.

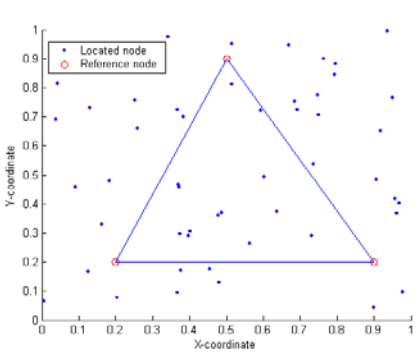


Fig. 3.

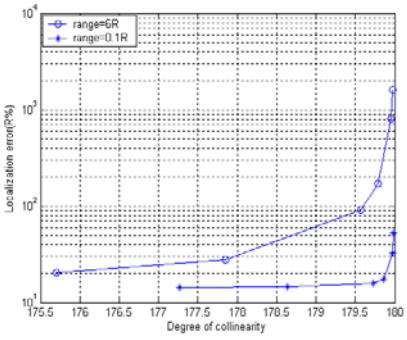


Fig. 4. (a)

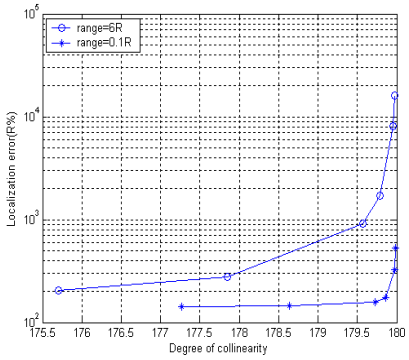


Fig. 4. (b)

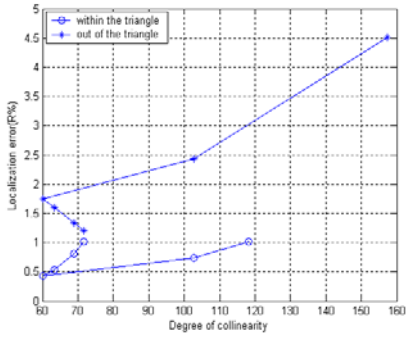


Fig. 5.

5. CONCLUSION

In this paper we study the reference nodes selection and placement mechanisms for localization with respect to the different geometric topology of the terns of the reference nodes. We also present and prove two theorems of reference nodes topology based on the analysis of localization error. The selection and placement mechanisms of reference nodes are proposed with respect to the two theorems. Extensive simulation experiments show that it is helpful to improve the localization performance of sensor nodes using the reference node selection and placement mechanisms, and the mechanisms do not need too much energy and computation complexity.

REFERENCES

[1] I. Stojmenovic, X.Lin, "Gedir: Loop-free location based routing in wireless networks", in Proceedings of IASTED International Conference on Parallel and Distributed Computing and Systems. Boston, MA, November 1999, 1025-1028.

[2] J.C. Navas, T. Imielinski, "Geographic addressing and routing", in MobiCom, Budapest, Hungary, September 1997.

[3] S. Meguerdichian, F. Koushanfar, et al, "Coverage problems in wireless ad-hoc sensor networks", in Proceedings of Infocom, Ankorange, Alaska, 2001.

[4] N. Bulusu, J. Heidemann, et al, "GPS-less low-cost outdoor localization for very small devices", IEEE Personal Communication.2000,7 (5):28–34.

[5] L. Doherty, K. Pister, et al, "Convex position estimation in wireless sensor networks", in IEEE Infocom 2001, Anchorage, AK, 2001.

[6] C. Savarese, K. Langendoen, et al, "Robust positioning algorithms for distributed ad-hoc wireless sensor networks", in USENIX Technical Annual Conference, Monterey, CA, 2002, 317–328.

[7] A. Savvides, W. Garber, et al, "On the error characteristics of multihop node localization in ad-hoc sensor networks", In 2nd Intl. Workshop on Information. Processing in Sensor Networks, 2003.

[8] S. Slijepcevic, S. Megerian, et al, "Location errors in wireless embedded sensor networks: Sources, models, and effects on applications", ACM Sigmobile Mobile Computation and Communication Review, 2002, 6(3):67-78.

Minimizing Delivery Cost in Hybrid Overlay Networks for Streaming Media

Jin Tan, XiaoZhu Chen
Department of Information Engineering, China JiLiang University
Hangzhou, Zhejiang 310018, China
Email: { tanjin,chenxiaozhu}@cjlu.edu.cn

Abstract

There are three key issues on streaming system in Internet: scalability, delivery cost and QoS. The scalability and QoS was improved by means of pushing the streaming objects to the client sides in a CDNs-based streaming system, but the cost of deploying and maintaining was very high. However it was difficult to provide good QoS in a P2P-based streaming system due to the heterogeneous peers and arbitrary network connection of peers. Under the hybrid architecture proposed recently, this paper presents a new content delivery approach with which the streaming media is segmented, cached on proxy in CDNs and stored on peers in P2P dispersedly, hence it takes less resource of peers and be fairer. It is shown that traffic in CDNs and workloads of edge server are reduced effectively, the reliability of streaming delivery in P2P is ensured by using this approach according to performance analysis.

Keywords: replication, streaming, P2P, CDN, delivery cost.

1. Introduction

To overcome the limited capacity of single server, overlay networks, such as Content Delivery Networks (CDNs) and Peer-to-Peer (P2P), are deployed to transmit streaming to a large scale of users [1,2]. The scalability and QoS is improved by means of pushing objects to the client sides in a CDNs-based streaming system, but the cost of deploying and maintaining is very high. However it is difficult to provide good QoS in a P2P-based streaming system due to the heterogeneous peers and arbitrary network connection of peers. In order to enable the two streaming technologies to complement each other, a novel hybrid architecture that integrates both CDN and P2P based streaming media distribution is proposed [3,4].

The replication of all streaming objects between origin server and replicas in CDNs not only consumes the large storage space of replicas but also brings much heavier traffic to the networks, some objects may be accessed rarely; In P2P, total delivery of a stream media is very difficult due to the heterogeneous peers, and the losses of data pocket are brought by come-and-go behavior of the peer. Hence, how to deliver media content will play important roles in minimizing delivery cost in hybrid overlay networks.

2. Related Works

Conventional replication methods for Web Objects in CDNs are not appropriate for streaming media due to the inherent properties of streaming, although some CDN companies advocate their streaming caching support, their technical details are not yet clarified nor verified [5]. Literature [1] discuss the design choices made during the evolution of Akamai's CDN for streaming media in order to ensure the network's scalability, quality of delivered content, and reliability while keeping costs low; Literature [5] discussed how to optimally replicate scalable coding streaming contents on to CDN servers.

In aspect of the P2P-based streaming systems, a multicast system for streaming delivery, called Zigzag, is proposed in Literature [2]; Literature [6] developed compatible system models for P2P streaming systems built upon replication and erasure-correction coding data redundancy schemes [6]. Both CDN-based and P2P-based architectures have their advantages and disadvantages, and each alone does not provide a cost-effective and scalable solution to streaming media distribution [3], a novel hybrid architecture that integrates CDN and P2P was proposed in Literature [3,4]. When a streaming object was published, it was replicated to all edge servers first; then the edge server supplied service acting as a "seed" to a P2P located in its service domain. The edge server could divide the streaming load between itself and the supplying peers, streaming sessions were supplied by supplying peers when the P2P streaming capacity is enough finally.

3. System model

Our strategy is based on the streaming delivery architecture shown in Figure 1. The model is similar to the hybrid structure proposed in [3] and [4] with some modifications. The main entities of the system are:

◆ Streaming Content: The media object a client requests. When an object is requested first, it is replicated from origin server to edge server totally, segmented with fixed length L (such as 10 seconds) and stored on edge server. We assume that all streaming Medias are Constant Bit Rate (CBR) streaming like [4].

◆ Server(called edge server in CDN): The roles of server are an actual streaming server and to maintain two tables which support our strategy.

◆ SP(Service Peer)：The SP is a requesting peer first, after receiving a streaming segment, it supply service to other peers with a limited contribution commitment. Upon joining the system, each SP announces its number of cached segments w ($w > 0$, this is the session capacity of supplying service also) according to its local resources.

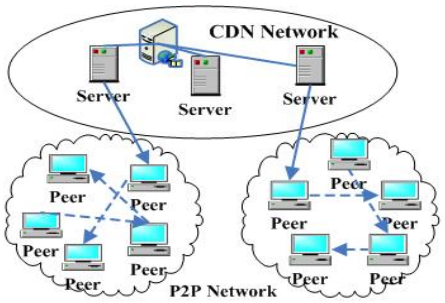


Fig. 1 The hybrid architecture of streaming delivery

◆ CP(Consumer Peer)：The peer that request streaming service only.

4. The Delivery Strategy

An edge server stores streaming segments and creates a "seed" to supply domain of its service.In our strategy, the system operates as follows：

(1) All peers in a service domain must registered to its edge server before they request a streaming object, meanwhile the server gets the information of this peer about ID, IP, W etc.

(2) All peers in a service domain request streaming objects from origin server through its edge server which act as a proxy.

(3) According to the distribution of streaming segments in P2P, the edge server supplies service to a requesting peer with one method selected from following:

- ◆ Service supplied by an edge server only
- ◆ Service supplied by an edge server combined with some SPs
- ◆ Service supplied by some SPs

(4) While a SP request a streaming segment, it caches the segment on its local storage in order to supply service for other peers.

(5) The segmentation of a object is transparent, the handoffs of segment are completed by media player in a peer with sending message ***Request (Segment ID)***.

An edge server creates a "seed" to supply domain of its service, achieves our strategy through maintaining Table 1 and Table 2. Table 1 records the information about ID, IP, Service Max, Cached Counts, Online and Serving Counts of the SP, and field Peer, ID is indexed; Table 2 records the information about a segment ID that has been cached by a SP (ID), and Segment_ID + Peer_ID is indexed. Table 3 is maintained by each SP, which records the Segment ID, File Name (Segment Content) and Servicing Counts (supplying services session counts to other peers).

When a streaming object ***I*** is requested first, the origin server replicates it to requested edge server at which it is segmented with fixed length L such as ***M (I, 1), M (I, 2), …, M (I, n)*** and stored. Meanwhile the edge server transmits first segment ***M(I,1)*** （it is ID）to the actual requested peer (SP or CP) with ***Send(Segment ID, Segment)*** message. If the requested peer is a SP, such as Peer, ID=1, and able to cache the viewing segment according to the replacement algorithm, it will send a ***Cached (Segment ID)*** message to the edge server. When the servers receive this message, it operates in Table 1 and Table 2 as follows:

Table 1 Information about registered SP

Peer ID	IP	Service Max	Cached Counts	Online	Serving Counts
1	IP1	5	4	Y	3
2	IP2	2	2	Y	2
3	IP3	3	1	N	0
…	…	…	…	…	…
789	IP789	1	0	Y	0
…	…	…	…	…	…

Table 2 Relation of Cached Media Segment and Peer ID

Segment ID	Peer ID
M(i,1)	1
M(i,1)	2
M(i,1)	3
M(i,2)	1
M(i,2)	2
M(i,3)	1

M(i,4)	1
…	…

(1) Locates requested SP ID In Table 1, the Cached_Counts is added 1if field Cached_Counts is less than Service_Max;

(2) The tuples of segment ID and peer ID, such as ***(M(I,1),1),*** is inserted into Table 2，this shows that there is one segment ***M(I,1)*** supplied from a SP in P2P, and the session number is W。

Table 3 Data Structure at Peer Side (Peer ID=2)

Segment ID	File Name	Serving Counts
M(i,1)	MI-1	1
M(i,2)	MI-2	1

When a segment cached in some SPs is requested again, such as ***M (I, 1)***, the edge server searches the Peer_ID which has cached the ***M (I, 1)*** in Table 2, if so (in Table 2, three peers have cached this segment), searches first record in Table 1 with condition of "list of Peer_ID which has cached ***M (I, 1)***" and "Service Counts < Service-Max" and "Online='Y'", for the searched result, the Serving Counts is added 1 and IP address is assigned to URL, message ***Redirect (URL, M (I, 1))*** redirect this request to the SP which is able to supply service. If no result searched in Table 1 or there is no peer that caches the segment ***M (I, 1)***, the edge server supplies service to requested peer directly acting as a "seed".

When a SP receive the message ***Redirect (URL, Segmetn ID)***, it supplies streaming session to requested peer acting as a server, sends ***Begin Service (Segment ID)*** message to the edge server, and the Serving Counts is added 1 in Table 3. After the session is end, the ***EndService(Segment ID)*** message is sent and the Serving Counts is subtracted 1 . At SP side, many sessions may be supplied with one streaming segment, but the sum of all Serving_Counts must be less or equal W . If cached segment need to be replaced, the item of Serving Counts=0 is released, and a message ***Uncached (Segment ID)*** is sent to. If a SP try to leave the P2P, it sends message ***Quit ()*** to the edge server first, waits for finishing session it supply (if need, the average waiting time is L/2), and then quits; When the SP enter P2P again, it sends message ***Online ()*** to the edge server.

According to the received messages, the edge server operates as follows:

BeginService(Segment ID)：The corresponding Service Counts of the SP sender is added 1 in Table 1.

EndService(Segment ID)：The corresponding Service Counts of the SP sender is subtracted 1 in Table 1.

Uncached(Segment ID)：The corresponding Cached Counts of the SP sender is subtracted 1 in Table 1, and the record that the SP caches the ***Segment ID*** is deleted in Table

2.

Quit(): The corresponding Service Counts of the SP sender is set zero, and the Online is set 'N' in Table 1.

Online(): The corresponding field Online of the SP sender is set 'Y' in Table 1.

Here is the pseudo-code of the algorithms at the server:

```
Do While (True)
  {
  Receive Message from Peer;
  Switch (message)
    {
    Case BeginService(Segment ID)
       Set Serving_Counts=Serving_Counts+1 in Table 1;
    Case EndService(Segment ID)
       Set Serving_Counts=Serving_Counts-1 in Table 1;
    Case Online ()
    Set Online='Y' in Table 1;
  Case Quit ()
    Set Serving_Counts=0 and Online='N' in Table 1;
  Case Cached (Segment ID)
    If Cached_Counts<Service_Max in Table
      Cached_Counts=Cached_Count+1;
      Insert (SegmentID, Peer_ID) in Table 2:
    Endif
  Case Uncached(Segment ID)
    Set Cached_Counts=Cached_Counts-1 in Table 1;
    Delete Record (SegmentID, Peer_ID) in Table 2;
  Case Request (Segment ID)
    Count Num of Record which Segmetn_ID=SegmentID
          in Table 1;
    If Num=0
      Send Send(Segment ID, Segment) to the Peer;
    Else
      Search Peer ID which cached Segment ID and
        Online= "Y" and Serving_Counts<Service_Max
        in Table 1 and Table2;
      Set URL=IP in Table 1;
      Redirect (URL Segment ID);
    Endif
  }
}
```

5. Performance analyses

The main purposes of our delivery strategy based on hybrid overlay networks for streaming are reducing the traffic in CDNs, decreasing the workloads of edge server and ensuring reliability of streaming transmission in P2P network.

When a streaming object is published, it is not replicated to the all edge servers totally like policy proposed in literature [4], but it is replicated to an edge server when it is requested first. If the probability of the media has been accessed is equal, such as 70%, the traffic in CDNs will be reduced 30%.

In our strategy, because a streaming media is divided into many segments, many peers become service ones with little storage and supplying short sessions; If a SP has not supply service, it can come and go, else it needs to wait a short time (maximum is L) before leaving P2P, the request never be rejected because a edge server act as a "seed".

The CDN is charged based on traffic, reducing workloads of an edge server is minimizing delivery cost. To simplify the analysis, a streaming media is considered.

Assumes that there are P SPs in a P2P located in the service domain of an edge server, the average capacity (average of w) is A and length of the media segment is L, the total storage capacity contributed by all SPs is denoted by C:

$$C = P \times A \times L \qquad (1)$$

Assumes that the request rate to the media is λ, the popularity of each segment is equal, if any SP can supply service for a segment, the cache is hit (only when the edge server supply service with a segment, the cache is missed), and the hit ratio is H, the saved bandwidths of the edge server is denoted B in time t:

$$B = P \times A \times L \times (1 - e^{-\lambda t}) \times H \qquad (2)$$

From formula (2), when P and A are fixed, the H is increased because the media is divided into many segments, and the bandwidths of the edge server are saved effectively.

6. Conclusions

The CDN-based streaming system was used in commerce[7], it is performance-effective but is also very expensive; The advantage of P2P-based system is that the total capacity (bandwidth, storage) grows quickly when the content it manages becomes more popular, but it is difficult to provide good QoS due to the heterogeneous peers and arbitrary network connection of peers. The hybrid architecture that integrates both CDN and P2P complements each other, is a better direction of development and a streaming service platform, the strategy proposed in this paper takes less resource of peers and be fairer, reduces the traffic in CDNs and workloads of edge server effectively, ensures the reliability of streaming delivery in P2P, and have good practicability.

7. Reference

[1] B. Mancuso, D. Shaw, et al, "A transport layer for live streaming in a content delivery network", In Proc. of the IEEE, 2004, 92(9): 1408- 1419.

[2] D. A. Tran, K.A. Hua, et al, "A Peer-to-Peer Architecture for Media Streaming", IEEE Journal on Selected Areas in Communications, 2004, 22(1): 121-133.

[3] D. Xu, S. Kulkarni, et al, "Analysis of a CDN-P2P Hybrid Architecture for Cost-Effective Streaming Distribution", ACM/Springer Multimedia Systems Journal, 2006, 11(4): 383-399.

[4] Y. C. Tu, J. Z. Sun, et al, "An Analytical Study of Peer-to-Peer Media Streaming Systems", ACM Transactions on Multimedia Computing, Communications, and Application, 2005, 1(4):354-376.

[5] S. Zhou, J. Katto, et al, "Optimal Replication Algorithm for Scalable Streaming Media in Content Delivery Networks", IEICE Transactions on Communications, 2004,

E87-D (12): 2723-2732.
[6] W. F. Poon, Y. B. Lee, et al, “Comparison of Data Replication Strategies for Peer-to-Peer Video Streaming”, Proc. of the 5th IEEE ICICS 2005, Bankkok, Thailand, 2005.
[7] G. Pallis, A. Vakali, “Insight and Perspectives for. Content Delivery Networks”, Communications of the ACM, 2006, 49(1): 101-106.

Research and Implementation of LFB Loading and Configuration Mechanism in ForCES Router

Hongjun He, Weiming Wang, Chuanhuang Li
College of Information and Electronic Engineering，Zhejiang Gongshang University
Hangzhou, Zhejiang 310035, P.R.China
Email: hjhe@pop.zjgsu.edu.cn, wmwang@mail.zjgsu.edu.cn

ABSTRACT

Forwarding and Control Element Separation (ForCES) is a protocol for open programmable QoS IP Router, which can achieve higher configurability and flexibility than traditional IP Router. A mechanism is presented for implementing Logical Function Block (LFB) loading and configuration in ForCES Router. After a brief introduction to the architecture of the ForCES Router, the mechanism is described in details. Then, it has been applied to the ForCES router. The experiments have shown that the LFB loading and configuration have been successfully realized in ForCES Router.

Keywords: ForCES LFB, Open Programmable, router.

1. INTRODUCTION

Emergence of Next Generation Network (NGN) greatly improves the flexibility of network and gives support to the promotion of new service and new technology. As an effective scheme of NGN, open programmable network can control and program network nodes through standard interface. The whole network resources can be dynamically configured through the distributed network nodes, which can be controlled dynamically. This scheme makes end-to-end QoS service possible and research on this filed has been a focus of attention.

ForCES, as a workgroup of IETF in routing area specializes the architecture and protocol of open programmable IP router and it is the top-rank organization of research on open programmable network at present. In ForCES router, CE can request FE to dynamically load various LFBs according to requirement of service from network user and then flexibly configure attributes of LFBs.

2. ARCHITECTURE AND MODEL OF FORCES ROUTER

The architecture and module of ForCES router are mostly defined in ForCES Requirement Protocol [1] and ForCES Framework [2]. According to fundamental idea, a ForCES router is a Network Element (NE) and it consists of one primary CE, some redundant CEs and multiple FEs up to hundreds of. The interface between CE and FE is defined by the ForCES protocol [3], which is called Fp reference point. CE primarily handles ForCES messages and is responsible for the management of LFBs in CEs. FE primarily handles and forwards data packets by line-rate. And FE forwards a majority of data packets dependently and some packets involved with routing and protocol messages are passed to CE and process further.

The Architecture of FE is defined by FE Model [4], in which the resources of FE are presented to various kinds of LFBs, each of which has a single function, e.g., classifier, scheduler, IPv4 or IPv6 forwarder. Those LFBs are connected with datapath, form the topology of LFBs and realize dynamic configuration of resources for providing various IP services. The topology and attributes of LFBs are controlled and managed by corresponding messages, which are defined in ForCES protocol.

3. LOADING AND CONFIGURATION MECHANISM OF LFB

The requirements of LFB concerned in ForCES router are defined in ForCES requirement protocol, e.g., (1) FE must know what service can be provided by every LFB; (2) supporting LFB's attributes dynamically adding, modifying and deleting; (3) supporting query of LFB topology; (4) supporting query of LFB attributes; (5) supporting LFB dynamically joining and leaving system, etc. In this paper, we design and implement LFB loading and configuration mechanism on the basis of those LFB requirements above.

3.1 Scheme of LFB Loading and Configuration

The analysis of LFB requirements indicates that LFB configuration from CE includes LFB attributes and LFB topology. LFB attributes are referred to LFB's visible and operable parameters to CE, e.g., configurable flag bit for selecting LFB operation mode, the number of input and output ports and lists CE can read/write through ForCES. The configurable attributes make CE flexibly specify the action of LFB. CE can confirm the state of LFB by querying the setting of LFB attributes.

*Project sponsored by National 863 Project (2005AA121310) , NSF China (60573116), and Zhejiang Key Sci&Tech Project (2005C21013)

Compared with the configuration of LFB attributes, the configuration of LFB topology is rather complex which is graphical expression. The configuration of LFB topology needs dynamically alter datapath including the alteration of LFB in datapath and setting and deleting connection topology of LFBs. CE provides special data service, e.g., DiffServ, VPN, to network user by building datapath in FE. The purpose of re-configuration to datapath makes CE dynamically alter network service according to users' demands when the system is running.

LFB loading and configuration module is shown in Fig.1, which needs a mechanism to describe the configurable ability of LFB topology in FE and report to CE for management. After knowing the ability of FE, CE configures both LFB attributes and LFB connection topology according to users' demands, sends messages to FE via ForCES interface which realizes LFB loading and configuration via packet and message data paths between modules.

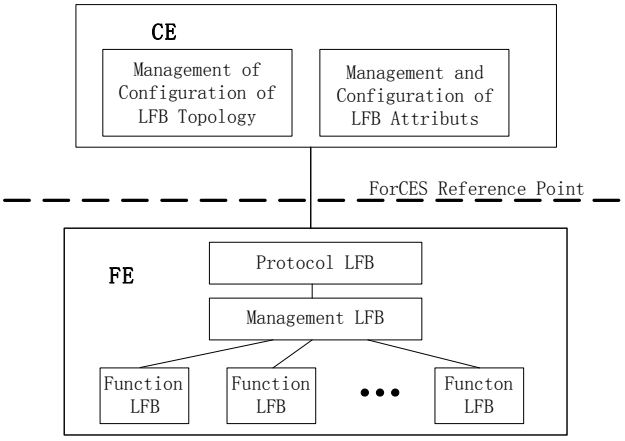


Fig. 1. LFB Loading and Configuration Module in ForCES

3.2 Implementation of LFB Loading and Configuration Mechanism

The key of realizing LFB loading and configuration includes the management of LFB attributes and topology and mechanism of packet and message data paths between modules.

1) CE's Role in the Management of LFB attributes and topology

The management module of LFB attributes is responsible for the process of modifying and updating FE&LFB attributes. When FE connects CE and reports its attributes, CE will obtain LFB attributes/abilities, display them via tree structure and save them as XML file. In order to conveniently manage LFB, we divide LFB attributes into two parts: one part shows some common attributes such as LFB type, ID and connection topology; the other part shows special attributes of various LFBs. The data structure of LFB attributes is shown below:

```
struct LFB_Instance
{
    int st ate; /*describe state of LFB*/
    int type; /*describe type of LFB, e.g., Receiver
            LFB and Forwarder LFB*/
    int instanceID; /*describe ID of LFB in FE*/
    int input_num; /*describe number of input ports*/
    int output_num; /*describe number of output ports*/
   LFB_Connect input [MAXINNUM];     /*describe
            input port connecting LFB*/
   LFB_Connect output [MAXOUTNUM]; /*describe
            output port connecting LFB*/
   Void *LFB_Attri; /*describe special attributes of
            LFBs*/
}
```

LFB topology module resolves dynamical display of LFB topology graphics, which makes CE operator distinctly know LFB's topology location and connection of LFBs. With the help of that information, CE operator has a rough view to policy of packet processing in router and enhances operability. The data structure concerned with LFB topology is below:

```
struct LFB
{
    int LFBID;
    int Input [MAXINNUM];
    int Output [MAXOUTNUM];
};   /*LFB structure, including ID, input and output*/
struct LFB LFBArray [MAXLFBNUM];
    /*structure of array including all LFBs*/
Int Datapath [MAXPATHNUM][MAXLFBNUM];   /*array
    including all datapaths*/
int LFBNum;     /*total number of LFBs*/
int PathNum;    /*total number of datapaths*/
struct LFB
*DatapathMatrix[MAXPATHNUM][MAXLFBNUM];
    /*LFB matrix*/
```

After receiving the response message of querying topology, CE fill LFB array, find all the datapaths, fill Datapath array of two dimensions via recursion transfer algorithm and realize the connection topology of LFBs via DatapathMatrix.

In order to support the management of LFB attributes and topology from CE, we need to design a management library of LFB attributes which is a set of function to configure and manage LFB attributes, responsible to obtain configurable attributes of various LFBs and design and realize the interface function of configuration and management concerned with LFBs' attributes. On the basis of the management library of LFB attributes, CE realizes the configuration and management of LFB attributes.

2) Mechanism of packet and message data paths in FE

In order to realize LFB loading and configuration, we use the mechanism of message correspondence between modules, which can suit the expansion, and management of

modules. Each LFB may comprise several modules and it is possible that one module may comprise some operation to several LFBs. Each module can have multiple inputs, each of which is associated with a different packet/message handler. Each input is associated with a hardware or software queue, and has a globally unique ID. The execution engine registers a packet/message handle for each input ID on behalf of a module. The set of IDs for packet handles and message handles have the same arrange. The IDs are allocated at compile time. Fig.2 shows how modules are interconnected to form message and packet data paths. Now asynchronous packet/message processing between module A and B is described as below:

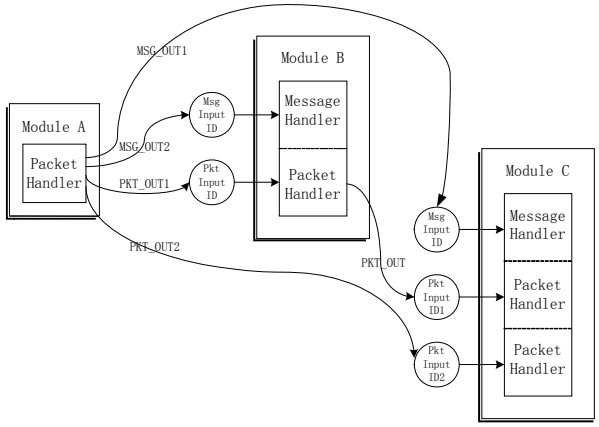


Fig. 2. Packet and Message Data Paths between modules

Module A send one message whose destination are determined by the globally unique ID, passing message type, such as obtaining certain state value or setting certain argument. When module A sends message to module B, it must provide argument value passed to module B, pointer to store return value and a callback function which is called by module B after ocessing message. After sending message to module B, the thread associated with module A enters sleep state and waits for module B processing message which can be realized via semaphore, and then the callback function must free the semaphore. Module B queries its message queue and if there is message needs to be handled, then call the registered message handle function, judge message type and execute corresponding operation. At last module B calls the callback function provided by module A, and the entire processing is completed.

When a single thread has multiple modules or one module is associated with multiple inputs including message/packet queue, we need to use a scheduling policy to manage various queues such as Round robin and Weighted Round-robin. Considering CE needs dynamic configuration, we can use a peculiar module, which is named management module to manage loading of the entire function module. Management module dynamically execute LFB's configuration and function modules realize the core function of LFB.

4. IMPLEMENTATION AND TEST

We use VC++ to implement CE and describe the connection and accessory relationship among NE, CE, FE and LFB via CtreeCtrl controller. Thus the operation of FE and LFB from CE can be changed into the operation of tree node, and we save the information about NE as .XML file for realizing the dynamical management of LFB. In our implementation CE is composed of three parts: NE tree information display window, topology display window and associated information output window.

Considering the flexibility and performance of FE, we choose Intel IXP2XXX network processor [5] to develop FE. The Xscale layer is the upper control unit of network processor and we use VxWorks operating system to receive the ForCES message from CE, which provides the path of control, query and event information to LFB's loading.

Now we have implemented multiple LFB's loading in ForCES router, and DSCP Classifier LFB is one of part of QoS module which primary function is sorting packets from inputs according DSCP field in IP header and processing packets via class rule. The sorting result is written to the description attribute of packet for the next LFB in order to provide DiffServ. In Classifier LFB every DSCP value set a stream of data and we can define some handle rules to provide various servers for every DSCP stream.

In the implementation of ForCES router, CE set DSCP Classifier LFB state --- LFBstate from OFF to ON via LFB's attribute tree for loading the LFB. The setting above is encapsulated into ForCES message via CE and sent to FE, here multiple LFBs are used such as protocol LFB, management LFB, DSCP LFB which are associated with a data/message handling entity/module in FE. After decapsulating the message via protocol LFB, FE analyzes its loading DSCP LFB and then sends to DSCP via management LFB, thus we realize the topology rebuilding of associated function LFBs.

After starting DSCP Classifier LFB, LFB topology is changed and makes ForCES router supporting QoS function. We can modify DSCP rule via CE tree node and forward packets in different policy. Testing in ForCES router shows that the loaded LFB works well and then prove the validity of the LFB loading and configuration module.

5. CONCLUSION

In this paper, we introduce and implement an loading and configuration mechanism, which is very important in ForCES router. First, it makes CE dynamically configure network services on system-run -time according to users' demands. Second, it promotes the flexibility of network equipment. Third, it facilitates the network reconfiguration and rebuilding. Last but not the least, it helps to form intelligently dynamical network and fasten the deployment of network upgrade and new service. The test proves that the mechanism is effective and practicable in ForCES router.

6. REFERENCE

[1]H. Khosravi, T. Anderson, "Requirements for Separation of IP Control and Forwarding", RFC 3654, 2003.

[2]L. Yang, R. Dantu, et al, “Forwarding and Control Element Separation (ForCES) Framework”, RFC3746, 2004.

[3] A. Doria, R. Haas, et al, “ForCES Protocol Specification”, draft-ietf, March, 2006.

[4]L. Yang, J. Halpern, et al, “ForCES Forwarding Element Model”, draft-ietf, September 2006.

[5] Intel, “Intel Internet Exchange Architecture Protability Framework Developer’s Manual”, 2004.

An Optimization Model Algorithm Based on MPLS Network

ZhanTing Yuan, QiuYu Zhang, ZhuPing Hua,JianMing Yu
School of Computer and Communication Lanzhou University of Technology
Lanzhou, Gansu Province, China 730050
Email:hzpalqy@163.com

ABSTRACT

Multiprotocol Label Switching has become the important technique in NGN'development, contributing for doing an efficient Traffic Engineering in order to provide the QoS and to optimize resource allocation. For this purpose, it is possible to apply Mixed Integer Mathematical Programming techniques to model routing problem minimizing implantation, operation and maintenance network costs. We present an optimization model applied to the MPLS network planning, which assign paths based on capacity and network architecture constraints, considering LSP survivability. The model is consolidated in two case studies, applying them to plan a MPLS network. The main characteristics of the MPLS concerning to the network planning approach are presented.

Keywords: MPLS, Label Switched Paths, Mixed Integer Mathematical Programming, Label Switching Router, Label Edge Router.

1. INTRODUCTION

MPLS technology integrated use respective strengths of their switching technology of core network and IP routing technology of network edge, which introduce the second layer of exchange rate into the third layer, and the exchange mode based on labels allow the router solely based on simple labels when it make the transmitted decision, instead of IP addresses based on the objective search for complex routing.MPLS become more important means to solve the problem of Traffic Engineering and VPN, and increasingly became important standard used in the core telecommunications network[3].

For planning an MPLS network in an efficient manner, it is necessary to consider changes in the architecture (installation of new links and routers, for example) and the use of Traffic Engineering (TE) to refine dimensioning. Traffic Engineering is a process, which aims to control the traffic in a network, in order to optimize the use of existent resources, improving network performance [1, 2]. For an effective Traffic Engineering, IETF introduced MPLS-TE [3] with Constraint-Based Routing [1, 4]. Multi-Protocol Label Switching (MPLS) is an advanced forwarding scheme for IP networks supporting Quality of Service (QoS). MPLS adds labels to the packet header to switch packets in a similar way to ATM.

2. MPLS NETWORK STRUCTURES

An MPLS network consists of edge label switch routers (edge LSRs) around a core of label switch routers (LSRs). Customer sites are connected to the carrier's MPLS net-work, or equivalently a large organization's net-work backbone. The customer premises equipment connect-ed to an MPLS network typically runs ordinary IP forwarding rather than MPLS, and is typically a router or a LAN switch. Since the customer equipment typically does not run MPLS, the edge LSRs are part of the carrier's network and under the carrier's administration.

A carrier's MPLS network will often be con-nected to one or more other IP networks as part of the Internet. An IP connection to another carrier might be an MPLS link, although use of MPLS on intercarrier links is usually not required. As with any interconnection of carrier networks in the Internet, the Border Gateway Protocol [6, 7] would typically operate over links to other carriers, to exchange routing information with them. The neighboring IP networks may use MPLS internally, but not necessarily.

The links between customer equipment, edge LSRs, and/or LSRs may be of virtually any type. Traditional non-MPLS switches may be used to carry connections to MPLS equipment, typically using some sort of permanent virtual circuits or optical light paths. Several LSPs may be established between a pair of LERs. These LSPs may choose different paths in order to uniformly distribute the traffic or for Traffic Engineering purposes. It could be a good strategy but the number of LSPs increases proportionally to the number of LER pairs. The policy of color distribution is an administrative resource used to forbid the utilization of certain paths by an LSP. These constraints decrease the number of possible paths between a given origin-destination pair. Network characteristics and its state change over time. For example, new resources become available, failed resources become reactivated, and allocated resources become de-allocated. This may lead to the availability of more efficient paths. Therefore, it may be desirable to dynamically change the paths of certain traffic trunks in response to changes in network state. This process is called re-optimization. In other scenarios, re-optimization might be very undesirable. In the following section an offline optimization model to improve the online CBR algorithm is proposed. This model considers interfaces and routing costs, equipment capacities and affinity [6,7].

3. LINK DIMENSIONING MODEL FOR MPLS NETWORK

We describe the mathematical model for link dimensioning, which determines the capacities of links and the LSPs. We also considere link survivability [5, 8].

The network topology is modeled by a directed graph *G (N, A)*, where N is the set of nodes, which represent routers and A is the set of arcs, which represent links. Each directed arc is represented by $(i, j) \in A$ and the traffic flows from *i* to *j*. LSPs arerepresented by the ingress and egress LERs denoted, respectively, by k and l. Or considering link survivability the demand d^{kl} is divided into two parts $d_1^{kl} = \alpha d^{kl}$ and $d_2^{kl} = \beta d^{kl}$, where $0.5 \le \alpha \le 1$

and $0.5 \le \beta \le 1$. There are two sets of possible LSPs: K and K^s, which will flow demands d_1^{kl} and d_2^{kl}, respectively. C_1^{kl} and C_2^{kl} are sets of links, which can not be used by LSPs $(k, l) \in K$ and $(k, l) \in K^s$, respectively. Other parameters that specify network characteristics are:

T - Set of transmission facilities type t.

b^t - Bandwidth of interface type $t \in T$.

γ^t -Maximum number of interfaces type $t \in T$ that can be allocated in a slot.

n_i -Number of available slots per node (router) i.

c_{ij} -Operational cost of link $(i, j) \in A$..

ω^t -Cost of interface type $t \in T$.

Hkl -Maximum number of intermediate hops for LSP $(k, l) \in K$.

δ^{kl} -Maximum number of intermediate hops for LSP $(k, l) \in K^s$.

μ_{ij} -Utilization of link $(i, j) \in A$.

The variables of the problem are:

λ_{ij}^t -Integer variable that indicates the number of interface of type $t \in T$ to be installed on link $(i, j) \in A$.

φ_{ij}^{kl} -Flow through link $(i, j) \in A$ from LSP $(k, l) \in K$.

ρ_{ij}^{kl} -Flow through link $(i, j) \in A$ from LSP $(k, l) \in K^s$.

f_{ij}^{kl} -Binary variable, which indicates the existence of flow on link $(i, j) \in A$ from LSP $(k, l) \in K$.

h_{ij}^{kl} -Binary variable, which indicares the existence of flow on link $(i, j) \in A$ from LSP $(k, l) \in K^s$.

The problem of determining the minimum cost of an MPLS network is formulated as a mixed integer mathematical programming problem [5, 8]:

$$\textbf{Min } \frac{1}{2}\sum_{(i,j)\in A}\sum_{t\in T}\omega^t\lambda_{ij}^t+\frac{1}{2}\sum_{(i,j)\in A}\sum_{(k,l)\in K\cup K^s} c_{ij}(\varphi_{ij}^{kl}+\rho_{ij}^{kl}) \textbf{ (1)}$$

$$\sum_{(i,j)\in A} f_{ij}^{ik} = 1 \quad i = ingress\ LER\ e\ (i,k) \in K \textbf{ (2)}$$

$$\sum_{(i,j)\in A} h_{ij}^{ik} = 1 \quad i = ingress\ LER\ e\ (i,k) \in K^s \textbf{ (3)}$$

$$\sum_{(j,i)\in A} f_{ji}^{ki} = 1 \quad i = egress\ LER\ e\ (i,k) \in K \textbf{ (4)}$$

$$\sum_{(j,i)\in A} h_{ji}^{ki} = 1 \quad i = egress\ LER\ e\ (i,k) \in K^s \textbf{ (5)}$$

$$\sum_{(i,j)\in A} f_{ij}^{kl} = \sum_{(i,j)\in A} f_{ji}^{kl} \quad i = \text{backbone LSR e } (k,l) \in K \textbf{ (6)}$$

$$\sum_{(i,j)\in A} h_{ij}^{kl} = \sum_{(i,j)\in A} h_{ji}^{kl} \quad i = \text{backbone LSR e } (k,l) \in K^s \textbf{ (7)}$$

$$\sum_{(k,l)\in K\cup K^e} (\varphi_{ij}^{kl}+\rho_{ij}^{kl}) \le \sum_{t\in T}\mu_{ij}b^t\lambda_{ij}^t \ , \quad \forall (i,j) \in A, \textbf{ (8)}$$

$$\lambda_{ij}^t = \lambda_{ji}^t \qquad \forall (i,j) \in A, \textbf{ (9)}$$

$$\sum_{(i,j)\in A}\sum_{t\in T}\lambda_{ij}^t \le \sum_{t\in T}\gamma^t n_i \ , \qquad \forall i \in N, \textbf{ (10)}$$

$$\sum_{(i,j)\in A} f_{ij}^{kl} \le \eta^k + 1 \ , \qquad \forall (k,l) \in K, \textbf{ (11)}$$

$$\sum_{(i,j)\in A} h_{ij}^{kl} \le \delta^k + 1 \ , \qquad \forall (k,l) \in K^s, \textbf{ (12)}$$

$$\varphi_{ij}^{kl} = \alpha d^{kl} f_{ij}^{kl} \qquad \forall (i,j) \in A \ and \ \forall (k,l) \in K \textbf{ (13)}$$

$$\lambda_{ij}^t \ge 0 \text{ and integer } \forall (i,j) \in A \text{ and } t \in T. \textbf{ (14)}$$

$$f_{ij}^{kl} + h_{ij}^{kl} \le 1, \qquad \forall (k,l) \in K \cup K^s \textbf{ (15)}$$

$$\varphi_{ij}^{kl} = 0 \text{ if } (i,j) \in C_1^{kl} \text{ and } \rho_{ij}^{kl} = 0 \text{ if } (i,j) \in C_2^{kl} \textbf{ (16)}$$

$$\varphi_{ij}^{kl} \ge 0 \ \forall (i,j) \in A \ \ and \ \ \forall (k,l) \in K, \textbf{ (17)}$$

$$\rho_{ij}^{kl} \ge 0 \ \forall (i,j) \in A \ \ and \ \ \forall (k,l) \in K^s, \textbf{ (18)}$$

$$\lambda_{ij}^t \ge 0 \text{ and integer } \forall (i,j) \in A \text{ and } t \in T. \textbf{ (19)}$$

The objective function **(1)** represents the total cost of the network as a function of the number of transmission facilities installed in each arc and the flows through the arcs. Constraints from **(2)** to **(7)** guarantee the balance flow of all nodes of the graph (network). Constraint **(8)** impose that the total bandwidth installed in each arc is large enough to support the bandwidth used by the LSPs that cross the arc and constraint **(9)** consider that the arc is bi-directional. The number of interfaces installed in each router is limited in the constraint **(10)**. It is imposed that each LSP has a limit in the number of hops, as represented by equations **(11)**、**(12).** The link survivability and color are represented, respectively, by constraints **(15)** and **(16)**. Constraints **(17)** to **(19)**are related to variables[8].

4. CASE ANALYSIS

4.1 Network description

In this section, we present two case studies considering one network composed by 6 Points of Presence (POP) and 4 backbone nodes (LSR), as shown in Figure 1.

Node 5 is, at the same time, a POP and a backbone node. In each POP there is a LER, which may be the ingress or egress LER of an LSP. They are connected to backbone through 2 links, with the exception of node 5 that is connected through 4 links. Backbone nodes are fully interconnected

We consider three types of transmission facilities: T34M, T68M and T155M, which are characterized by cost, bandwidth and the number of slots to plug them, as shown in Table 1. We also assume that the operation and maintenance cost per unity of flow (cij*)* is l, independently of types of link or flow. The utilization factor is equal to 1 for all links. The Survivability parameters are α=1.0 and β=1.0.

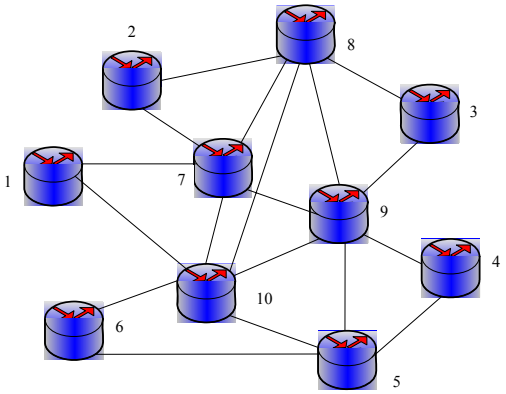


Fig. 1. Topology of the MPLS network

Table 1. Transmission facilities characterization

Interface	Cost(c^t)	Bandwidth in Mbps(b^t)	Maximum number of interfaces per slot(γ^t)
T34M	50	30	2
T68M	100	60	1
T155M	150	80	0

Another constraint imposed is affinity, also known as color, which specifies the class of resources, which are to be included or excluded from the path of the LSP In this example, we impose that just backbone nodes are able to route packets, with the exception of node 5 that is at the same time an access and backbone node .In Table 2. below POPs.

Table 2. Traffic matrix among POPs

POPs	1	2	3	4	5	6
1	-	5	12.5	10	12.5	5
2	5	-	12.5	10	12.5	5
3	12.5	12.5	-	12.5	12.5	5
4	10	10	12.5	-	25	5
5	12.5	12.5	12.5	25	-	2.5
6	5	5	5	5	2.5	-

For the case studies, we consider different maximum numbers of hops (η^{kl} and δ^{kl}) and maximum numbers of slots per router (*ni*).

4.2 Result Research

In Table 3, we observe the total traffic in each link, considering survivability. We obtained two values of the objective function: 2210 for case study 1 and 2280 for case study 2 and the traffic is distributed according to Table 3. The differences are in links (5, 9), (5, 10), (8, 9) and (8, 10).

Table 3. Results: total traffic per link (bi-directional) considering survivability

Link (i,j)(j,i)	Total traffic per link (bi-directional)	
	Case study 1	Case study 2
(1,7)(7,1)	45	45
(1,10)(10,1)	45	45
(2,7)(7,2)	45	45
(2,8)(8,2)	45	45
(3,8)(8,3)	55	55
(3,9)(9,3)	55	55
(4,5)(5,4)	62.5	62.5
(4,9)(9,4)	62.5	62.5
(5,6)(6,5)	22.5	22.5
(5,9)(9,5)	100	82.5
(5,10)(10,5)	50	77.5
(6,10)(10,6)	22.5	22.5
(7,8)(8,7)	0	0
(7,9)(9,7)	80	80
(7,10)(10,7)	0	0
(8,9)(9,8)	27.5	0
(8,10)(10,8)	57.5	75
(9,10)(10,9)	0	0

The difference in objective function values is mainly consequence of the number of interfaces chosen, as we can observe in Table 4. As in study case 2 there are more constraints, the objective function value; higher than the one in case study 1, as expected.

We note from Table 4. , for example, that when we reduce the maximum number of hops (η^{kl} and δ^{kl}) and the maximum number of slats per router (*ni*). (case study 2) there is an increase in the number of interfaces T155M.

Table 4. Results: number and types of interfaces for each link of the network

Link (i,j)(j,i)	Interfaces					
	Case study 1			Case study 2		
	34 M	68 M	155 M	34 M	68 M	155 M
(1,7)(7,1)	0	1	0	0	1	0
(1,10)(10,1)	0	1	0	0	1	0
(2,7)(7,2)	0	1	0	0	1	0
(2,8)(8,2)	0	1	0	0	1	0
(3,8)(8,3)	0	1	0	0	1	0
(3,9)(9,3)	0	1	0	0	1	0
(4,5)(5,4)	1	1	0	1	1	0
(4,9)(9,4)	1	1	0	0	0	1
(5,6)(6,5)	1	0	0	1	0	0
(5,9)(9,5)	0	0	1	0	0	1
(5,10)(10,5)	0	1	0	0	0	1
(6,10)(10,6)	1	0	0	1	0	0
(7,8)(8,7)	0	0	0	0	0	0
(7,9)(9,7)	1	1	0	0	0	1
(7,10)(10,7)	0	0	0	0	0	0
(8,9)(9,8)	1	0	0	0	0	0
(8,10)(10,8)	0	1	0	0	0	1
(9,10)(10,9)	0	0	0	0	0	0
Total	6	11	1	3	7	5

The reason is the constraint in the number of slots per router, which is more restrictive in case study 2, mainly in routers 7 to 10, where $n_i = 1$ for each type of interface. For these reason, for example, in link (7, 9) of case study 2, the interface chosen is a T155M and not T34M+ T68M, which has enough capacity to flow the traffic (80) and is cheaper (50+100=150).

Moreover, η^{kl} and δ^{kl} are higher in case study 1. Compared with case study 2. As a consequence, the algorithm directs the traffic for certain path in order to use as maximum as possible the capacity of links.

When the maximum number of hops is reduced, even if there is available link capacity, sometimes it may not be used because the traffic had to flow through a path longer than the specified limit

Related to LSPs considering survivability, we observe in Fig. 2. An example of the paths adopted for the LSP (2, 4). As we considered in the model disjoint link survivability, the LSPs did not use the same link, even though they use some of the same nodes.

In both case studies,we adopted survivability parameters as α=1.0 and β=1.0. This means that in case of failing a link in one LSP, there is another LSP with exactly the same capacity to flow the traffic of the failed LSP. This choice increases the cost of the network and sometimes it may be infeasible its adoption. More realistic survivability parameter values are, for example, α=1.0 and β=0.5，which guarantees that 50% of the traffic is flown in case of failing a link.In this model, we adopted the same values of α and β，$\forall (i,j) \in A$. This is a simplified approach. In real networks, we can specify different values of α and β for each LSP, for example. Using our notation, this parameter would be denoted by $\alpha^{kl}$and β^{kl}.

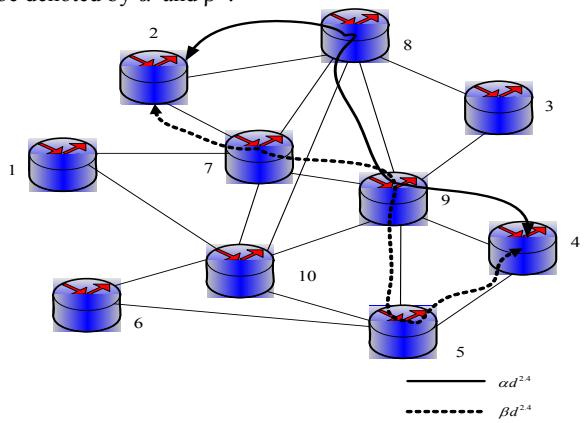


Fig. 2. LSP2-4: paths considering survivability

5. CONCLUSION

In this work we present some important characteristics of an MPLS network aiming at network planning and describeLSP's attributes. These attributes contribute for the Constraint-based Routing-CBR, which can be online or offline.We propose an optimization model to minimize administration routing and interface costs for designing LSP's routes, based in constraints like Affinity link and router capacities.

For future works, the model may be improved incorporating other constraints or modifying the existent ones. For example, the constraint **(8)** would not consider the whole flow $\sum_{(k,l)\in K\cup K^s} \rho_{ij}^{kl}, \quad \forall (i,j) \in A$, at each link but its maximum value. This modification considers that the probability of one link failure is higher than the probability failure in more than one links simultaneously. Moreover, disjoint node survivability may be incorporated as presented in[8]. Another modification in the model that we suggest for future works is to specify different values of α and β for each LSP, the clients could obtain a specific services.

6.REFERENCES

[1] E. Rosen, A. Viswanathan, et al, "Multiprotocol Label Switching Architecture," RFC 3031, Jan. 2001.
[2] G. Armitage, "MPLS: The Magic Behind the Myths," IEEE Commun. Mag., vol. 38(1), Jan. 2000, pp. 124–31.
[3] IETF RFC 3031 Multiprotocol Label Switching Arcititectzrre[S], Jan., 2001.
[4] IETF RFC 3026 Label Distribution Protocol Specification [S], Ian. 2003.
[5] Maria, S. Marcos, An Optimization Model for MPLS Network [J], NETWORKS, 2004 11th (285-290).
[6] Karasan, E, Yetginer, "Robust Path Design Algorithms for Traffc Engineering with Restoration in MPLS Networks", IEEE Trans Communication, 2003, Vol. E86-B(5), May,2003, (1632-1640).
[7] L. Cardoso, J. Patrao, "Combining Node and Link Dimensioning for MPLS". Networks 2002, Munique. Proceedings. Munique, 2002.
[8]Z.Y. Li, L. Wang, et al, "Study on Traffic Engineering Based on Constraint-Based Routing in MPLS Network",Computer Engineering and Applications August 2004(165-167).

Distributed Particle Filter for Maneuvering Target Passive Tracking In Sensor Networks

Feng Xue, Zhong Liu and Xiaorui Zhang
Electronics Engineering College, Naval University of Engineering
Wuhan, Hubei 430033, China
Email: {xfmilk, hunter, zxr_hgxt}@sohu.com

ABSTRACT

To improve the passive tracking performance of maneuvering target and reduce communication cost in wireless sensor networks, we proposed a distributed particle filtering (DPF) algorithm in this paper. The centralized particle filtering (CPF) was presented based on the turn rate model. The dynamic clusters of sensor nodes were organized to apply particle filters distributively, and cluster heads (CHs) were selected as local processing nodes whose local estimates and covariances were transmitted to the sink node for global estimate fusion. The particle information exchanged between CHs was approximated by Gaussian mixture model. Detailed implementation steps of DPF were deduced. Computer simulations were conducted to compare tracking performance and to analyze communication data amount. Simulation results show that the DPF has similar good performance with the CPF in tracking accuracy but lower communication cost than the CPF.

Keywords: Distributed Computing, Target Tracking, Passive Tracking, Maneuvering Target, Sensor Networks.

1. INTRODUCTION

Recent advances in electronic and wireless technologies have greatly improved the processing and communication capacities of wireless sensor networks (WSNs)[1]. One of the most important applications in WSNs is target tracking, which adapt traditional tracking algorithms to deal with unique problems in WSNs. In the military application, stealthy operations require WSNs to obtain the information passively.

It is well known that Particle filters (PF) are very suitable for nonlinear/nonGaussian applications[2]. Hence, high accuracy state estimation can be obtained by PFs applied in the general passive tracking. However, in the context of WSNs, if the PF is used directly in a fusion center to process all observations from sensor nodes, unbalanced communication and heavy computation will cost the limited energy of WSNs quickly, which will lead to the center node failure. Parallel PF is investigated in recent literature[3], and it provides a new way to run PF distributively in WSNs.

In this paper, a distributed PF (DPF) is presented for passively tracking an maneuvering target through a WSN where dynamic clusters are constructed. The particle set is divided into subsets processed by cluster nodes distributively. The goal of the proposed algorithm is to perform on-line, distributed estimation of the current state at multiple nodes, whilst attempting to minimize communication overhead and computation on single node. Computer simulation results have shown that not only the tracking accuracy is improved but also the energy cost is reduced by the DPF in WSNs.

2. CENTRALIZED PASSIVE TRACKING IN WSN

2.1 State space model

To describe the state space evolution of maneuvering target precisely, the tracking model is constructed by the turn rate[4]. Define the state vector as $\mathbf{x}_k = (r_{xk}, v_{xk}, r_{yk}, v_{yk}, \varepsilon_{k+1})$, and considering the two-dimensional passive tracking problem, discrete state and measurement equation are given by

$$\mathbf{x}_k = \mathbf{\Phi}_{k/k-1}\mathbf{x}_{k-1} + \mathbf{\Gamma}\mathbf{w}_{k-1} \tag{1}$$

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{v}_k, \mathbf{u}_k) = \tan^{-1} r_{xk} / r_{yk} + \mathbf{v}_k \tag{2}$$

where the transition matrix is

$$\mathbf{\Phi}(\varepsilon_k) = \begin{bmatrix} 1 & \frac{\sin(\varepsilon_k T)}{\varepsilon_k} & 0 & -\frac{(1-\cos(\varepsilon_k T))}{\varepsilon_k} & 0 \\ 0 & \cos(\varepsilon_k T) & 0 & -\sin(\varepsilon_k T) & 0 \\ 0 & \frac{(1-\cos(\varepsilon_k T))}{\varepsilon_k} & 1 & \frac{\sin(\varepsilon_k T)}{\varepsilon_k} & 0 \\ 0 & \sin(\varepsilon_k T) & 0 & \cos(\varepsilon_k T) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The turn rate ε_k is assumed velocity dependent according to the following model $\varepsilon_k = a_{typ} / \sqrt{v_{xk}^2 + v_{yk}^2}$, where a_{typ} is the typical maneuvering acceleration, which is modeled as a set of three discrete values, having a Markovian switching structure.

2.2 Centralized particle filtering algorithm

In a centralized particle filtering (CPF) architecture, all measurements are sent to a central site for processing. We assume that measurements at the individual nodes are independent conditioned on the states. Hence, combined data likelihoods for all sensors can be factored into the product of data likelihoods at the individual sensor nodes.

$$p(\mathbf{z}_k^{1:M} \mid \mathbf{x}_k) = \prod_{m=1}^{M} p(\mathbf{z}_k^m \mid \mathbf{x}_k) \tag{3}$$

where M denotes the number of sensor nodes. So the posterior density is given by

$$p(\mathbf{x}_k \mid \mathbf{z}_k^{1:M}) = p(\mathbf{x}_k) \prod_{m=1}^{M} p(\mathbf{z}_k^m \mid \mathbf{x}_k) \propto \prod_{m=1}^{M} p(\mathbf{z}_k^m \mid \mathbf{x}_k) \tag{4}$$

Hence choosing $p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$ as the proposal density, the normalized weight of N particles and state estimate can be computed by

$$w_k^i = w_{k-1}^i \prod_{m=1}^{M} p(\mathbf{z}_k^m \mid \mathbf{x}_k) / \sum_{j=1}^{N} [\prod_{m=1}^{M} p(\mathbf{z}_k^m \mid \mathbf{x}_k)] \tag{5}$$

$$\hat{\mathbf{x}}_k = \mathrm{E}[\mathbf{x}_k \mid \mathbf{z}_{1:k}] = \int \mathbf{x}_k p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) d\mathbf{x}_k \approx \sum_{i=1}^{N} w_k^i \mathbf{x}_k^i \quad (6)$$

If effective particles become degenerate, the resampling step usually is taken as follows

$$\left\{x_k^i, w_k^i\right\} \rightarrow \left\{x_k^j, 1/N\right\} \quad (7)$$

This architecture is theoretically optimal since the central site has access to all information. However, significant bandwidth is needed for communication and the central site is a single point of failure.

3. DISTRIBUTED PARTICLE FILTERING

3.1 Dynamic cluster management

Sensors nodes are organized into clusters that are constructed and destroyed according to the target's position. The sensor node closest to the target is selected as the CH. Sensor nodes do not statically belong to a cluster and may support different clusters at different times.

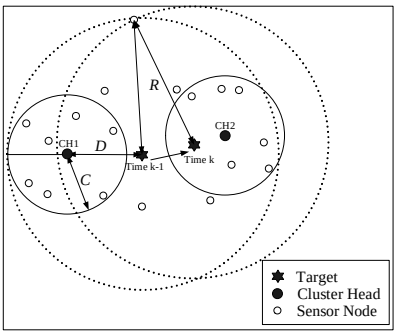


Fig. 1. Dynamic cluster generation and destroy

During the tracking, the first CH detecting the target becomes the processing node and initializes the tracking with a state estimate, and each CH runs the processing algorithm in parallel. The belief state $\mathbf{x}_k$ is passed from the CH to the sink node that completes the fusion of position estimation.

3.2 Distributed particle filtering in the cluster

Considering a cluster with N child nodes and a CH, n_j denotes the particle number on the node j, and $\{\mathbf{x}_k^{i,j}, w_k^{i,j}\}$ is the ith particle at sample time k. PF processing is usually divided into four steps: state sampling, weight updating, state estimate and resampling. In the first two steps, each child node runs PF independently, so these steps can be easily pipelined.

$$\mathbf{x}_k^{i,j} \sim q(\mathbf{x}_k \mid \mathbf{x}_{k-1}^{i,j}, \mathbf{z}_k) = p(\mathbf{x}_k \mid \mathbf{x}_{k-1}^{i,j}) \quad (8)$$

Weight updating is given by

$$w_k^{i,j} \propto w_{k-1}^{i,j} \frac{p(\mathbf{z}_k \mid \mathbf{x}_k^{i,j}) p(\mathbf{x}_k^{i,j} \mid \mathbf{x}_{k-1}^{i,j})}{q(\mathbf{x}_k^{i,j} \mid \mathbf{x}_{k-1}^{i,j}, \mathbf{z}_k)} = w_{k-1}^{i,j} p(\mathbf{z}_k \mid \mathbf{x}_k^{i,j}) \quad (9)$$

Then, each child node computes the aggregated data as

$$S_k^j = \sum_{i=1}^{n_j} w_k^{i,j} \ , \ X_k^j = \sum_{i=1}^{n_j} \mathbf{x}_k^{i,j} w_k^{i,j} \quad (10)$$

$$G_k^j = \sum_{i=1}^{n_j} (w_k^{i,j})^2 \ , \ P_k^j = \sum_{i=1}^{n_j} w_k^{i,j} \mathbf{x}_k^{i,j} (\mathbf{x}_k^{i,j})^{\mathrm{T}} \quad (11)$$

where X_k^j and S_k^j are unnormalized local estimate and weight. G_k^j and P_k^j are used to compute the estimate errors and to control degeneration. These quantities represent aggregated data on the local subsets of particles. All nodes send these data to their CHs. CH runs the state estimate and sums the weights from its child nodes as follows

$$C_k = \sum_{j=1}^{N} S_k^j \quad (12)$$

State estimate and covariance are computed by

$$\hat{\mathbf{x}}_k \approx \sum_{j=1}^{N} X_k^j / C_k \ , \ P_k = \sum_{j=1}^{N} P_k^j / C_k - \hat{\mathbf{x}}_k \hat{\mathbf{x}}_k^{\mathrm{T}} \quad (13)$$

Then estimate and covariance of CH are transmitted to the sink node at each sample time. On each node in the cluster, state estimate is updated by

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) \propto \hat{\mathbf{x}}_k \quad (14)$$

On the CH, the effective particle number is computed by

$$N_{eff} = C_k^2 / \sum_{j=1}^{N} G_k^j \quad (15)$$

If N_{eff} is smaller than the advance defined threshold, each child node performs local resampling independently as Eq. (7). However, to maintain the consistency of particle, a global resampling is needed periodically by

$$\left\{\mathbf{x}_k^{i,j}, w_k^{i,j}\right\} \rightarrow \left\{x_k^{l,j}, 1/\sum_{j=1}^{N} n_j\right\} \quad k = mC \quad (16)$$

When a previous cluster is destroyed and a new one is generated, the particle information must be transmitted from the former to the latter. A Gaussian mixture model (GMM) is used to describe particles[5].

$$p_k^i \sim \hat{p}_k^i = \sum_{j=1}^{N} \hat{\lambda}_k^j N(\hat{\mu}_k^j, \hat{\sigma}_k^j) \quad (17)$$

where n is the number of Gaussian mixtures, $\hat{\mu}_k^j$ and $\hat{\sigma}_k^j$ are the mean value and covariance of GMM.

3.3 Implementation of DPF algorithm

To reduce the energy cost of the CH, DPF runs only on child nodes. The implementation of DPF algorithm is detailed below.

1) Initialization: If CH(m) has a detection, n_j particles are allocated to jth sensor node. These particles are spread uniformly along the detection geometry in x-y space, viz. $p(\mathbf{x}_0 \mid \mathbf{z}_0)$. Each particle has equal weight.

2) Particle generation in nodes: At sample time k, previous state $p(\mathbf{x}_{k-1} \mid \mathbf{z}_{1:k-1})$ has been known. the state predicted by Eq.(1). When present observations are available, the likelihood is computed by Eq. (3).

3) Estimate computation in nodes: Weights sum is computed at each node by Eq. (10). Node cumulative estimate is computed Eq. (11).

4) Up network communication: The aggregated data (S_k^j , X_k^j , G_k^j , and P_k^j) are transmitted to the CH.

5) Estimate computation in CHs: Collecting aggregated data from N sensor nodes, and the local estimate is obtained by Eq. (6). Weights are summed up from their children nodes by Eq. (12). State estimate and covariance are computed by Eq. (13). If the effective particle number Eq.

(16) gets lower than a given threshold, the flag *f* is set which makes nodes resample the particle same as Eq. (17).

6) Down network communication: CH state estimate $\hat{\mathbf{x}}_k(m)$ and resample flag *f* are transmitted to each node in the cluster.

7) Global estimation: Estimate and covariance data of each CH are transmitted to the sink node, which constructs the estimate sequence according to the time order. Then the optimal smooth method is used by data fusion.

8) Particle exchange: When a previous cluster is destroyed and a new one is generated, the particle sets are transmitted from the former to the latter by GMM.

4. SIMULATIONS

To test the performance of the DPF, it is compared with the CPF in the same simulated scene. 100 wireless sensor nodes consist of a passive tracking scene. Each sensor node is modeled as a passive sensor to get bearings of the target. Nodes are located randomly between coordinates (0, 0) and (10000, 10000). The deployed sensors had a detection radius of 500m. The initial tracking conditions are given as follows:

(a) The sampling period T = 10s.

(b) The simulated target performs random constant acceleration and coordinated turn movement to give the noisy observations.

(c) The probability of false alarm in detecting a signal from the sensor was P_d= 0.05.

(d) The original motion parameters of the target are $\mathbf{x}_0 = [5000, 8660, 10, 6]$.

(e) The CPF and DPF have same particle number *N*=500.

(f) Define the turn rate set as [-2, 0, 2] degree/sec. The transition probability matrix is selected as

$$P_{ij} = \begin{bmatrix} 0.98 & 0.01 & 0.01 \\ 0.01 & 0.98 & 0.01 \\ 0.01 & 0.01 & 0.98 \end{bmatrix}$$

The algorithm of the DPF presented is right through the simulation results compared with the CPF. Trajectory tracking is shown in Fig. 2.. Results show that two PF algorithms have similar good performance in tracking maneuvering target. From Fig. 3., the DPF yields almost the same accuracy of state estimation as the CPF. However, as shown in Fig. 4., average number of communication bits in the CPF increases proportionally to particle number, while the increase speed of the DPF is slowed down. Finally, communication amount of the DPF is 60% less than the CPF. Thus, the DPF is much better than the CPF in communication cost if PF has large particle set.

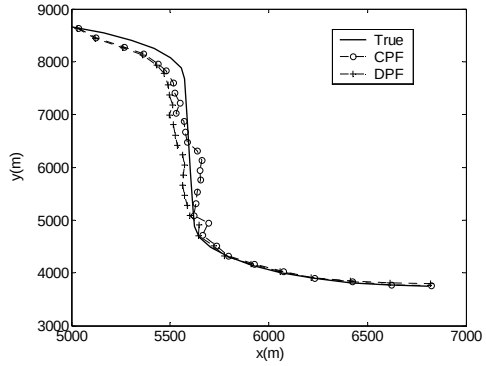


Fig. 2. Tracking comparison of CPF and DPF

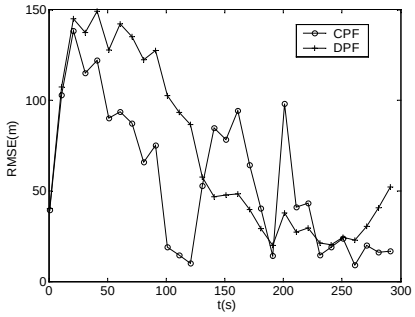


Fig. 3. Range RMSE of CPF and DPF

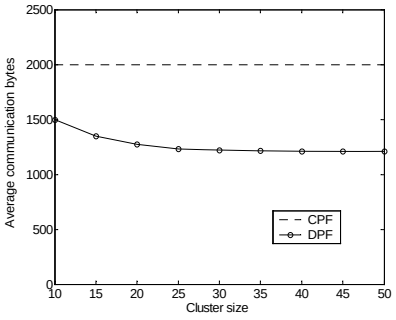


Fig. 4. Average number of bytes per sample time

5. CONCLUSIONS

We proposed a distributed PF algorithm for passive tracking in WSNs. Due to the dynamic clustering structure, the energy cost of communication and processing is balanced in WSNs. Hence, during the tracking, the central node failure is avoided. Compared with the centralized particle filter, the DPF shows the same performance in tracking accuracy but lower communication costs. However the proposed DPF still needs much communication cost between CHs and their children, thus the future researches should focus on how to reduce communication cost further to realize light-weight distributed tracking.

6. REFERENCES

[1] C.Y. Chong, S.P. Kumar, "Sensor Networks: Evolution, Opportunities, and Challenges", Proceedings of the IEEE, Aug. 2003, 91(8): 1247-1256.

[2] S. Arulampalam, S. Maskell, "A Tutorial on Particle Filters for On-line Non-linear/Non-Gaussian Bayesian Tracking", IEEE Transactions of Signal Processing, Feb.2002， 50(2): 174-188.

[3] M. Bolic, P.M. Djuric, et al, "Resampling algorithms and architectures for distributed particle filters", IEEE Transactions on Signal Processing, July 2005,53(7): 2442-2450.

[4] M. Efe, D.P. Atherton, "Maneuvering Target Tracking Using Adaptive Turn Rate Models in The Interacting Multiple Model Algorithm", Proceedings of the 35th Conference on Decision and Control, Vol.3, Kobe (Japan), Nov. 1996, 51-56.

[5] X. Sheng, Y.H. Hu, et al, "Distributed particle filter with GMM approximation for multiple targets localization and tracking in wireless sensor network", Fourth International Symposium on Information Processing in Sensor Networks, April 2005, 18-188.

Real-time Gate Signal Choosing Technology*

Tao Shen, Jianshe Song
the Second Artillery Engineering College Section 602,xi'an Shannxi 710025,China
Email: luckyshentao@163.com

ABSTRACT

According to the semi-active laser guidance weapon's technology status that its impulse repeating frequency coding is easily interfered, the possibility of high repetition rate laser interference with semi-active laser guided weapons operating based on real-time gate signal was studied. The technology principal of laser impulse width coding is analyzed, and the influence of real-time gate signal is discussed.

Keywords: Laser Guidance Weapon, Real-time Gate Signal Code Technology, Impulse Width.

1. INTRODUCTION

At present, the laser guidance weapon especially the half initiative laser guidance weapon system adopts the signal coding technique mostly, the laser guide leads look for the target by using the reflected laser signal , but the laser target indicator sends the laser designation signal constantly toward the target consecution in the meantime, also exposed to threat easily. In addition, the guide head separates from the indicator that makes it difficult to keep synchronous strictly in the transmission and receive, and this makes the guide lead to be subjected to beguilement interference easily. The real-time date signal choosing technology mainly set up impulse matriculate gate in the tracking system, the object indicative signal adopts coding, and this can distinguish itself guide signal correctly among the many guide signal and interfere signal.

2. REAL-TIME GATE SIGNAL CHOOSING TECHNOLOGY

To the laser half initiative guide weapons, the indicator sends out the coding impulse signal to the object, and the object reflects signal to the guide head of the weapon, at last the laser weapon destroyed the object. The laser guide weapon adopts laser coding and decoding technology to solve the interference each other in the same area. Each indicator has its own coding style, and uses different code to different object, and the guide bomb can fly to its own object independently with the decoding ability in accordingly.

At present, many laser guide weapon systems lead to code technique to identify the indicator signal, also combine real-time gate signal to choose own synchronous signal, make the system anti- interference ability strengthen, not accept non-synchronously interference signal. The main principle of the real-time gate signal is the single impulse signal as the synchronous point of departure that opens a door time next time. The advantage of this is the total amount error margin that removed and the gate door can designed very narrow.

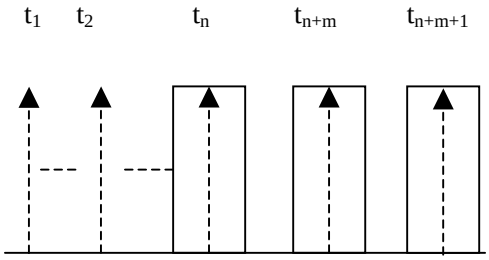


Fig. 1. Real-time Gate

3. THE PRINCIPLE OF DECODE IMPULSE WIDTH

The selection of the decoding impulse width in the laser receives of the laser guidance system has something to do with the wave door of the impulse coding time. And the selection of code time wave door of the impulse is smaller, the guide system resists various interference stronger, but the width of the impulse code time wave door is not infinitesimal, it has to consider the error margin in fixed time of the laser blast-off machine and during the deliver time. The laser impulse of the laser guide head received has certain time deviation with the standard laser impulse. Through the analysis, the deviation can be expressed with the following type:

$$T = \Delta T_1 + \Delta T_2 + \Delta T_3$$

T -- the time deviation between the laser impulse of the laser guide head received and the standard laser impulse

ΔT 1--impulse interval changing time that the frequency stability of the coder crystal oscillator

ΔT 2-- the time deviation between the laser impulse eradiation and trigger impulse

ΔT 3--the time deviation between the received impulse and standard impulse.

From the type, it is very important to choose the decoding impulse width correctly based on the coding impulse width. As the factor caused by the laser frequency whipping, optical path difference and the clock stabilization, the decoding impulse width should be 30-50 microsecond.

4. ANALYSIS OF THE TECHNICAL PARAMETER

4.1 REPETITION FREQUENCY

To the laser weapon must generate impulse frequency $f \geq 1/\tau$ interference. Because the real-time signal gate is the

* This work is supported by the National Nature Science Foundation of China under Grant NO: 60272022

first matriculated signal, if the real-time signal gate and the pilot signal are strictly synchronous, that is to say, the impulse peak appears at the middle of the wave gate, the interference impulse repetition frequency f $\geq 1/\tau$ can interfere 100%. Because of the effect of many factors, the impulse peak of the pilot signal can appear at the any time of the wave gate appeared time τ .Supposed the probability between the pilot signal and the interference signal appeared in the time τ is equality, when f=$2/\tau$, the interference probability is:

$$\eta = \frac{1}{2}\int_{-\tau/2}^{0} \frac{\tau/2+x}{\tau/2}\frac{dx}{\tau/2} + \frac{1}{2} = 75\%$$

in fact, the impulse peak of the pilot signal should be controlled to appear at the middle of the wave gate to ensure the pilot signal can be received by guide head. If the pilot signal can be controlled in the

[—$4/\tau$,$4/\tau$], the interference probability is:

$$\eta = \frac{1}{4} + \frac{1}{2}\int_{-\tau/4}^{0} \frac{\tau/4+x}{\tau/4}\frac{dx}{\tau/4} + 1/2 = 87.5\%$$

so, the repetition frequency of the laser

f=$2/\tau$ can get the interference request. From the width of the real-time signal gate is 30 us- 50 us, the repetition frequency of the interference laser should be 40kHz- 70kHz.

4.2 REQUEST OF POWER

If it makes use of a false target to carry on interference, namely, light upon the interference signal to the false target, make the diffuse reflection signal enter into guide head to implement interference, suppose the interference machine to output the power as P_1, enter into guide head a the power as P_2, then

$$P_2 = \frac{t_1 t_2 \cos\theta_1 e^{-\mu(R_1+R_2)} \rho_1 P_1 A_r}{\Omega_1 R_2^2}$$

t1--the transmission of the laser eradiate system

t2-- the transmission of the guide head receive system

θ_1 --the entrance angle of the laser signal irradiation false target

μ --the modulus decay of the atmosphere

ρ_1 --the reflectivity of the false target

R1—the distance of the laser to false target

R2—the distance of the guide head to false target

Ar—the receive area of the guide head

Ω_1 --the reflect solid angle of the false target

Suppose the power of the indicator as P_3, the power of guide head received as P_4, the relation as follow:

$$P_4 = \frac{t_2 t_3 \cos\theta_2 e^{-\mu(R_3+R_4)} \rho_2 P_3 A_r}{\Omega_2 R_4^2} \quad (4)$$

t_3 -- the transmission of the indicator system

θ_2 -- the entrance angle of the indicator signal irradiation false target

ρ_2 -- the reflectivity of the target

Ω_2 -- the reflect solid angle of the target

R_3 --the distance of the indicator to target

R_4 -- the distance of the guide head to target

type（4）shows the relation of P_3, P_4 when the target is big object. What is called the big object is the area that the indictor irradiatesS1 $\geq$ S2the indictor faucal area. If S1 $\leq$ S2, type（4）can be changed as follow:

$$P_4 = \frac{t_2 t_3 \cos\theta_2 e^{-\mu(R_3+R_4)} \rho_2 P_3 A_r S_1}{\Omega_2 R_4^2 S_2} \quad (5)$$

for valid interference, P2 $\geq$ P4, from type (3)(4)(5) :

$$\frac{P_2}{P_4} = \frac{t_1 \cos\theta_1 e^{-\mu(R_1+R_2)} \rho_1 P_1 \Omega_2 R_4^2}{t_3 \cos\theta_2 e^{-\mu(R_3+R_4)} \rho_2 P_3 \Omega_1 R_2^2} \quad \text{(big object)}$$

$$\frac{P_2}{P_4} = \frac{t_1 \cos\theta_1 e^{-\mu(R_1+R_2)} \rho_1 P_1 \Omega_2 R_4^2 S_2}{t_3 \cos\theta_2 e^{-\mu(R_3+R_4)} \rho_2 P_3 \Omega_1 R_2^2 S_1} \quad \text{(small object)}$$

to type （6） （7） :R2 >>R1 、 t1 $\approx$ t2 、 $\theta_1 \approx \theta_2$、 $R_2 \approx R_3 \approx R_4$、 $\rho_1 \approx \rho_2$, both the true and false object are diffuse reflection. so, $\Omega_1 = \Omega_2 = 2\pi$ 。 type（6）、（7）as follow：

$$\frac{P_2}{P_4} = \frac{P_1}{P_3} e^{\mu R_2} \geq 1 \quad \text{(big object)}$$

$$\frac{P_2}{P_4} = \frac{P_1}{P_3} e^{\mu R_2} \times \frac{S_2}{S_1} \geq 1 \quad \text{(small object)}$$

the power request to the laser machine can be get as follow:

$$P_1 \geq P_3 e^{-\mu R_2} \quad \text{(big object)}$$

$$P_1 \geq P_3 e^{-\mu R_2} \times S_1 / S_2 \quad \text{(small object)}$$

5. CONCLUSION

This text from the theories and the technique request analyzes adoption up solid hour wave the door chose the correspondence number technique to carry on the laser half to make to lead the possibility that the weapon system lead actively, from analysis the pulse frequency is more high, output the power is more big, the result of the interference more good.

6. REFERENCES

[1] F. Huang， Analysis of Pulse Width Coding Anti-interfering Technology Apply on Semi-active Laser Guidance Weapon [J], Laser & infrared, 2005.2

[2] X.C. Xie， Research on Decoding Pulse Width of Laser Seeker [J] ， Electro-optic technology

application ， 2005.8.
[3] B.Q. Yang，Y. Chen， Study on Evaluation Methods of Laser Angle Deception [J], Electro-optic technology application ， 2005.8.
[4] New kind scheme of laser code in the laser terminal guidance ammunition round [J], Infrared and Laser Engineering, 2005.10
[5] Z. N. Hou， Research on laser deception jamming technology [J], optical application, 2002.1

A Scheme on How to Solve Anycast Scalability

Xiaonan Wang[1,2]
[1.] **Computer and Technology School, Nanjing University of Science & Technology**
[2.] **Software Engineer Department, Changshu University of Technology**
Nanjing, Jiangsu 210094, China
Email: wxn_2001@163.com
And
Huanyan Qian
Computer and Technology School, Nanjing University of Science & Technology
Nanjing, Jiangsu 210094, China
Email: hyqian@mail.njust.edu.cn

ABSTRACT

At the present time, the communication models of implementing Anycast service are all confronted with one problem: scalability. Therefore a new kind of Anycast communication model is proposed in this paper, which utilizes a new kind of Anycast communication method and a new kind of address allocation. This communication model not only solves Anycast scalability but also allows Anycast members to freely leave and join Anycast group. In this paper the feasibility and validity of this communication model are deeply analyzed and its error tolerance and performance are discussed according to the experiment data in IPv6 emulation, which argue that this model is able to support the large-scale Anycast group.

Keywords: IPv6, Anycast, ICMPv6, Router, Address.

1. INTRODUCTION

The original definition of Anycast in RFC1546[1] is: "A host transmits a datagram to an Anycast address and the internet is responsible for providing best effort delivery of the datagram to at least one, and preferably only one, of the servers that accept datagrams for the Anycast address." In practice, Anycast is a point-to-point flow of packets between a single client and the nearest destination server identified by an Anycast address. The idea behind Anycast is that a client just wants to send packets to a server offering a particular service or application, but in his opinion the fact that which server is chosen is not important. To accomplish this, a single Anycast address is assigned to one or more servers within a so-called Anycast group. In Anycast, the communication between client and server are the same as the one in Unicast for client and server are unaware that Anycast is used. Thus, when a host sends packets with Anycast address as its destination address to one server of Anycast group these packets will attempt to be delivered to the nearest server to the source host [2].

Anycast can support various web services. For example, client would no longer have to be configured with the IP addresses of DNS, but rather could send a request to a well-known DNS Anycast address, which makes it easy for that client to move to any new networks by his pleasure.

To sum up, it can be easily inferred from the previous words that Anycast service is very important and useful and can be applied to many fields. With the increasing growth of new network applications and services, Anycast service is needed urgently. However, the study on Anycast is still on the primary stage so inevitably there exist many problems in fulfilling Anycast services, one of which is scalability.

2. ANYCAST SCALABILITY ANALYSIS

IPv6 allocates Anycast addresses from Unicast addresses so Anycast address is not distinguishable from Unicast address, which makes it possible to mitigate routing table expanding explosion caused by Anycast service and in the meanwhile makes full use of the existing routing resources. Traditional Unicast routing can aggregate routes to destinations with the same prefix into one routing entry, which makes it possible for Unicast to scale. However, unfortunately Anycast defies this kind of hierarchical aggregation.

An Anycast address represents a node group with the same characteristics whose members can be distributed to every Internet region. If Unicast routing protocol is used to route Anycast packets each member of global Anycast group will be advertised as a separate entry in all the relevant routing tables, which makes the routing tables expand proportionally to the number of the global Anycast groups in the entire Internet so that it results in Anycast scalability.

IPv6 confines all the members of each Anycast group to particular topological regions with the same address prefix. Within these regions each member of Anycast group is advertised as a separate entry in the routing tables and outside the regions all the members of Anycast group are aggregated into one routing advertisement for the same prefix. By confining each member of Anycast group to a predetermined region, IPv6 lessens Anycast scalability problem to some extent, but it does not radically solve it.

In this situation, a scheme on solving Anycast scalability problem is proposed in this paper. The following sections give a detail analysis and discuss of this scheme.

3. COMMUNICATION MODEL

This scheme is built on the basis of layered structure of IPv6 address

Anycast address model adopted by IPV6 is totally different from the one recommended by RFC1546. The former suggests that Anycast addresses be allocated from Unicast addresses so that Anycast address structure is not distinguishable from Unicast one, but the latter recommends that Anycast address adopt an independent model [3]. This scheme chooses the latter recommendation.

IPv6 address format is different from Ipv4 one, and it consists of 8 address fragments each one of which contains 16 bits. In addition, IPv6 still designs a layered address structure. IPv6 address consists of 3 layers which divide the whole Internet into different domains. The first layer in IPv6 address divides the entire Internet into some domains. On

the basis of partition on the first layer the second layer divides the domain on the first layer into some sub-domains. And on the basis of partition on the second layer, the third layer divides the domain on the second layer into some sub-domains [4]. According to this kind of layered address structure, the address format in this scheme is proposed, just as the following Fig. 1. Shows.

3	13	8	24	16	64
FP	TLA	RES	0	SLA ID	Interface ID
FP	TLA	RES	NLA	0	Interface ID
FP	TLA	RES	NLA	SLA ID	0

Fig. 1. Anycast Address Partitoin

From the above figure it can easily inferred that only in local routing table is Anycast address distinguishable from Unicast address and in the routing table on the topper layer Anycast routing information and Unicast routing information can be aggregated into one entry. As a result, Anycast address never appears as a separate entry in the routing table on the topper layer, but aggregates Unicast address in the same domain into one entry in the routing table. For example, a router on the second layer domain only needs to store and maintain routing information on Anycast address assigned on the second layer, and routing information on Anycast address assigned on the third layer but included in the second layer address space never appears in the routing table of this router. Therefore, the process of dealing with Anycast service in this scheme hardly influences the performance of Unicast service but this Anycast address partition effectively solves the expanding explosion of Anycast routing table.

According the previous analysis, the scheme proposes the following Anycast communication model, just as the Fig. 2. Shows.

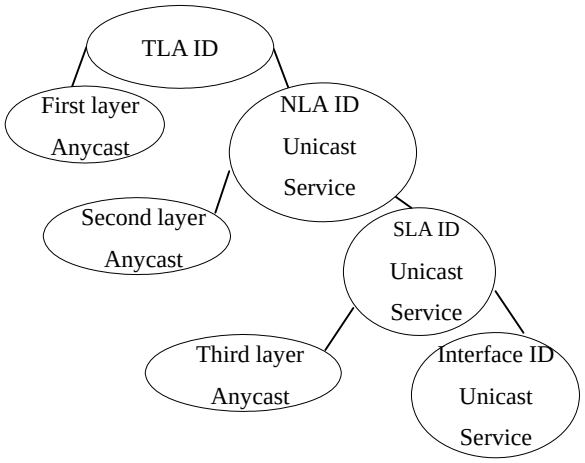


Fig. 2. Anycast Communication Model

Since the communication model is layered the Anycast address on the different layer may be assigned to different service according to that service scale magnitude. In general, the service with the large-scale magnitude may apply for the Anycast address on the first layer. Thus, all the Anycast members may be distributed across the entire domain identified by that Anycast address. And the service with the small-scale magnitude may apply for the Anycast address on the third layer.

In this scheme the domain on every layer is configured with one (or more) controller(s) who is responsible for maintaining Anycast tree (topological construction on IP layer) on the local layer domain, finding out the Anycast member with the shortest path, fulfilling the transition from Anycast address to Unicast address, authenticating the identity of new hosts applying for joining Anycast group, supervising the status of all the Anycast members on the local layer, and so on. All the controllers are distributed all over the entire network to form a large distributed system.

If a host wants to join an Anycast group it must know the Anycast group addresses [5,6]. A host takes the following steps to join an Anycast group:

1) A host lodges a request for joining Anycast group with Anycast address as its destination address;

2) After this request is routed to the controller (reference to the following chapters for routing details), if controller accepts the host as its new Anycast member it will reply to the host with an Accept message, and insert the new member into its Anycast tree according to the routing information and store its relevant information (for example, new member's Unicast address), or controller will reply to the host with Reject message;

3) After the host receives the reply message it will check its type. If it is an Accept message the host will broadcast its new identification as Anycast member to all the member of the local domain, or it will abandon it.

The above process may be achieved by adding new types of messages to BGP and IGMP [7]. To avoid malicious attacks, the information interactions between hosts and routers should take some security measures.

The entire Internet may be considered as a domain set so routing a packet from source host to destination host includes two instances: routing between domains and routing in one domain [8].

Before a host sends a request for service it first checks if the requested service is Anycast service according to the destination address. If it is Anycast service the host will send an address-transition packet (transition from Anycast address to Unicast address) with Anycast address as its destination address. Here, if Anycast address belongs to the local domain the host will send this packet to the local domain border router. After the domain border router receives this packet it can learn that it is an Anycast packet according to the type of destination address so it will deliver it to the local controller to further deal with it.

If the destination address is not within the local domain the routing of this address-transition packet will pass three consecutive parts: the source domain routing from source host to source domain border router; the backbone routing from source domain border router to destination domain border router; the destination domain routing from destination border router to destination host. In the first and second parts, the Anycast routings are the same as Unicast ones. But in the third part, after the destination domain border router receives the address-transition packet it will check the type of Anycast address. If the value of the next domain address is zero the router will deliver the packet to the local controller or it will continue to route this packet in the normal way.

The following words describe how the controller deals with the address-transition packet. First, controller searches its local Anycast tree for the Unicast address of the member which is the nearest to the source host (in this scheme, the so-called distance between the Anycast member and source host is calculated according to topological construction on IP layer. Here according to different requests different measurement methods may be adopted). Then, the controller encapsulates the acquired Unicast address of Anycast member into a reply packet and sends it to the source host.

After the source host receives the reply packet it begins to communicate with Anycast member using its Unicast address.

Here the Anycast member with the shortest path may be unreachable. In this situation, the scheme proposes different disposal processes according to service quality and scale. If service quality is preferred, before controller sends reply packet to the source host, controller will first send a test packet to the Anycast member with the shortest path to confirm that it is reachable. If the member is unreachable the controller will go on searching its Anycast tree for the second nearest member to the source host and mark the unreachable member (for further test) to update the tree information. The above steps will repeat until a reachable member is found. In some extreme situations, all the members of Anycast group may be unreachable. In this situation the controller may send an ICOMv6 error message to source host notifying it that the destination server is unreachable.

If real-time attribute of Anycast service is preferred controller will directly encapsulate the unicst address of Anycast member with the shortest path into a reply packet and send it to source host. If the member is unreachable the network system will deal with this error in the normal way. In this situation, controller needs to update tree information by regularly querying the status of all the members of Anycast tree in order to ensure service quality.

In this communication model, a host requests for Anycast service according to the following steps:

1) Host acquires Anycast address by DNS;.

2) Host sends an address-transition packet with Anycast address as its destination address;

3) Controller deals with this packet and obtains the Unicast address of Anycast member with the shortest distance to source host, and sends this Unicast address to the source host;

4) After receiving this Unicast address host begins to communicate with Anycast member using the acquired Unicast address.

Since the communication model adopts layered structure a controller only needs to maintain the data and information on Anycast members in the same layer domain. Therefore, Anycast has little influence on the entire network performance.

In this scheme it is very important for a controller to provide source host with correct and valid Unicast address of Anycast member. The following words discuss how to ensure quality and validity of service in some unexpected situations:

1) In this scheme the role of controller is very important so this scheme adopts mirror controller to avoid some unexpected instances, such as abrupt breakdown of controller, etc;

2) In some extreme situations, all the members of Anycast group may be unreachable. In this situation the controller sends an ICOMv6 error message to source host notifying it that the destination server is unreachable;

3) Anycast member may become unreachable when host uses Anycast member's Unicast address to communicate with it. In this situation, the network system may deal with it in the normal way.

4. PERFORMANCE ANALYSIS

This performance analysis refers to the comparison between Anycast service performed on IP layer and in this scheme in simulated IPv6 environment.

It is assumed that the topological structures of all the Anycast members in these two instances are the same. The parameter *ShortestPath* refers to the shortest path value acquired in this scheme and *ShortestPathIP* refers to the one acquired on IP layer, and the following formula is adopted to analyze the Anycast performance.

$$R = ShortestPath / ShortestPathIP \quad (1)$$

It can be inferred that more approximate to 1 R is lower the layer of Anycast address is. According to the experiment data the average value of the shortest paths acquired on Anycast address on the second layer in this scheme stays within 1.2 times of the one obtained on IP layer.

In this scheme, an address-transition mutual process is added, but that process has little influence on backbone network performance since the data transmitted in this process are very small. In addition, the Anycast routing table in this scheme is manageable because the backbone routers only maintain the Unicast routing information. In the scheme, the fact that Anycast addresses are distinguishable from Unicast addresses means that Anycast service can be fulfilled separately from Unicast service. As a result, Anycast service hardly affects the performance of Unicast service. The most important is that this scheme solves the Anycast scalability problem.

5. CONCLUSION

Anycast is a new characteristic of IPV6 and supports various kinds of services. In IPV6 simulation, a new communication model on performing Anycast services is brought forward and its validity and feasibility are analyzed and discussed. As a new kind of communication model, it is promising, but there still exists many problems in performing Anycast services and these problems need further study and analysis.

6. REFERENCES

[1] C. Partridge,T. Mendez, et al, "Host Anycasting service", RFC 1546, 1993.

[2] S. Deering, R. Hinden, "Internet Protocol Version 6 (Ipv6) specification", RFC 2460, 1998.

[3] R. Hinden, S. Deering, "IP version 6 addressing architecture", RFC 2373, 1998.

[4] D. JohnSon, S. Deering, "Reserved Ipv6 Subnet Anycast addresses", RFC2526, 1999.

[5] I.H. Jurrichiro, K. Ettikan, "An analysis of Ipv6 Anycast Internet Draft", Internet Engineering Task Force, 2001.

[6] D. Katabi, J. Wroclawski, "A framework for scalable global IP-Anycast(GIA)", In: Proc of SIGCOMM, New York: ACM Press,2000, 3-15

[7] T. Narten, E. Nordmark, et al, "Neighbor discovery for IP version 6 (Ipv6)", RFC 1970, 1996.

[8] C. Huitema, "Routing in the internet", Prentice Hall, 1996.

An Efficient Round Robin Scheduling Algorithm for DiffServ Networks *

Gongchao Su, Xiaohui Lin and Hui Wang
College of Information Engineering, Shenzhen University
Shenzhen, Guangdong Province, P.R.China
Email: {gcsu, xhlin, wanghsz}@szu.edu.cn

ABSTRACT

There have been a lot of studies on designing scheduling algorithms for differentiated service (DiffServ) networks. In this context three important issues have to be considered: simplicity, fairness and priority-based services. This paper introduces an efficient flow-based scheduling algorithm, called DS-DRR, for DiffServ networks. By adopting appropriate sorting strategies, this approach can provide low latencies for traffic of high priorities. Moreover, this scheme inherits advantages such as simplicity and fairness from traditional Deficit Round Robin schemes. Performance analysis and simulation results show that this scheme ensures strict priority-based services under dynamic network scenarios, while maintaining fairness for traffic of all priority service classes.

Keywords: Differentiated Service, Scheduling, Round Robin, Service Priority, Latency.

1. INTRODUCTION

Nowadays the Internet provides access to a wide variety of services ranging from ordinary mail systems to real-time audio and video communications. These applications put highly varied requirements on network transport functionality, which can be characterized by a set of parameters commonly referred to as Quality-of-Service (QoS) requirements. To implement QoS in current IP networks, the Differentiated Service architecture has been proposed by the IETF [1]. The Differentiated Service architecture defines multiple service classes by setting DS parameters in packet headers. QoS are provided by routing nodes that deploy certain queuing disciplines. Many fair scheduling algorithms have been developed [2, 3, 4, 5, 6, 7], so that traffic of higher classes can have more bandwidth and get better services.

One of the most common and widely accepted fair scheduling schemes is the Deficit Round Robin (DRR) algorithm. It exhibits O(1) complexity[2] and can be easily implemented in hardware. However, it does not ensure QoS for multi-class services, in terms of delay under dynamic traffic conditions. When it is applied in differentiated service architecture, this becomes a problem for delay-sensitive applications such as real-time multimedia communications.

This paper presents the DS-DRR scheme for fair scheduling in DiffServ networks. The objective of our approach is to ensure traffic of higher priority classes can experience lower delay in DRR architecture. By adding alternative active list for traffic flows and deploying appropriate sorting strategy, our technique allows DRR to provide lower delay for higher traffic classes, while maintaining the fairness and complexity.

The rest of the paper is organized as follows. Section II briefly introduces the DRR algorithm. The DS-DRR scheme proposed is described in Section III. Performance evaluation and simulation results are presented in Section IV. We draw conclusions in Section V.

2. RELATED WORK

Round-robin algorithms are flow-based fair packet scheduling algorithms and have received intensive studies in literature [2, 3, 5, 6, 7].Compared with other round-robin algorithms such as WRR [3], Deficit Round Robin (DRR) exhibits good fairness and has O(1) complexity under specific conditions. In DRR, backlogged flows are stored in a FIFO list (called the Active List). Each flow within the list is serviced cyclically. Each flow is assigned a given quantum, which stands for the maximum amount of traffic to be serviced in a round, and a deficit counter, which specifies the portion of the flow's quantum that has not been used within a round. The counter is decreased by the length of each transmitted packet, and transmission of a packet is allowed as long as the counter does not fall below zero, when a flow is serviced. After the flow's turn ends the remaining deficit is carried over onto the next round as a credit. By setting the quantum parameters for each flow in the Active List, the DRR algorithm can support bandwidth reservation for various flows, and provides good fairness among these flows.

However, the DRR scheme does not provide packet delay differentiation for traffic flows [5, 6, 7], while delay time is very crucial for real-time communications. This indicates that the DRR approach can not sufficiently support real-time traffic, which is usually of high priority among aggregated traffic flows.

A modification of DRR called DRR+ [2] has been proposed to provide latency-bounded services to real-time traffic. In DRR+, flows are roughly separated into two classes: latency-critical and best-effort. Latency-critical flows have a lower latency bound than best-effort flows. This scheme provides a framework to reduce latency bound for real-time traffic. However, DDR+ does not provide differentiated services to flows of the same class (latency-critical or best-effort) and does not take into account the priorities of traffic flows. This makes it inapplicable for DiffServ networks.

3. THE DS-DRR SCHEME

From our previous discussion we note that the Active List determines the sequence each flow is served within a round. Apparently, traffic delay is directly influenced by the order

* This work is partially supported by Nature Science Foundation of Guangdong Province (No.5010494) and Shenzhen Sci. and Tech. Plan (No.200334).

it is served: The earlier it is served, the less delay it will experience. The Active List in traditional DRR scheme is maintained as a FIFO list: each active flow is inserted into the tail of the list once it has been served. This prevents the flows with high priority to occupy the head of the list in order to get lower delay.

The DS-DRR approach we propose aim to maintain the Active List in such an order that flows with higher priorities are always on the top of the list. To achieve it, we propose to maintain two lists in DS-DRR approach: the Active List and the Backup List. Flows within the Active List are served in a FIFO order, and once any flow has been served, it is immediately removed from the Active List and inserted into the Backup List. The Backup List is maintained in a priority-based order: Flows with higher priorities are always on top of the list. Once all the flows in the Active List are served (the Active List becomes empty), the Backup List is directly designated as the Active List and afterwards the Backup List becomes empty.

As a comparison with [2], we present the pseudo code for DS-DRR in Fig. 1.It is noted that compared with DRR approach, the DS-DRR approach change the sequence that flows are served, but do not change the chances that flows are served within a round. This indicates the DS-DRR approach maintains the same fairness among flows as it is done in DRR scheme. We prove the conclusion in various simulation scenarios. However, since flows with higher priorities are always on top of the Active List, packets of these flows experience less delay. We discuss this issue in more detail.

```
Initialization:
    For (i=0;i<n;i=i+1)
      DC_i=0;
Enqueuing module: on arrival of packet p
    i=ExtractFlow(p)
    If ((Exists in Activelist or Backuplist)==FALSE)
    Then InsertBackupList (i); /*add into BackupList by priorities of flow i */
      DC_i=0;
    Enqueue(i,p); /*enqueue packet p to queue i */
Dequeuing module:
    While (TRUE) do
      If ActiveList becomes empty then
        Exchange ActiveList with BackupList
      If ActiveList is not empty now then
        Remove head of ActiveList, flow i
        DC_i=Q_i+DC_i;
While ((DC_i>0) and (Queue_i not empty)) do
          PacketSize = Sizeof (Head of Queue_i);
          If (PacketSize <DC_i) then
            Send(Dequeue(Queue_i);
            DC_i=DC_i-PacketSize;
          Else Break;
        If (Queue_i Empty) then
          DC_i=0;
        Else InsertBackupList(i);
```

Fig. 1. Pseudo code for DS-DRR

Discussion: Fig.2 depicts a scenario of N flows backlogged in the router, with M flows listed in the Active List ($M \leq N$). Handling of the round robin list in both DRR and DS-DRR schemes are illustrated in Fig. 2.

Consider a packet p for a new flow, denoted as f_{N+1}, arrives at the router. Each flow has a predefined quantum Qi, $i \in [1,N+1]$. Define the backlogged delay T_{N+1} for f_{N+1} as the time interval between the time t_1 when the first packet of f_{N+1} arrives at the router, and the time t_2 when this packet is sent out. The router has a constant service rate of C.

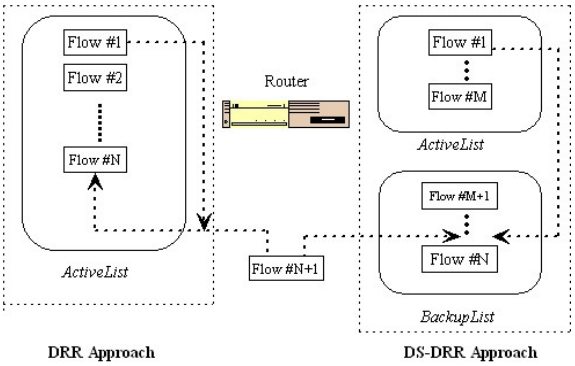


Fig. 2. Comparison of DRR and DS-DRR

.

In DRR scheme, T_{N+1} can be denoted as:

$$T_{N+1} = \sum_{i=1}^{N} Q_i / C \tag{1}$$

In DS-DRR scheme, if flow f_{N+1} has the highest priory compared with other flows currently in the Active List, T_{N+1} can be denoted as:

$$T_{N+1} = \sum_{i=1}^{M} Q_i / C, M \leq N \tag{2}$$

If flow f_{N+1} has the lowest priory compared with other flows currently in the Active List, T_{N+1} can be denoted as:

$$T_{N+1} = \sum_{i=1}^{M+N} Q_i / C, M \leq N \tag{3}$$

For more comparison, we assume $Q_i=Q_j=Q$, $i,j \in [1,N]$, equation (1) can be rewritten as:

$$T_{N+1} = NQ / C \tag{4}$$

For DS-DRR scheme, T_{N+1} can be denoted as:

$$T_{N+1} = MQ / C, M \leq N$$

Since each flow has the same quantum and the router has a constant rate, each flow occupies the same time portion in a round. Thus it can be seen that for any possible value of M,

$$P[M = m] = 1/N, m \in [1, N]$$

For flow f_{N+1} with the highest priority, the average value of T_{N+1} can be denoted as:

$$E[T_{N+1}] = \left(\sum_{m=1}^{N} \frac{1}{N} \times mQ \right) / C = \frac{(N+1) \cdot Q}{2C} \tag{5}$$

Thus it can be seen from equation (1)~(5) that in DS-DRR scheme flows with higher priorities endure lower delay than flows with lower priorities, which is necessary for providing differentiated QoS services for DiffServ networks.

4. SIMULATION ANALYSIS

We build our simulation scenarios and validate our results in OPNET Modeler 8.0 c. The simulated model is depicted in Fig.3. The traffic source model has N flows, each of which sends packets of variable length at a corresponding Poisson rate and has a predefined priority. Ill-behaved flows among these N flows send packets at three times the rate at which other flows send packets. The Schedule model implements DRR or DS-DRR algorithms. We use an infinite buffer size to eliminate the effect of packets drop due to lack of buffers. Statistics are collected in Sink model.

Fig. 3. Node model in OPNET simulation

A. *Fairness in DRR and DS-DRR schemes*

We use the following simulation settings in this scenario, as seen in Table 1. Note that each flow has the same quantum (Q) assigned, and flow #10 is the only ill-behaved flow among the aggregated traffic. Each flow has the same Poisson rate (except for the ill-behaved flow, which has three times the Poisson rate designated), and it is set at different level with DRR implementation and with DS-DRR implementation for comparison. The simulation runs for 10 minutes.

Table 1. Simulation Parameters in Experiment A

	DRR	DS-DRR
Number of Flows (*N*)	10	
Output Link Rate (*C*)	50 Megabytes/s	
Packet Mean Length	1000 bytes	
Assigned Quantum	5000 bytes	
Ill-behaved flow	Flow #10	
Average rate of a flow	4 M bytes/s	4.5 M bytes/s

The actual bandwidth used by each flow is depicted in Fig. 4. It can be seen that both DRR and DS-DRR schemes maintain good fairness among the mixed traffic of ill-behaved flows and other flows, since each flow acquires almost the same share of available bandwidth.

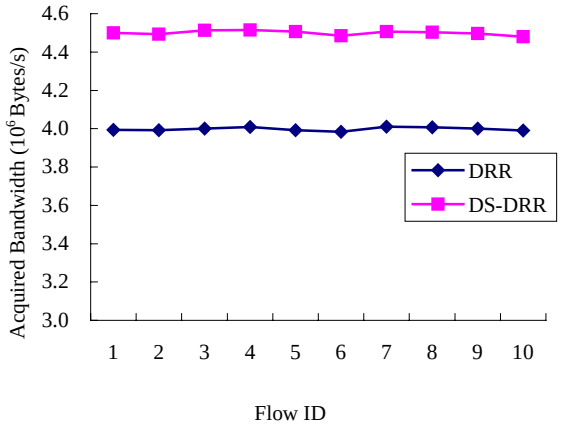


Fig. 4. Bandwidth allocation (DRR and DS-DRR)

B. *Latency differentiation in DRR and DS-DRR schemes*

In this experiment we define a priority sorted in a descending order for each flow. Flow 1 has the highest priority and flow 10 has the lowest priority. Traffic source rates are set at the same level in both the DRR and DS-DRR schemes. Simulation parameters are similar to Experiment A, as seen in Table 2.

Table 2. Simulation Parameters in Experiment B

	DRR	DS-DRR
Number of Flows	10	
Output Link Rate	50 Megabytes/s	
Packet Mean Length	1000 bytes	
Assigned Quantum	5000 bytes	
Ill-behaved flow	Flow #10	
Average rate of a flow	4.5 M bytes/s	4.5 Mbytes/s

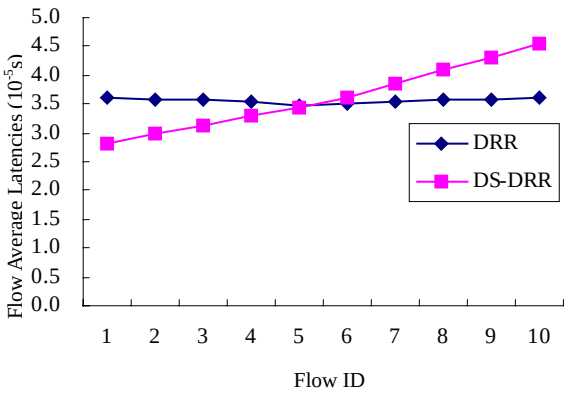


Fig. 5. Flow average latencies (DRR and DS-DRR)

5. CONCLUSIONS

In this paper, we propose the DS-DRR scheme for packet scheduling in DiffServ networks. We demonstrate that by keeping flows of high priorities on the head of the Active List in DRR scheme, flow latencies can vary according to their priorities in round robin schedulers. Our approach inherits the advantages such as simplicity and fairness from DRR scheme, and does not require a major modification to widely deployed DRR implementations. Our scheme can be a reasonable choice to deploy round robin schedulers in DiffServ networks.

We also note that prediction on latency differentiation in our approach depends on quantum size assigned. In a scheduling environment with highly varied quantum size and link rate, latencies experienced by flows can be difficult to remain stable and low. Recent studies [5, 7] claim that for optimal performance, frame size should be smaller and this can be achieved by maintaining multiple Active Lists. Our future work should cover this issue.

6. REFERENCES

[1] S. Blake, “An Architecture for Differentiated Service,” IETF RFC 2475, December 1998.

[2] M. Shreedhar, G. Varghese, “Efficient fair queuing using deficit round- robin,” IEEE/ACM Tran. Networking,Vol.4, pp.375-385, June 1996

[3] M. Katevenis, S. Sidiropouloset, “Weighted round robin cell multiplexing in a general-purpose ATM switch chip,” IEEE Journal on Selected Areas on Communications, Vol. 9, Issue 8,pp.1265--1279, October 1991.

[4] N. McKeown, “The iSLIP Scheduling Algorithm for Input-Queued Switches,” IEEE/ACM Tran. Networking, Vol.7, pp.188--201, April 1999.

[5] G. Chuanxiong, “SRR: an O(1) time-complexity packet scheduler for flows in multi-service packet networks,” IEEE/ACM Transactions on Networking, Vol.12, Issue 6, pp: 1144-1155, December 2004

[6] S. S. Kanhere, H. Sethu, et al, "Fair and efficient packet scheduling using elastic round robin," IEEE Trans. Parallel and Distributed Systems, Vol. 13, Issue. 3, pp. 324--336, Mar. 2002.

[7] L. Lenzini, E. Mingozzi, et al, “Tradeoffs between low complexity, low latency, and fairness with deficit

round-robin schedulers," IEEE/ACM Tran. on Networking, Vol.12, Issue 4,pp.681 - 693, August 2004

Design Considerations and Finite Element Analysis of Piezoresistive Cantilever for low flow sensing applications

Muhammad Akram Bhatti, Changxi Li
Department of Control Science and Engineering, Huazhong University of Science and Technology, Wuhan, 430074, Hubei Province, China.
Email: muhammadakramb@yahoo.com
Ahmed N. Abd Alla
College of Electric and Electronics Engineering, Huazhong University of Science and Technology
Email: najem200420032@yahoo.com

ABSTRACT

We present a design and finite element analysis for piezoresistive cantilever sensor. The mechanical analysis and design of piezoresistive cantilevers for flow sensing applications is studied. The change of relative resistivity of piezoresistive cantilevers has been analyzed for simple and paddle cantilever sensors. For a cantilever with specified spring constant and bandwidth requirements, the geometric dimensions of the cantilever were varied within certain constraints. The sensitivity of the sensing technique is studied. The calculated sensitivity (($\Delta R/ R$)/ y(0)) for simple cantilever sensor is 0.818-5.933x10^{-6}/nm and for cantilever paddle sensor is 0.86-6.75x10^{-6}/nm. These values fit well with the values predicted by finite element analysis (FEA) technique. These cantilever paddles have potential application in flow sensors.

Keywords: finite element analysis, piezoresistive cantilevers, low flow, sensitivity, stiffness and resonant frequency.

1. INTRODUTION

Piezoresistive detection is a popular technique when simple integrated sensors are required, and is often used for atomic force microscope (AFM) cantilevers instead of conventional optical detection. The convenience of a piezoresistive cantilever comes at the expense of resolution, thus they are primarily used in cases where laser detection is difficult. These include high-vacuum AFM [1], large arrays of cantilevers, ultra-small high-bandwidth cantilevers, and portable cantilever-based sensors [2].

The amount of bending of a cantilever beam can be detected by several read-out systems, including optical detection, capacitive detection, tunneling detection and interferometer detection. The optical level technique and the piezoresistive method are usually used to detect cantilever beam deflection. In general, the deflection is caused by its interaction with measurands under circumstances of stress, a small force, and a change of mass or temperature.

Micromachined silicon cantilever beams have been applied in fluid flow volume sensing [3, 4]. The piezoresistive cantilever is highly sensitive compared with any other cantilever probes. Due to their high spatial and temporal resolution, micromachined cantilevers can be used as versatile tools for interrogating micro/nano jets. For example liquid jets from micro scale nozzle can be examined with piezoresistive cantilevers

The piezoresistive cantilevers can measure deflection which can be correlated into thrust and velocities, due to momentum transfer and detect mass of liquid droplets when the liquid sticks to the cantilever surface in a resonant mode. When a piezoresistive material such as doped silicon is strained, its electrical conductivity is changed. The change in the resistivity can be most conveniently measured by using a Wheatstone bridge.

We first present the theoretical mechanical characteristics of silicon cantilever beams, such as the spring constant, deflection and relative change in piezoresistivity. An optimal structure is sought by using structural analysis and numerical finite element analysis. The sensitivity of piezoresistive cantilever sensor in this paper is addressed because of their simple geometry and quantifiable performance merits. The method and most of the results from this analysis should, however, be directly applicable to other types of piezoresistive sensors, including commercially available pressure sensors and accelerometers.

2. MECHANICS ANALYSIS

In this section, the simple and paddle cantilever sensors as shown in Fig. 1. and Fig, 2. were modeled using the static equations of mechanics. To calculate the amount of deflection at the tip of a cantilever beam, the differential equation of a cantilever beam for a small deflection is given by [5].

$$E I \frac{d^2 y(x)}{dx^2} = M \qquad (1)$$

Where M is the bending moment, E is Young's modulus, $y(x)$ is the deflection along the cantilever beam and I is the area moment of the cross section with respect to the neutral axis of the cantilever. $M = Px$ when a single force P is applied on the free end of the cantilever. $M = qx^2/2$ under a flowing fluid situation, where q is a force element at the position x along the cantilever beam and is proportional to the surface area facing towards the flowing fluid and drag force. The drag force is proportional to the fluid density, the drag coefficient of the cantilever, and the flow velocity squared in a turbulent flow or flow velocity in a laminar flow.

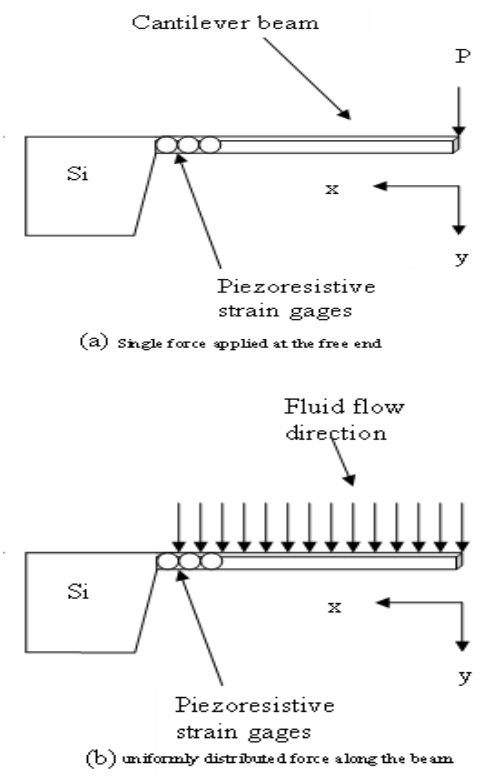


Fig. 1. Simple cantilever

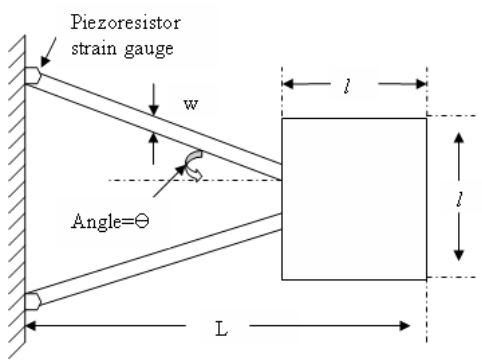


Fig. 2. Paddle cantilever

When the x-axis origin is selected at the free end of the cantilever beam, the boundary conditions are given by

$x=L$, then $dy/dx=0$

$x=L$, then $y=0$

Now we integrate the differential equation for cantilever deflection and use the above mentioned boundary conditions.

$$\frac{dy(x)}{dx} = \frac{1}{EI}[Px^2/2 + K_1] \tag{2}$$

Eventually, the deflection of the cantilever beam when a single force is applied at the free end of the cantilever is given as

$$y(x) = \frac{2P}{Ewt^3}[x^3 - 3L^2x + 2L^3] \tag{3}$$

Fracture Stresses

The stress in the beam is also an important factor. Stresses that are too high may cause problems like the beam to fracture or break. The stress is calculated as given below

$$\sigma = Ey\frac{d^2y(x)}{dx^2} \tag{4}$$

Where, y is equal to the distance from the neutral axis.

The maximum stress developed during bending the beam is always at the base of the cantilever beam. Therefore, piezoresistors are always implanted at the base of the cantilever beam to get maximum sensitivity.

The stress increases as the thickness of the beam gets smaller and also for an increase in beam length. However, increasing the beam thickness to reduce the stress will cause a reduction in the beam tip deflection. Before choosing any design criteria, the maximum or ultimate stress that the device can withstand must be found. The yield strength of silicon is $7\text{x}10^9$ N/m^2 [6]. Suppose that the beam withstand under this ultimate stress and the ultimate load to the cantilever beam can be found from [7].

$$\sigma_u = \frac{P_u}{A} \tag{5}$$

Spring constant or stiffness

The spring constant for cantilever beam has been defined as

$$K = \frac{3EI}{L^3} = \frac{Ewt^3}{4L^3} \tag{6}$$

Under the same load, a soft cantilever will deflect more than a stiff one, therefore has higher force sensitivity. From Eq.(6) one might be tempted to design a cantilever that is thin, narrow, and long. But long, slender cantilevers also tend to be very floppy with very low resonant frequency. The mechanical resonant frequency of a cantilever limits the bandwidth of its measurements and is another very important parameter that needs to be considered.

Resonant Frequency

The term resonant frequency refers to the first mode of natural frequency in free vibrations. The mechanical response of a sensor attenuates above its resonant frequency. Therefore the resonant frequency of a sensor often determines the maximum range of frequencies that can be detected. The resonant frequency of a device is a function of its stiffness and mass. For a homogenous cantilever with uniform rectangular cross section, the stiffness is defined by Eq. (6), and mass by

$$m = \rho(Lwt) \tag{7}$$

Suppose the mass, m, is distributed evenly along the beam, the fundamental resonant frequency of this cantilever is not simply $\sqrt{\frac{K}{m}}$ but we introduced a correction factor and resonant frequency is as

$$\omega_0 = \sqrt{\frac{K}{0.24m}} \tag{8}$$

A sensor with a high resonant frequency will be able to detect high-frequency signals and thus improves the measurement bandwidth.

3. COMPUTER MODELING

The cantilever beam was modeled using ANSYS software. A

finite element analysis software. The premise was to model a simple cantilever structure and verify that the two techniques, theoretical and computer provided the same answer. If this was true, then one could use the computer analysis to solve for more difficult structures that might pose a problem using theoretical analysis. Fig. 3. is screen snapshot of how *ANSYS* modeled the paddle cantilever beam. Modeling begins by choosing an element type, a beam, truss element, frame, solid, shell, etc. Solid 92 was chosen as the element type because it was a three-dimensional element and had the capability of six degrees-of-movement at each node point. Node locations were entered in Cartesian coordinates, and these were used to construct an area and then a volume. The volume was then meshed, which fills in node points throughout the entire volume for the finite element analysis. Before a solution is run, loading of the beam which includes both constraints and actual loads must be added. In Fig. 3., the left side two ends of paddle cantilever of the beam were completely restrained in all degrees of freedom. A gravitational force was applied and then the solution was implemented. Fig. 4. shows the structure in its deformed and undeformed shape after analysis. The maximum deflection of the beam occurred at the beam tip.

Using this ANSYS software the simple and paddle cantilever were modeled for a series of structures and found the sensitivities and also compared with theoretical calculated sensitivities, this will be covered in detail in analysis and results section.

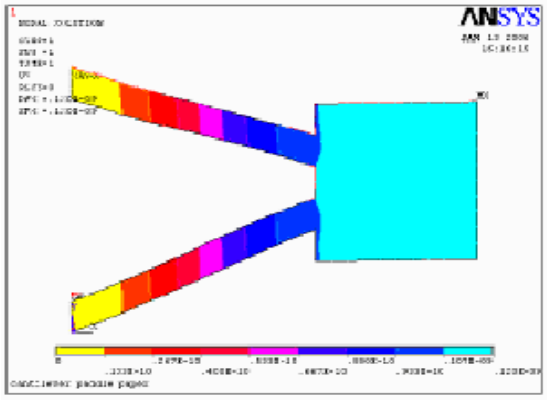


Fig. 3. Shows the structure in its deformed shape after analysis.

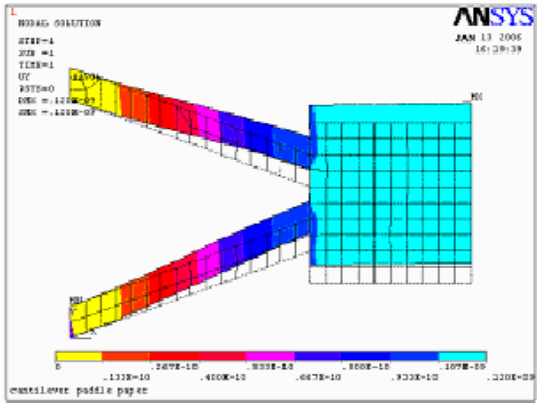


Fig. 4. Shows the structure in its deformed and undeformed shape after analysis.

4. ANALYSIS AND RESULTS

In piezoresistive cantilever sensors, the main device parameters are sensitivity of the device and maximum detectable measurand which is related to the maximum relative change of the resistance within the fracture strength of silicon.

Assuming that the deflection of the cantilever occurs only in the force direction, transverse shearing stress is zero. When a force acts on the free end in the direction of the symmetrical axis of the cantilever, on the basis of mechanics theory [8], the force constant of the simple and paddle cantilever are as below:

$$K_{sp} = \frac{E w t^3}{4 L^3} \tag{9}$$

Where ksp, is force constant for simple cantilever, E is the Young's modulus; L, w and t are the length, width and thickness of the simple cantilever respectively as shown in fig. 1.

$$K_{pd} = \frac{E w t^3}{2[L^3 - l^3 + 2wl^2]} \tag{10}$$

Where kpd, is force constant for paddle cantilever, E is the Young's modulus; *l*, L, w, t are the paddle size, length, width and thickness of the cantilever paddle respectively as shown in Fig. 2..

When the cantilever beam is pushed at the free end by a mechanical probe on a micrometer then the resistance of the piezoresistor on the cantilever changes under the strain. The relationship between the relative change of the piezoresistance and the deflection of the Simple cantilever at the free end is:

$$\frac{\Delta R}{R} = \beta \frac{3\pi E t L}{2L^3} y_1(0) \tag{11}$$

The relationship between the relative change of the piezoresistance and the deflection of the Paddle cantilever at the free end is

$$\frac{\Delta R}{R} = \beta \frac{3\pi E t L}{2[L^3 - l^3 + 2wl^2]} y(0) \tag{12}$$

Where R and ΔR are the resistance and it's the change under the strain, π is the longitudinal piezoresistive coefficient of silicon. β is a correction factor, which allows for the position of the resistors on the cantilever. The Young's modulus E and the longitudinal piezoresistive coefficient of silicon are taken as 1.5×10^{11} N m^{-2} and 45×10^{-11} m^2 N in our design.

Eq. (12) indicates that the relative change of resistance of the piezoresistors is proportional to the thickness and deflection of the cantilever, and approximately inversely proportional to the length squared of the cantilever.

The relative changes of the resistance versus the deflection of the cantilevers are shown in Fig. 5. and Fig. 6. for the designed cantilever paddles. The symbols are the calculated results and the solid lines are the fitted results by using the derived formulae. The slopes of the lines are the sensitivities of the cantilever devices.

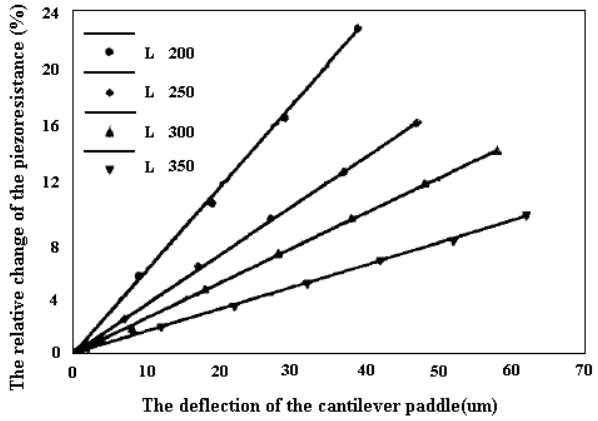


Fig. 5. The relative change of the resistance versus the deflection of the cantilever paddles—1 (*l= 100, w = 25, t =* 2.5μm).

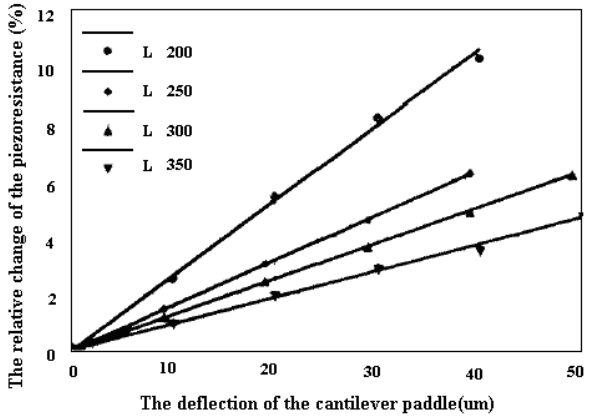


Fig. 6. The relative change of the resistance versus the deflection of the cantilever paddles—2 (l= *150, w = 30, t=* 2.5 μm).

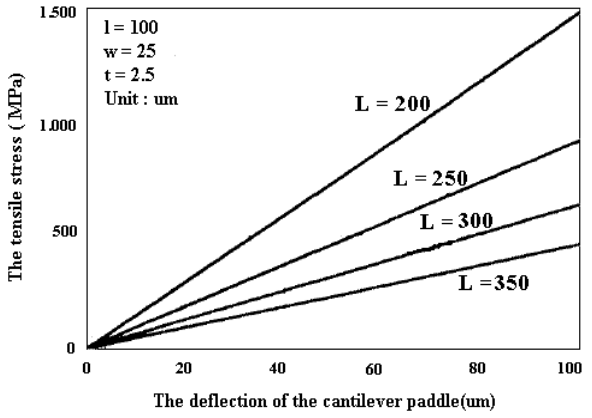


Fig. 7. The tensile stress versus the deflection of the cantilever paddles.

The sensitivities for a set of simple and paddle cantilever sensors have been calculated for theoretical and FEM techniques and tabulated in Table 1. and Table 2.. The tensile stress on the root of the cantilevers versus the free end deflection is shown in figure 7. The measured fracture strength of the silicon cantilever is about 1GPa, which is lower than others' results [9].

Based on the average fracture strength of silicon, the maximum relative change of the piezoresistance is calculated to be 44% when the correction factor is taken as 0.40; however, the maximum relative change of the piezoresistance is only 18.0% for the designed cantilever devices because of their lower fracture strength. This estimate is based on viscous drag theory.

TABLE 1

THE CHARACTERISTICS OF SIMPLE CANTILEVER AND DIMENSIONS (10^{-6}M)

L	w	t	K_{sp} (Nm^{-1})	S_{th} (10^{-6}/nm)	S_{FEM} (10^{-6}/nm)
200	25	2.5	1.831	5.933	5.822
250	25	2.5	0.938	3.848	3.721
300	25	2.5	0.5425	2.695	2.562
350	25	2.5	0.3417	1.993	1.926
300	30	2.5	0.6510	2.695	2.599
350	30	2.5	0.4100	1.993	1.901
400	30	2.5	0.2747	1.533	1.492
450	30	2.5	0.1929	1.215	1.197
400	40	2.5	0.3662	1.533	1.487
450	40	2.5	0.2572	1.215	1.197
500	40	2.5	0.1875	0.987	1.001

TABLE2

THE CHARACTERISTICS OF PADDLE CANTILEVER AND DIMENSIONS(10^{-6}m)

L	*l*	*w*	*t*	K_{pd} (Nm^{-1})	S_{th} (10^{-6}/nm)	S_{FEM} (10^{-6}/nm)
200	100	25	2.5	3.91	6.75	6.63
250	100	25	2.5	1.94	4.18	4.12
300	100	25	2.5	1.11	2.87	2.72
350	100	25	2.5	0.69	2.09	1.98
300	150	30	2.5	1.41	3.04	2.97
350	150	30	2.5	0.86	2.17	2.12
400	150	30	2.5	0.57	1.63	1.58
450	150	30	2.5	0.40	1.28	1.21
400	200	40	2.5	0.792	1.71	1.63
450	200	40	2.5	0.543	1.32	1.28
500	200	40	2.5	0.391	1.05	0.987
550	200	40	2.5	0.29	0.86	0.79

5. CONCLUSION

In this paper, a set of simple and paddle cantilever sensors were modeled using the mechanics equations of moments and finite element analysis techniques. These techniques were used to model a simple and paddle cantilever beam in order to verify the solution from both and be able to compare them. The calculated sensitivity ((ΔR/ R)/ y(0)) for simple cantilever sensor is 0.818-5.933x10-6/nm and for cantilever paddle sensor is 0.86-6.75x10-6/nm. These values fit well with the values predicted by finite element analysis (FEA) technique.

6. REFERENCES

[1] F. J. Giessibl, B. M. Trafas, "Piezoresistive cantilevers utilized for scanning tunneling and scanning force microscope in ultrahigh vacuum," Rev. Sci. Inst., 1994, 65(6). 1923–1929.

[2] D. R. Baselt, G. U. Lee, et al, "A high-sensitivity micromachined biosensor," Proc. IEEE, Apr. 1997, 85, 672–680.

[3] V. Gass, " Nanofluid handling by micro-flow sensor based on drag force measurement" Proc. MEMS 93 (New York: IEEE), 1993. 167–72.

[4] T. Nishimoto, S. Schoji, et al, "Buried piezoresistive sensors by means of Mev ion implantation" Sensors and Actuators A 43, 1994, 249–53.

[5] S. Timoshenko, "Strength of Materials, Part 1 Elementary Theory and Problems" 3rd edn (Princeton, NJ: Van Nostrand), 137–40.

[6] K.E. Petersen, "Silicon as a Mechanical Material," Proceedings of the IEEE, vol. 70, pp. 420457, 1982.

[7] F.P. Beer, J. E.R. Johnston, "Mechanics of Materials" 2nd edn New York: McGraw-Hill, Inc., 1981, pp. Chapter 8.

[8] Y. Kanda "A graphical representation of the piezoresistance coefficients in silicon" IEEE Trans. Electron Devices ED-29, 1982, 64–70.

[9] S. Johansson, J.A. Schweitz, "Fracture testing of silicon microelements in situ in a scanning electron microscope" J. Appl. Phys. 63,1982, 4799–803.

Configurable and Flexible System-on-Chip Design

Hua Hu, Luke Wang, Jianguo Xing
College of Computer and Information Engineering, Zhejiang Gongshang University
Hangzhou, Zhejiang, 310035, China
Email: huhua@mail.zjgsu.edu.cn, wangluke@126.com, jgxing@zjgsu.edu.cn

ABSTRACT

The development of EDA (Electronic-Design-Automation) technology promotes the level of SoC (System-on-Chip) design [5]. Here we introduce a new SoC design method. Unlike traditional SoC design approach, which is RTL-based and starts with hardware describe language, the new approach implements SoC based on a configurable and flexible processor that can be customized by designers from instructions extension to hardware block definition to reach product's performance requirement. Due to the configurable and flexible features, the new method provides SoC designers with great convenience and possibility to create a high performance, low power dissipation and small silicon area processor for application-specific SoCs or other embedded systems quickly.

Keywords: SoC, Configurable, Flexible, RTL-based, Processor-based.

1. INTRODUCTION

Compared with traditional SoCs, Modern SoCs designed for consumer products such as cell-phone, audio & video player, PDA and other digital products hold features of cheap, fast upgrade, low power consumption, short lifetime. Traditional SoC design method cannot provide SoC designer with ability to develop products quickly, effectively adapt frequently changing of requirement and control cost. Thus research on configurable and flexible SoC design method to satisfy the need of presently SoC development becomes more and more prevalent. This paper introduces a new SoC design method with features of configurable and flexible. The core of the new approach is a configurable ASIC processor of high performance or an application specific standard product (ASSP) combined with the flexibility of an FPGA. Configurable processors couple a microcontroller and memory with a block of programmable logic that can be used to create custom peripherals [9]. By this new technology, SoC designers can also define new instructions for the new customized processors to enhance its execution performance for served applications. By this way, we can design application specific processor, not general-purpose processor, only fit for our target system. Accompanied with the emergence of automatic development tool kit, the configurable and flexible SoC design method greatly shrinks development cycle, which is becoming more and more important for the products' success.

The work of this paper will be organized as follows. Section 2 overviews the traditional SoC design and points out the drawbacks of it. Section 3 describes a new method of SoC development and shows how it is more excellent than traditional one. Section 4 gives a case——Xtensa processor to present the new way of SoC design. Section 5 shows other companies' related work on configurable SoC and Section 6 summarizes this paper.

2. THE TRADITIONAL SoC DESIGN——RTL—BASED DESIGN

The typically traditional SoC design flow [1, 2], shown in Fig.1, reflects a historical separation between hardware and software development. The flow is composed of 4 stages.

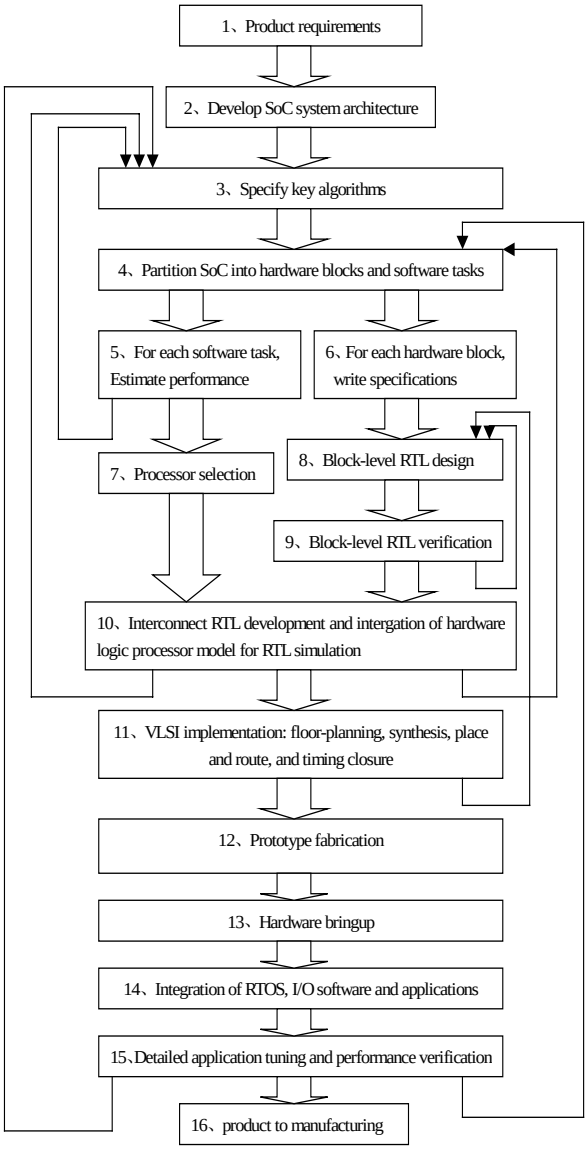


Fig. 1.

Stage 1 which is formed from Block 1 - 4 is called SoC concept. Designers in this stage can complete architecture design, arithmetic argumentation and partition system into

hardware blocks and software tasks.

Stage 2 that is formed from block 5 - 9 is called block-level design. In this stage, on one hand, the designers have to select a proper general-purpose processor for the target system based on the performance requirement by estimating each software task. On the other hand, the hardware designers begin to design RTL blocks following verification all of them.

Stage 3 which is formed from block 10 -13 is called hardware integration. The task of this stage is interconnection of RTL blocks, VLSI implementation and building prototype fabrication.

The last stage which is formed from block 14 -16 is called software implementation and integration. It includes integration of RTOS, I/O software and applications, detailed tuning and performance verification.

In the flow, many of the key architectural decisions are made quite early in the design process, long before any detailed performance or implementation feedback is available. The long delay from key algorithm selection and system-performance measurement on hardware sharply increases the risk of surprises and cost of fixing issues.
There are three dimensions for the flow worth nothing. The first one is architecture design. Architecture must always evolve to overcome limitations in the initial design. The evolving architecture must also accommodate changes from end-product requirements and exploit new system-design to improve performance, cost, reliability, and functionality. Software design is the second dimension. Often the software team's early participation in the design is focused on development and validation of only the most important algorithms or new software components. Limitations in model capacity and authenticity typically delay all hardware/software system-integration until hardware prototypes are available. The last dimension is hardware design. The pace of growth in both system complexity and silicon capacity demands a substantial increase in the reuse of existing hardware blocks, even though many hardware block are too rigid and specialized to allow wide reuse.

The necessity of getting hardware blocks and VLSI implementation absolutely correct makes the design and integration a long, iterative process. Designers must evolve block designs to allow the VLSI implementation to meet cycle-time targets; to fit gate-count and area goals; and to satisfy placement, routing, signal-integrity, power-distribution, and total power-dissipation requirements. Preliminary estimates from block diagrams, floor plans, and initial RTL code are typically crude and optimistic, so block design and block interfaces must change in response to simulation results from actual block layouts. Changes in the interface, size, timing, and power characteristics of one block will spread to other blocks, triggering even more design changes. The inevitable growth in complexity of system functions and silicon has particularly terrible consequences on the hardware design flow. Increasing block complexity means that the SoC designer faces a greater complexity also leads to more blocks, with more interfaces to document and test.

Growing chip Increasing complexity also has a cascade effect on VLSI design characteristics. More complex blocks require more gates and wires; higher gate and wire counts increase average wire length; longer wire means that more of the total clock cycle is spent in wire-delay; longer wire-delay means that placement and routing optimization increasingly influence overall circuit size, performance, and power dissipation.

3. THE CONFIGURABLE AND FLEXIBLE SOC DESIGN——PROCESSOR-BASED DESIGN

The new design-flow process is sketched in Fig.2. In this design, the block arrows show the basic SoC flow of refinement from high-level application requirement down to detailed hardware implementation. Within the main flow, there are four major subflows.

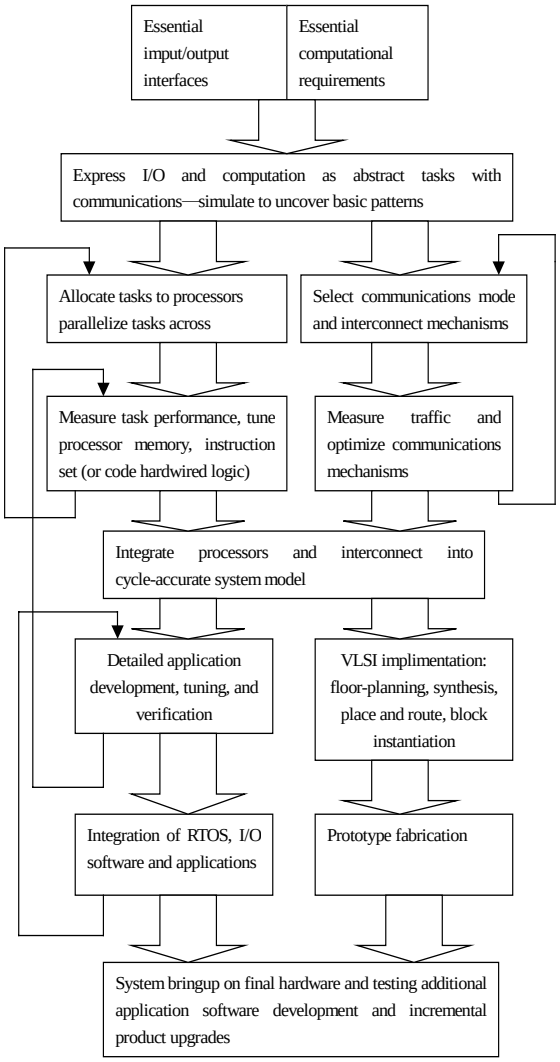


Fig. 2.

1. Computing optimization by processor configuration after code analysis.

2. Communicating optimization by selecting and tuning interconnection among blocks, especially processors.

3. Detailed software development, including operating systems and I/O code.

4. Detailed VLSI implementation, including logic design, verification, physical design, and silicon fabrication.
Some of the traditional conflicts between hardware and software are avoided because the team shares a common reference model. In particular, hardware prototype development is off the critical path for much of the software development. Final hardware/software integration should go smoothly because, in truth, hardware and software are integrated early in the processor-based approach to SoC design [6].

The new SoC design flow implies a number of

innovations relative to SoC design techniques commonly employed today. Compared with traditional method, the new SoC design concept possesses advantages as following:
1) Prefer simulation over hardware prototypes
While the complexity makes SoCs hardware prototyping expensive and time consuming. The new SoC design approach, with same effort invested in developing good system-simulation models, can result in virtual platforms that are available earlier, easier to debug, and require no hardware maintenance;
2) Defer hardware/software partitioning
The availability of application-specific processors increases the range of tasks that can be implemented in a software-based manner. Functions that once mandated early assumption of major RTL-based hardware design now can remain in software form, assuming extended processors as implementation vehicles. Of course, the design team may still decide to move particularly simple and efficiency-sensitive tasks into a hardware-only implementation, but the fraction of the design that follows this path is significantly reduced;
3) Configuration of communications
The greater interface flexibility of configurable processors creates new system architecture options. Instruction-mapped and memory-mapped data queues, low-overhead shared memories, and wide crossbar and bus interconnections all provide potential order-of-magnitude reductions in communication latency and order-of-magnitude bandwidth improvements relative to conventional 32-bit processor bus communications;
4) Anticipate design reuse [7]
Compared with a purely hardwired design approach, the advanced SoC design flow promises easier chip reuse. To build a new but related system, building a new chip with block reuse is unnecessary, reprogram the processor fabric in the existing chip is instead;
5) Develop hardware and software in parallel
The traditional model of SoC development forced most software development to wait until silicon prototypes were complete. The new SoC design methodology makes the entire development process more software-centric, allows the software team to start development sooner, and encourages tighter collaboration between the hardware and software teams.

4. TENSILICA'S XTENSA PROCESSOR — — A CASE OF CONFIGURABLE SOC DESIGN

Tensilica Company began to bring automatic, configurable and flexible processor technology in system-level SoC design in 1999. The core of this technology is Xtensa processor generator [10], which produce optimized processor core for embedded application based on served application. The designers can define new extended functions by themselves to content with special tasks requirement. By adopting this means, design team shall fast and flexibly create product satisfying requirement and specification. Tensilica's Xtensa LX processor [3] can provide high throughput and computation performance for users. Tensilica Company also develop XPRES compiler, which can automatically analyze any C, or C++ program and then by processor generator produce microprocessor to fit that program. Characteristic to implement configurable and flexible processor design with Xtensa LX as follow:

4.1 Instruction-set configurations

1) Operator fusion: This technique notes the frequent occurrence of simple –operation sequences in program loops. XPRES combines these operation sequences into one enhanced instruction, which accelerates code execution by cutting the number of instructions executed within the loop, making the loop run faster, as well as reducing the number of instructions that must be fetched from memory, thus decreasing bus traffic.
2) SIMD vectorization: Single Instruction Multiple Data. Some loops within application process an array of data in the same operation. XPRES compiler can vectorize these loops and build a multiple execution units to execute multiple data items in parallel.
3) FLIX: Flexible Length Instruction Extension. Unlike the operator fusion and SIMD, FLIX instructions consist of multiple independent operations. Tensilica's XCC compiler can pack these operations into a FLIX-format instruction as needed to accelerate code.

4.2 Multiport access conquer bus bottleneck[8]

In order to conquer bus bottleneck limitation, Tensilica Company adds two features to Xtensal LX processor. First, Xtensa LX processor has two load/store units, as Fig.3 shown. The two-load/store units possess of their own bus, which connect local memory with execution units of processor. With this additional unit, the system's throughput performance will be enhanced greatly. In actually, the second load/store unit's throughput will reach to 8-16 times than fixed instruction set architecture processors by bringing in 128-bit bandwidth bus. This new technique has been used in DSP system to satisfy requirement of high I/O bandwidth. However, some applications need higher I/O bandwidth performance that two load/store unit also cannot give, such as video coding and decoding. So Tensilica Company adds another feature to Xtensa LX processor, called TIE (Tensilica instruction extension) port and queue technique, which will avoid bus bottleneck limitation absolutely. SoC designers can define as many as 1024 ports connecting Xtensa LX execution units directly. The port bandwidth can also reach to 1024 bit. The systems with this new feature can exchanges information with Xtensa LX processor at a speed of 350,000Gbps. This performance will content with all processors I/O bandwidth requirements. The sketch map is shown in Fig.3.

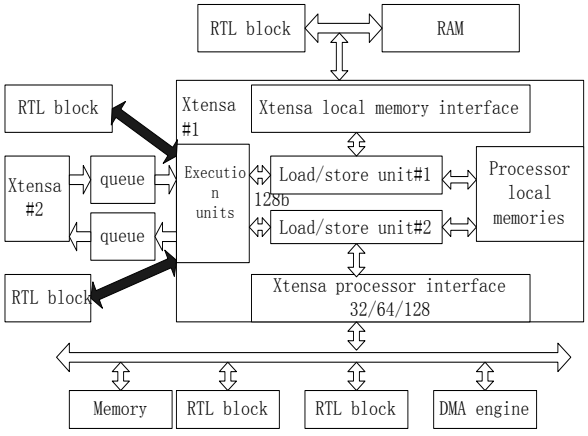


Fig. 3.

Tensilica Company's TIE (Tensilica instruction extension) language, which consists of Verilog language and C language extents, the processor's instruction sets. Previously mentioned TIE port and queue feature realization just uses this technique. This function can enable designers freely extent instruction sets in order to enhance the system

performance. The following fragment of TIE language describes an instruction that takes two sets of four 8-bit values packed into two 21-bit entries (a and b) in the AR register file, multiplies together the corresponding 8-bit values, and sums adjacent 16-bit multiplication results into a pair of 16-bit accumulated values. The result value(c) is written back into a third 21-bit AR register file entry:

```
Operation mac4*8 {in AR a, in AR b, out AR c} {} {
Assign
c={a[31:24]*b[31:24]+a[23:16]*b[23:16],a[15:8]*b[15:8]+a
[7:0]*b[7:0]};
}
```

From this instruction description, the following actions take place automatically:

1. New data-path elements, including four 8*8 multipliers and two 16-bit adders are added to the processor hardware description.
2. New decode logic is added to the processor to decode the new mac4*8 instruction using a previously unallocated operation encoding.
3. The integrated development environment, including instruction-set simulator, debugger, profiler, assembler, and compiler are extended to support this new mac4*8 instruction.
4. Plug-in extensions for third-party tools, including debuggers and RTOS are generated so these tools also include support for the new instruction.

   Then the software team can directly expose the new instruction in the high-level programming environment as following.

```
Int a, b, c;
C=mac4*8(a, b);
```

TIE language can also describe new registers, register files, and custom data types such as 24-bit data for audio applications, 56-bit data for security processing, 256-bit data types for packet processing.

Because extended processors employ firmware instead of RTL-defined hardware for their control algorithms, it's easier and faster to develop and verify processor-based task engines for many embedded SoC tasks than to develop and verify RTL-based hardware to perform the same tasks. Compared with other configurable processors, such as ARCtangent-A5 (ARC International) or Configurable ARM processors, Tensilica's Xtensa processors have much more configurable features than others. The variable instruction length, variable GPR (General-Purpose Register) and FPR (Float Point Register), variable address space and variable memory format (big/little endian) supply quite more configuration to users and make the target processor core fits for the target system better.

5. SOME OTHER RELATED WORK

5.1 ARC International ——user configurable cores[4]

In 1998 ARC invented a configurable microprocessor core that was licensed by Nintendo. In 1999 the firm introduced a complete IDE and soon followed with an RTOS and other software and middleware. The ARCtangent microprocessor is a 32-bit user-customizable core for ASIC, SOC, ASSP and FPGA development. Developers can modify and extend the instruction set for specific applications to optimize performance, I/O throughput, power consumption, silicon area and cost. Designers can add DSP functionality and merge RISC and DSP functions onto a single processor, thereby saving even more silicon area and power consumption. Due to multiple CPU I/O interfaces and a low gate count, the ARCtangent processor lends itself to multiprocessor designs. The processor is supported by development tools including a configuration tool, which has a graphical "point and click" user interface. The tool has a range of options to build HDL, synthesis scripts, test bench and HTML documents. The MetaDeveloper tool suite includes a C/C++ compiler, assembler, linker, profiler and debugger. This tool chain fully supports the capabilities of the ARCtangent processor including multiprocessor debugging and extensibility for processor customization. ARC also provides a royalty free real time operating system. Although ARC's processor core can not be so configurable and flexible as Tensilica's, it has own attraction——ARC XY advanced DSP. ARC XY Advanced DSP adds the power of a true DSP engine to ARC CPU cores, enabling conventional and signal processing computation within a single unified architecture. ARC XY Advanced DSP's separate memory banks for X and Y operands deliver data at register speed, eliminating main memory fetch cycles. In addition, ARC XY Advanced DSP's address generators eliminate the CPU cycles to determine the address of the performance of a dedicated DSP engine. ARC also offers DSPlib, a library of frequently used signal processing functions that have been verified and optimized for ARC XY Advanced DSP. The library takes full advantage of ARC's extendible instruction architecture to maximize the performance of each function. ARC cores with ARC XY Advanced DSP provide a complete solution for many complex computation problems in SoCs that are targeted at communications, media processing and many other applications.

5.2 Elixent —— configurable processing array

Company Elixent has been delicated in developing the concept of a Reconfigurable Algorithm Processing (RAP) platform. The basis for company's technology is the D-Fabrix processing array. The components of the array are 4-bit ALUs, registers, and the 'switchbox', which are incorporated into a 'tile'. Hundreds or thousands of tiles are then combined to create the fine-grained D-Fabrix array. Special functions are then distributed through the array and algorithms can be implemented in the hardware. The result is a solution that combines the performance, power and area overhead benefits of hardware with the flexibility of a software configuration.

RAP, usually not used alone, is always integrated with configurable processor core. It is a powerful approach to implementing algorithms that need high arithmetic throughput and low cost. Elixent's D-Fabrix RAP platform implements algorithms in "Virtual Hardware", allowing the creation of a hardware accelerator for every algorithm in a system. By virtue of reconfigurability, it can implement multiple hardware accelerators in the same silicon area, giving high silicon utilization. Further, this reconfigurability allows functionality to be added or changed post-fabrication, allowing bugs to be fixed, new functions to be added, or even the whole chip to be customized. RAP is similar with Tensilica's firmware. But it has more powerful and more flexible ability than the latter.

5.3 PicoChip Designs Limited —— reconfigurable signal processors picoChip offers a scaleable, multi-processor baseband IC that combines the computational density of a dedicated ASIC with the programmability of a traditional high end Digital Signal Processor along with a rich programming environment and comprehensive system libraries. The picoArray itself is a massively parallel array of

individual processors linked by a deterministic high speed interconnects fabric of 32-bit buses, with about 400 cores on a single die, well described as “Software System On Chip” (SSoC). Each of these is a capable 16-bit device with local data and program memory, roughly equivalent to an ARM9 for control tasks or a TI C5x for DSP roles. Because each of these cores can operate in parallel or in concert, and because of the huge bandwidth of the on-chip buses, the Pico Array can deliver a huge amount of processing power. Multiple array elements can be programmed together as a group to perform particular functions ranging from fast processing such as filters and correlators, through to the most complex control tasks. Each element is allocated a series of simple tasks to avoid problems of statistical multiplexing of resources or run-time scheduling and so performance is entirely deterministic, simplifying development and verification.

Unlike other MSoCs (Multiprocessore Systems-on-Chip), pico Array only use VHDL to describe the inter-processor relationships. There is no actual VHDL programming and no need for VHDL simulation; only the structural elements are used to define the relationship between elements. This approach allows the algorithms to be efficiently partitioned and mapped onto specific processing elements at a relatively high level.

5.4 IPFlex, Inc. —— dynamically reconfigurable processors IPFlex, a Japanese corporation, was founded in March 2000 as a fabless semiconductor company focused on developing dynamically reconfigurable processors and its integrated development software. Its main product is DAP/DNA-2 (Digital Application Processor /Distributed Network Architecture). DAP/DNA dynamically reconfigurable processor is designed as a dual-core processor comprised of a high-performance RISC core (DAP) and a dynamic reconfigurable processor core (DNA). And it is a platform that provides hardware performance while maintaining software flexibility. The DAP/DNA dynamic reconfigurable processor series is provided with the DAP/DNA-FW II as the integrated software development environment. It provides compilers for algorithms written in MATLAB/Simulink and C with data flow extension, thus realizing high-abstraction level algorithm design as well as leveraging existing intellectual properties of users. The DAP/DNA-2 is a microprocessor that contains multiple processing elements (PEs) and can optimally configure internal circuits to best suit the application in demand.

DAP/DNA distinguishes itself by its dynamically reconfigurable feature. The function of each PE, as well as connections among PEs, can be reconfigured not only when building the system, but also when it is running, enabling instant (within one clock cycle) reconfiguration to suit the application at hand. The DAP/DNA-2 lays out this PEs in a two-dimensional array so that it can quickly and flexibly change their function and the connections between them. This distinguishing feature will become prevalent technology in future SoC design, which makes better use of hardware and gives more intelligence to target system.

At last, I would like to say the difficult that configurable SoC is facing. Many semiconductor companies are dedicated in configurable SoC hardware development now. But few of them research in how can efficiently program on these devices. So we will often meet with such problem that software cannot make full use of hardware. In the future, the research on this aspect should be paid more emphasis.

6. SUMMARY

RTL-based SoC design is appropriate for the small, simple and stable system. However, with the development of IC design, large-scale, complexity and fast upgrade are becoming the topics of SoC design. Traditional SoC design (RTL-based) method cannot be qualified for this trend because of itself disadvantages, such as difficult design, slow verification, high cost and poor scalability for complex problems. Configurable SoC and flexible design method overcomes the limitation of traditional method. The features —configuration and flexibility—enable SoC designers create an individualized SoC quickly and easily with relatively little manual effort. Realizing the strongpoint of configurable SoC design approach, more and more semiconductor companies launch out into related products. In this situation, the new approach is becoming mainsteam in the VLSI design field.

7. REFERENCE

[1] C. Rowen, Engineering the complex SoC: Fast, Flexible Design with Configurable Processors, China machine press. September 2005.

[2] R. Rajsuman, System-on-a-Chip: Design and Test, Buaapress, August 2003.

[3] S. Leibson, J. Kim, Configurable Processors: A New Era in Chip Design, IEEE Computer Society, 2005 July, pp.51-59.

[4] J. Horgan, “ International system-on-chip conference”, EDA Weekly, May 03, 2004.

[5] S. Wallner, “A congurable system-on-chip architecture for embedded and real-time applications”: concepts, design and realization, Journal of Systems Architecture 51 (2005) pp.350–367.

[6] R. E. Gonzalez, Xtensa: “A configurable and extensible processor”. IEEE Micro, 20(2): March-April 2000, pp. 60--70.

[7] M. Margala, H.F Wang, “New approach to design for reusability of arithmetic cores in systems-on-chip”, INTEGRATION, the VLSI journal 38 (2004) .185–203.

[8] A. Wang, E. Killian, et al, "Hardware/Software Instruction Set Configurability for System-on-Chip Processors", Proceedings of the Design Automation Conference, June 2001,184-190.

[9] E. Ricardo, Gonzalez, “Configurable and Extensible Processors Change System Design”.

[10] Tensilica, Xtensa 6 Processor Core Provides Fastest Customization Path, Lower Power and Advanced Security Provisions, Available from www.tensilica.com/news_events/pr_2005_10_24.htm.

Research on High-performance Memory System Design

Qingbao Li , Ping Zhang , Rongcai Zhao , Xuerong Gong
Department of Computer Science and Technology, University of Information Engineering
Zhengzhou, Henan 450002, CHINA
Email: lqb215@vip.371.net

ABSTRACT

The rapidly development of the semiconductor technology have been greatly affecting the system configuration of computer system. The continuous growing gap between processor and memory speed has become an important drawback in the overall computer performance. The reason caused "Processor-Memory Performance Gap" is first discussed from the point of the different semiconductor technology of processor and memory. After analyzing the locality principles of application programs, the design of memory hierarchy structure and memory interleaving is discussed which are two techniques based on principles of locality to improve memory performance. Then two novel techniques are presented. Multi-windows cache is used to improve cache hit ration in multi-core or hyper-thread processors. Intelligent memory system is used to increase bandwidth by integrating computational power with memory. All these techniques are effective to improve memory performance and attenuate the disparity between processor and memory.

Keywords: Principles of Locality, Memory Hierarchy Structure, Memory Interleaving, Multi-window Cache, Intelligent Memory.

1. INTRODUCTION

VLSI technologies have been increasing the speed and integration of processor. The work frequency presents the processor speed while compute speed presents processor's ability of computing. Although speed is a very important fact to improve computer's performance, High speed of processor only can't bring good performance of whole computer system. The improvement rate in processor speed by far exceeds the one in memory by an improvement rate of 60%/year in processor performance, while the access time to main has been improving at less than 10%/year [7, 10], as shown in Fig.1. These have been introducing a larger, growing performance gap between processors and main memory. This increasing processor-memory performance gap is now the primary obstacle to improved computer system performance and it will increase in the next few years.

A number of reasons are in the genesis of this growing disparity of processor and memory, but the prime reason is the division of the semiconductor industry into processor and memory fields [8]. As a consequence their technologies head in different directions: the first one has increased in speed, while the latter has increased in capacity and this lead to the fact that different semiconductor technologies are used in their products. The latter technology is more difficult, it is not easy to improve the clock speed of memory and the control logic of processors is not so regularity as the one of memory. In order to keep stability of memory chips, the semiconductor technology trends to maximize the area of capacitance, and use multi-layer silicon, but not use metal-layer. While the technology of processor logic trends to use more metal-layers and optimize the transistors switch speed. So the improvement rate in processor speed by far exceeds the one in memory.
Computer pioneers have correctly predicted that programmers would want unlimited amounts of memory, so the memory manufactories will not change their direction of pursue larger capacity and it is hard to improve memory speed to map processor speed by improve technology of memory manufactory [6]. Thus the way to solve the problem is to improve the design and implement of memory system.

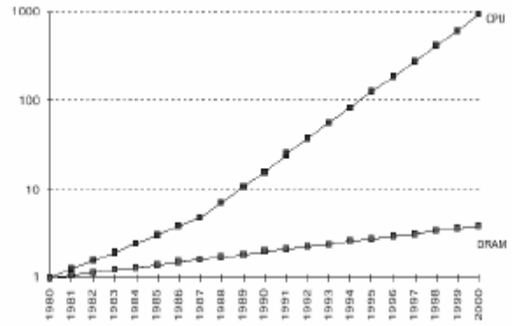


Fig.1 Processor-memory performance gap: starting in the 1980 performance, the processor and memory performance over the years

The remainder of this paper is structured as follows: Section 2 discusses the principles of the locality when instructions and data access memory which is the basic of high performance memory system design. Section 3 introduces techniques having been used to improve system performance. Section 4 presents several new ideas in memory system design. Section 5 concludes this paper.

2. PRINCIPLES OF LOCALITY

In order to improve the total system performance, we must study the architecture of computer system, analysis the mechanism of program control flow and emphasis on the locality of time and space in system design.

Due to program accesses in program control mechanism computer tend to be clustered in certain regions. Before delving into the depths of a memory system, it is important to acquaint ourselves with the basic underlying principles. These

are called the *Principles of Locality* [1]. There are two types of locality:

- Temporal Locality - This law states that recently referenced instructions or data are likely to be referenced again in the near future. Such as iterative loops, once a small code segment is entered, will be accessed repeatedly many times.
- Spatial Locality - This law emphasizes on the fact that if an item is referenced, items with addresses close by will tend to be accessed in the near future. For example operations on tables or arrays involve accesses of a certain clustered area in the address space.

The locality in control flow computer provides us the basis to solve the problem of "Processor-Memory Performance Gap". We can work on memory system architecture to attenuate the disparity between processor and memory.

3. TODAY'S TECHNIQUES TO IMPROVE MEMORY PERFORMANCE

Based on the principles of locality, many researches have been done to attenuate the disparity between processor and memory. Two primary measures are discussed in this section: memory hierarchy structure and memory interleaving.

Memory Hierarchy

Based on the characteristic of control flow of program, Cache is inserted between the processor and the memory to solve the problem of speed gap. Cache is a small amount of faster and expensive memory [9]. The processor, caches and main memory are organized as a hierarchy as depicted in Fig.2.

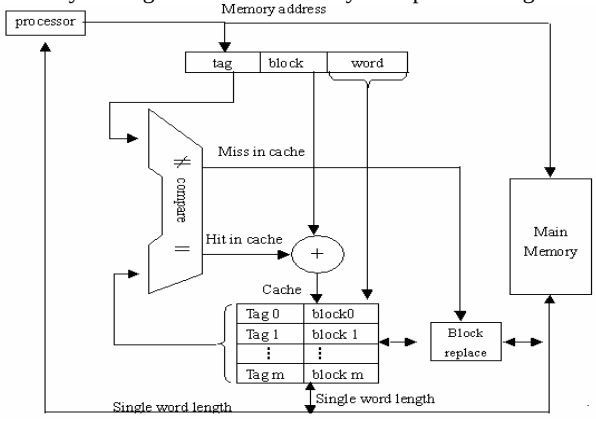


Fig.2 Memory Hierarchy

Most multiprocessor systems use private caches. When the processors access main memory, the addressed data/instruction is first checked in the cache. A cache hit occurs when the data/instruction is found in the cache. The processor may directly read/write data from/to the cache quickly. Otherwise, a cache miss occurs, the processor must read data/instruction from main memory, and the cache is loaded with a block of addressed data/instruction from the memory.

Obviously, the architecture can improve the efficiency of memory system [4]. A way to analyze the performance of a cache-memory hierarchy is through the average memory access time(T_{avg}), using the following expression:

T_{avg} = hit time + miss rate $\times$ miss penalty .

Assume the cache access time is T_c, the main memory access time is T_m, $T_c << T_m$, The hit ratio at cache is H, the miss ratio at cache is 1-H , then the average access time of processors access this cache-memory hierarchy is expressed as follow:

$T_{avg} = T_c \times H + T_m \times (1-H)$

When H close to 1 $T_{avg} \approx T_c$, The cache hit ratio is affected by the cache size and the block size. Programs addressed flow and block replacement policies also affect the cache hit ratio. In a word the cache-memory hierarchy can enormously improve speed of processor accessing data/instruction.

Memory Interleaving

By using memory interleaving,, memory adjacent cells can be overlapped in a pipeline fashion. As a result, multi-way interleaving can increase memory bandwidth effectively.

The main memory is built with multiple modules. These memory modules are connected to a system bus or a switching network. Once presented with a memory address, each memory module returns one word per cycle. It is possible to present different addresses to different memory modules so that parallel access multiple words can be done simultaneously or in a pipelined fashion [1].

Based on the local characteristic of control flow computer, the low-order interleaving can store the contiguous memory words into M memory modules. The low-order interleaving is shown in Fig. 3.

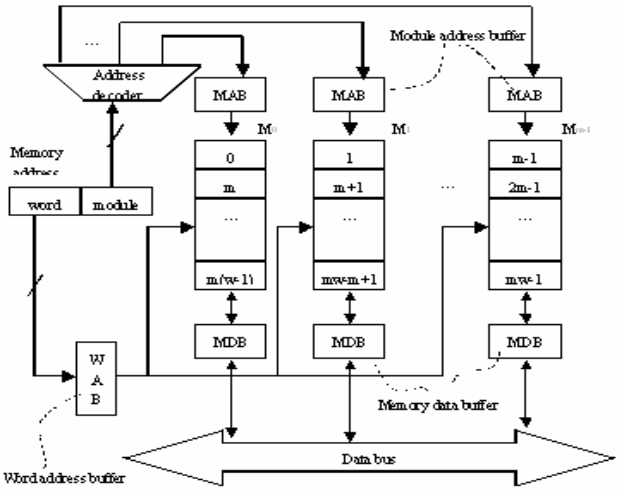


Fig.3 Memory interleaving

In the low-order interleaving, low-order bits are used to identify module, high-order bits present word addresses within each module. A word address is sent synchronously to present all memory modules. The m-way memory modules return with one word to private data buffer per cycle. Hence the low-order interleaving implements block data access in a pipeline fashion. Let θ be the memory access cycle, namely the m words can be accomplished in a θ cycle. the access time of each word is τ cycle, these two cycle times are related as follows:

$$\tau = \theta / m$$

If other factors, such as memory-access conflicts, are ignored, the effective memory bandwidth is m times to the bandwidth of a single module.

4. CHALLENGES AND SOLUTIONS

Now processor manufacturers are trying to improve processor performance by using new techniques, such as multi-kernels, hyper-threading and increasing cache size in a processor chip. All these techniques will increase the gap between processor and memory, but we still can not exploit the full potential of these processors. The primary reason lies in memory, from long term views, the memory manufacturers can not manufacture higher performance/cost memory products satisfying the demand of processor. Thus, these new techniques bring new challenges for us to improve memory system performance and fully exploit processor potential. In this section, we first present multi-windows cache technique which is used to improve cache hit ration in multi-core or hyper-thread processors, then we introduce intelligent memory system [5] which increase bandwidth by integrating computational power with memory.

Multi-window CACHE

The trend of the development of future processor is multi-core and hyper-thread. Integration of multiple processors onto the same chip offers a number of potential advantages: the inter-processor communication latency and bandwidth can be greatly improved. But when multiple threads executing simultaneously, each thread may be defined as a stream of addresses associated with the instructions and data of particular sequence of code that has been scheduled within processor. Multiple threads may cause the primary cache to be largely overrun by the data lines associated with a particular thread. Such a condition may be referred as cache pollution and may slower the execution of threads other than a particular thread.

In order to to avoid such cache pollution and provide a balance between the throughput obtained by the different threads in multi-core/hyper-thread processors. We present a new method: multi-window cache, as shown in Fig.4. The cache is dynamically partitioned into multiple windows according to the number of threads, each thread corresponds to a window and the cache windows are switched through switches. Thus each thread load a block of instruction or data into its own cache-window, not destroy the other thread's caches, in other words it can keep each thread locality.

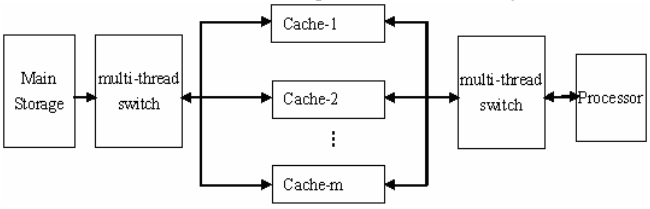


Fig.4 Multi-windows cache

The cache memory is partitioned among a set of threads. When a cache miss occurs, a replacement line is selected in a partition of the cache memory which is allocated to the particular thread from which the access causing the cache miss originated, thereby preventing pollution to partitions belonging to other threads. The partition may be static. Alternatively, the partition may be dynamic and may be used to control relative throughput associated with the threads.

The cache management method comprises steps of:

- Partitioning the cache memory into a set of separate cache windows and allocating a separate subset of the partitions to each of a set of threads.
- Performing a cache lookup operation to the cache memory in response to an access to a data line by the processor.
- If a cache miss occurs, then selecting a replacement line in one of the window of the cache memory which is allocated to the particular one of the threads from which the access originated and do cache fill operation with new data line.

Intelligent Memory

The contemporary model of a processor-centric architecture will allow performance to scale acceptably as long as two conditions hold: first, that the processing core has sufficient work to do, to mitigate cache miss latencies [2]; second, that the processor has enough bandwidth to load changes to the cache set without excessive delay. Both these conditions have already been stated, and are getting more and more difficult to meet. So the idea of memory-centric architecture become more attractive and the subjacent technology has been considered promising enough, as the number of transistors per chip (that exponentially grows) begins to be sufficient to the implementation. Memory-centric architecture fuses the storage component with the processing core in a single chip, creating memories with processing capacity. Such memories is called intelligent memories [3] and the idea is based on the observation made by D. Elliot, which stated that memory chips have huge internal bandwidth [7]. As he pointed out, the connection pins are responsible for the external degradation of the bandwidth, thousands of times slower than internal bandwidth. Eliminating the connections not only improves the bandwidth, but also improves the latency, as logic and storage are closer to each other.

Intelligent memory systems improve memory system performance by integrating computational power with memory [10]. This design strategy can explore the massive on-chip bandwidth of the memory, as the computation elements are integrated directly into the memory chip. Architectonically, this is a simple approach and theoretically capable to achieve the high performance. The intelligent memory naturally has extraordinary internal bandwidth. An on-chip processor can tap that bandwidth, reduce the latency of access memory five to ten times and improve memory bandwidth fifty to hundred times.

5. CONCLUSION

Until few years ago almost all the research and production effort (divided into microprocessor and memory fields) was respectively channeled to speed and to capacity. The result is a continuous growing gap between the processors and memoirs speeds. This paper discusses several techniques to improve the performance of memory system to lessen this disparity from the point of memory system design. All these techniques are effective and are likely to shape the computer industry in the future.

6. REFERENCES

[1] H. Kai, Advanced Computer Architecture. New York:McGraw-Hill book Co,1993.
[2] K. Diana, O. Mark, et al, "Cache Coherence in intelligent Memory Systems". IEEE Transactions on Computers, 52(7):960-966,July 2003
[3] D. Patterson, T. Andersson, et al, "A Case for Intelligent RAM": IRAM.IEEE Micro,April 1997.
[4] J.L. Hennessy, D.A. Patterson, Computer Architecture: a Quantitative Approach, third ed., Morgan Kaufmann Publishers, UK, 2002.
[5] M. Oskin, F. Chong, et al, "Active Pages: A Computation Model for Intelligent Memory". Proceedings of the 25th Annual.
[6] D. Pattasson, T. Anderson, et al, "A Case for Intelligent Ram": IRAM Computer Science Division/EECS Department university of California Berkeley CA 94720-l776 2/8/97
[7] D. Elliott, "Computational Ram: A Memory-SIMD Hybrid and its Application to DSP". In Proceedingsof the Custom Integrated Circuits Conference, Boston, MA, 3 1992.
[8] D. Burger, J. R. Goodman, et al, "Memory Bandwidth Limitations of Future Microprocessors". In Proceedings of the 23rd Annual International Symposium on Computer Architecture, 1996 78–89.
[9] N. Dinis, "Cache why Level it?" In Proceedings of the 3rd Internal Conference on Computer Architecture, Universidade do Minho, 2002
[10] T. Justin, "Performance Characteristics of an Intelligent Memory System", Master's thesis University of Maryland. August 2004.

Energy Efficiency Based Optimal Task Allocation for Cluster-Based Target Tracing

Xihuang Zhang, Chen Jin
School of Information Engineering, Southern Yangtze University
214122 Wuxi, Jiangsu, P.R.China
Email: zxhsytu@sytu.edu.cn, chenjing_jing@163.com

ABSTRACT

For cluster-based target tracing in wireless sensor networks (WSN), tracking by only one cluster is not effective. Tracking tasks are collaborative, dynamic and distributed tasks and are executed in several gateways. In order to maintain a certain degree of service quality and a reasonable system lifetime, energy needs to be optimized at every stage of system operation. Sensor node clustering is another very important optimization problem. Nodes that are clustered together will easily be able to communicate with each other. Considering energy as an optimization parameter with multi-cluster WSN, the model of target tracking and the mode of energy consumption have been presented. For energy optimization, an optimal genetic algorithm for task allocation among these gateways is studied which balance the energy consumption among concerning gateways. Experiments are designed and simulation results show that energy savings can be obtained with the proposed algorithm.

Keywords: Target Tracing, Wireless Sensor Networks, Optimal Task Allocation, And Genetic Algorithm

1. INTRODUCTION

Recently, the wireless sensor networks (WSN) have gained increasing importance due to their potential for some civil and military applications. Each sensor node distributed in an area is capable of detecting surrounding conditions such as temperature, sound, or the presence of mentioned objects. One of the most promising applications of WSN is the target tracking. [1,2]. During target tracking with WSN, energy recovery, energy efficiency and real time are to be considered. Many useful protocols among all WSN layers are researched. The cluster-based target tracking is one of them.

During a cluster there is a node called gateway node that is used to collect all sensing data from all sensor nodes belong to the cluster. The sensor periodically sends such sensed data, via radio transmitter, to gateway node directly. The gateway can process the sensed data as well as communicate with sensor nodes and base station. In addition the gateway can be as a router for communication among nodes belong to different clusters and execute the complex task in distributed way. In target tracking, sensor data indicates detection of a target while fusion of multiple sensor reports can be used for tracking and identifying the detected target. Using gateways of clusters to transmit processed data to base stations, which could minimize the number of sensor nodes that take part in long distance communication. This directly affects the overall system energy consumption.

While most of the previous research focused on the optimal use of sensor's energy, very little attention has been paid to the efficiency of energy usage at the gateway [3]. Tasks need to be allocated to gateways in such a way that maximizes the life of these cluster-heads and eventually the whole network. How to allocate tasks to gateways optimally or near optimally is the key of the problem. It is well known that task allocation problem is NP-complexes, which means that it is hard to solve it by searching all solving possibilities. Genetic algorithm (GA) is a methodology within the artificial intelligence area [4]. It is a set of ransomed methods for optimization an objective function by breeding a population of possible solutions. In nature the evolution may treat many problems. To treat the big solving space problem, Genetic algorithm is a useful tool.

In this paper, a multi-clustering is detailed and the energy model is established. For real-time and energy metric, we need to allocate tasks to these gateway nodes. The optimal genetic algorithm is used for task allocation. Objective for task allocation are proposed and simulations are given.

2. SENSOR NETWORKARCHITECTURE

The wireless sensor network architecture for target tracking is shown in Fig. 1. It is cluster based and sensor nodes are grouped into clusters controlled by base stations. Sensors are responsible for sensing the area where is covered by the target sensing area. Every cluster has a gateway node that manages sensors in the cluster. Clusters can be formed based on many criteria such as communication range, number and type of sensors and geographical location [5,6].

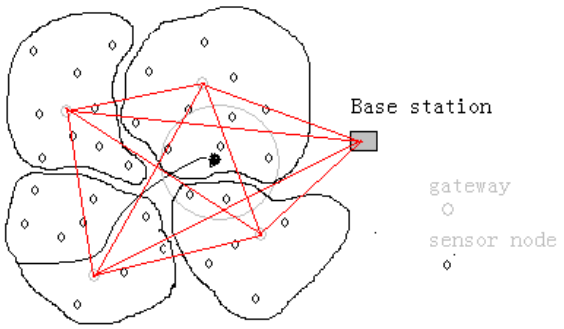


Fig. 1. The wireless sensor network architecture for target tracking

It is assumed that sensor and gateway nodes are stationary and the gateway node is located within the communication range of all the sensors of its cluster. The base station performs clustering the sensor network. Every gateway has its address and location. Gateways can track targets using the data from sensors in any clusters as deemed by the base station. However, sensors that belong to a particular cluster are only accessible via the gateway of that cluster. Therefore, a gateway should be able to route sensor data to other

gateways.

Heavy use of a gateway node can consume most of the node power and thus significantly shorten the life of the gateway, which allows for efficient and optimized use of the gateway nodes. Optimized task allocation among gateways, can expand the life of a certain gateway node by reducing the load (number of tasks) on that node.

2. ENERGY CONSUMPTION MODEL

For tracking applications, N wireless sensors are randomly distributed in an L*L square area with the density $\rho = N/L^2$. These nodes are clustered into k clusters, $\{M_1, M2 \ldots MK\}$. Every cluster has its own gateway G_i to communicate to base station with R_i distance. Every gateway G_i can communicate to its member node in its cluster in a circle area with radius r_i. After initialization, all nodes know their positions and the architecture of WSN is stable and nodes are immoveable.

There are two operation of sensor node energy consumption for tracking a target: sensing, transmission and receiving for gateway node. Otherwise, for getaways there are many operations to energy consumption: transmission and receiving with nodes in cluster, transmission and receiving with base station, data processing and routing. Usually a moving target has its own sensing area noted by a circle area with radius Tr. The area may covers several clusters. That is to say these clusters should track the target cooperatively. Assume n_i sensor nodes in cluster M_i sense the target, note as N_i. The energy consumption model is expressed as following:

$$E = \sum_{M_i}\Big(\sum_{N_i}(k_{si}E_{sen} + 2k_{ri}E_{rece} + 2k_{ti}E_{trans}) + \sum_{N_i}(K_{ri}E_{rece} + K_{ti}E_{trans} + K_{pi}E_{proc} + K_{ri}E_{routing})\Big) \quad (1)$$

Where E. is the energy consumption corresponding operating, such as E_{sen} is sensing energy consumption. k_i expresses the times of operating for sensor nodes. K_i expresses the times of operating for gateway nodes.

3. TASK ALLOCATION TO GATEWAYS

3.1 Task Allocation

During some applications such as target tracking, a gateway node should communicate a set of sensors that belong to another cluster. So there are too much tasks executed in gateway nodes. It can consume most of the gateway node power and thus significantly shorten the life of the gateway. Optimized task allocation among gateways, can expand the life of a certain gateway node by balance the load (number of tasks) on gateway.

To implement the task allocation among gateway nodes is generally constrained by the accessibility of sensors. If a sensor belongs to a cluster mastered by a gateway M, the sensor has to be reached via the gateway M. In additions, task allocation also should meet the timing constraints of applications. But task allocation and task scheduling are NP-Hard problems. To reduce the complexity of above problems, it is assumed that the task of data processing is in the period after the data is sensed, which omit the dependency between data sensing and data processing and thus ensure data availability when scheduling the data processing task.

Target tracking is a cooperation application job, which consists of multiple tasks that execute data collection, communication, data processing routing and so on. The goal of task allocation optimization is to maximally extend the life of all gateways by balancing the load proportional to the energy that each gateway has. By the feature of wireless sensor network, there is no task migration, which means a task would not be reallocated during a new allocation.

3.2 Objectives

To establish the objective function, some definitions are detailed as following. Let N be the Number of gateway nodes and n sensor nodes in sensor network. A task $T \in \Phi$ has *one* process and n communications to n sensor nodes. $\Phi = \{T_1, T2 \ldots t_m\}$ is the set of tasks to be allocated to N gateways. If a set of tasks were allocated to gateway i, note φ_i. $E_i(t)$ express the remaining energy of gateway i at time t, and $rE_i(t)$ Rate of energy consumption per cycle of gateway node i at time t, which is the sum of all energy consumption for computation and communication of all assigned task to gateway node i.

After allocating, all tasks should be successfully allocated and no task is assigned to two or above two gateways. If the relationship between power consumption and both CPU processing and communication is known as above detailed, the power consumption rate for task can be estimated, here noted as $rET_i(t)$ and $rEC_i(t)$ respectively. So after allocation, time $t+\delta t$ after t, $rE_i(t+\delta t)$ is:

$$rE_i(t+\delta) = rE_i(t) + rET_i(t) + rEC_i(t) - rE_i(t-T) \quad (2)$$

Here assume that a task finish T cycle after it be allocated.

The life of the gateway at time t is calculated by following formula:

$$Life_i(t) = E_i(t)/rE_i(t) \quad (3)$$

There are some constrains in task allocation optimization. Enough gateway processing capacity and communication bandwidth ensure each task is finished before deadline. To reduce the problem in task allocation, it is acceptable that the max number N_{max} of tasks is set to every gateway. This shows that the formula (2) is a rough computing.

To implement the task allocation optimization, the objective metric is needed. From formula (3), the life of all gateways at time t can be calculated. The goal is to maximize the life of all gateways. The sum of weighted-value of the remaining life of all gateways may be the one of objective metrics. A High weight reflects the importance of extending the life of a particular gateway for current and future missions.

$$\sum_{i=1}^{N} \omega_i * Life_i(t) \rightarrow \max \quad (4)$$

Another objective metrics is the total energy remains.

$$\sum_{i=1}^{N} E_i(t) \rightarrow \max \quad (5)$$

And the average remaining energy is also an important objective metric that ensures the all gateways have similar lifetime.

$$\sum_{i=1}^{N} (E_i - E_{avg})^2 \rightarrow \min \quad (6)$$

3.3 Task Allocation with Genetic Algorithm

Genetic algorithms are modern search techniques that start from initial population of potential solutions to the problem. Better solutions are got through a repetitive of genetic operators such as crossover and mutation. Genetic algorithm is applied to task allocation as following:

1. Allocate Φ to N gateways randomly which satisfy constrains. m*N metric A represents the allocation. $A_{ij}=1$ means task i is allocated to gateway j. Obviously A is a member of initial population. The initial population is represent by $P=\{A\}$.
2. Evaluate objective metrics (4,5,6)
3. Create new population by repeating following steps
4. Calculate the average objective metrics. If objective metrics of a member population is better than the average, the member survives. Otherwise, the member dies. (Selecting)
5. With crossover operator, generate new offspring to replace the members that have worse objective metrics. (Crossover)
6. With mutation operator, generate new offspring. (Crossover)
7. Evaluate objective metrics (4,5,6)
8. Check constrains and add new offspring to population P
9. If the best member in P is satisfied, return the task allocation solution. Otherwise go to 3.

Crossover operator and mutation operator are described in [4].

3.4 Related work

Task allocation to computing resources is studied in [7,8]. Resource consumption models and the mathematical formulation of the allocation strategy are the main variations among resource allocation problems. Distributed resource management is also modeled as a distributed constrained satisfaction problem. Approaches for distributed constrained satisfaction are achieved. One approach to distributed resource management that handles dynamic changes in resource requirements through the lifetime of a single task is discussed in [9,10]. The idea is to continuously monitor resource usage at each node via a resource manager module, which is main idea of this paper. When a node is about to run out of resource capacity risking the fulfillment of task timing constraints, the resource manager of that node will establish negotiations with its counter parts at other nodes to migrate some tasks and free up some capacity of its local resources. Such approach is resource demanding in itself and might be justified in time-critical application where task timeliness is of great importance.

However, these works lack energy consideration. In addition, for WSN, there is little research on task allocation for nodes. It is reasonable that distributed computing is rarely considered in WSN application. But when a optimal task allocation model of gateways is established, there will be a balance energy consumption for all gateways. The base stations perform task allocation among the gateway with consideration the need of the task and the efficiency of performing such a task while maximizing the life of the gateway.

4. EXPERIMENTS AND ANALYSIS

4.1 Experimental Setup

Among three metrics and corresponding objective functions, the remaining energy is the core of the objective function. The area of wireless sensor network is 100*100 meter^2. 64 sensor nodes and other 5 gateways are placed randomly which are clustered into 5 clusters. The energy consumption for computation and communication for a gateway node is by the model in section 2. According to these objective functions and target-tracking environment, we have design a set of parameters chosen for the target-tracking experiments that are listed in Table 2. The cover area of target is a circle area with radius 5 meters. To compare and analyze experiment results, we have specified 5 trajectories that are randomly used for experiments.

Based on the above model, we calculated the energy required by a gateway to route data to another gateway. The distance between the two gateways is considered to be between 2 and 20 km. Sensors are picked for the generated tasks using a normal distribution.

4.2 Result Analysis and Conclusions

After the optimal allocation of tasks to the gateways, we calculated the above three objective functions mentioned in section 3.1. During different times target tracing, results obtained by genetic algorithm are compared with results obtained without performing any optimization.

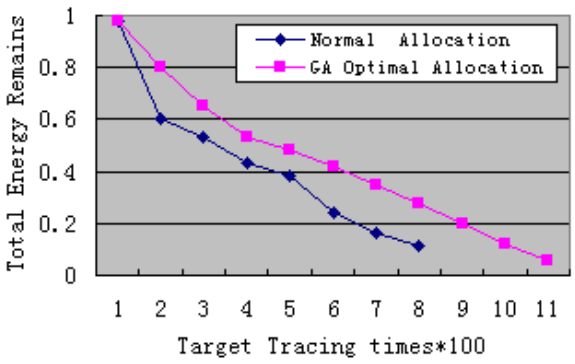


Fig. 2. Total remaining energy of gateways for target tracing times*100

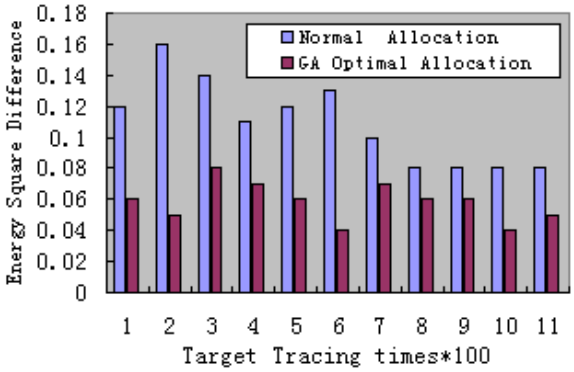


Fig. 3. Energy square difference of gateways for target tracing times*100

The results shown in Fig.2 show that optimized task allocation by genetic algorithm prolong the remains lifetime due to balancing tasks in all gateways. The remains lifetime is nearly a linear function of target tracking times. But curves clearly show that the energy consumption rate by GA

task allocation is less than that of by the allocation method without optima ion. The total remaining energy by GA task allocation is reasonably more. After about 800 times tracking, all gateways cannot execute tracking task without optimal task allocation whereas all gateways can implement about 1100 times target tracking, which can be explained by Fig.3. One of goals of optimal task allocation is to balance the energy consumption for all gateways. It is wanted that remaining energy among all gateways is nearly same and the metric calculated by (6) is shown in Fig.3. It is clear that there are many performance gains achieved. Such performance gains are due to our optimal handling of tasks. For example, long distance communications between gateways and base station would be implemented by multi hop communications among gateways and finally to base station, which may balance and reduce the energy consumption.

A genetic algorithm based optimal task allocation in gateways for target tracing in a cluster-based sensor network is presented in this paper. The approach maximizes the lifetime for the gateways and WSN for target tracking application. Several different objective functions are described and used for optimal method. Simulation results have been analyzed which show that optimized allocation of tasks have better results than non-optimized case in terms of energy usage.

5. REFERENCES

[1] S. Patten, S. Poduri, et al, “Energy-quality tradeoffs for target-tracking in wireless sensor networks”, Lecture Notes In Computer Science 2634(2003): 32-46

[2] H. Yang, B. Sikdar, “A Protocol for tracking mobile targets using sensor networks”, Processing of IEEE workshok On Sensor Network Protocols And Applications, May 2003.

[3] B. Krishnamachari, R. Bejar, et al, “Distributed Constraint Satisfaction and the Bounds on Resource Allocation in Wireless Networks”, Proceedings of the International Symposium on Communications Theory and Applications, July 2001.

[4] Y.X. He, Distributed Operating System, Beijing: Higher Education Press, 2005.

[5] L. Subramanian, R. H. Katz, “An Architecture for Building Self Configurable Systems”, IEEE/ACM Workshop on Mobile Ad Hoc Networking and Computing, August 2000.

[6] M. Younis, M. Youssef, et al, “Energy-Aware Routing in Cluster-Based Sensor Networks”, Proceedings of the 10th IEEE/ACM International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS2002), Fort Worth, Texas, October 2002.

[7] H. Siegel, S. Ali, “Techniques for Mapping Tasks to Machines in Heterogeneous Computing Systems”, Journal of Systems Architecture, Special Issue on Heterogeneous Distributed and Parallel Architectures: Hardware, Software and Design Tools, 2000.

[8] M. Maheswaran, “Dynamic mapping of a class of independent tasks onto heterogeneous computing systems”, Journal of Parallel and Distributed Computing, Nov. 1999, Vol.59: 107–131.

[9] D. Fernandez-Baca, “Allocating modules to processors in a distributed system”, IEEE Transaction on Software Engineering, November 1989, Vol.SE-15: 1427–1436.

[10] B. Krishnamachari, R. Bejar, et al, “Distributed Constraint Satisfaction and the Bounds on Resource Allocation in Wireless Networks”, Proceedings of the International Symposium on Communications Theory and Applications, July 2001.

Research of Migrating Linux to Embedded Systems and Analysis of Real-Time Performance of Embedded Linux

Hanxiao Shi, Guiyi Wei
Zhejiang Gongshang University
Hangzhou, 310035, P.R. China
Email: hxshi@mail.zjgsu.edu.cn
weiguiyi@tom.com

ABSTRACT

At present, the embedded systems are becoming more and more popular for their portable and multi-functional properties. They are used in many domains and located ubiquitous. In this paper, we describe a framework for migration from normal Linux platforms to embedded Linux systems. By the detailed analysis of current real-time performance and process management methods in normal Linux, we reformed its scheduling mechanisms and provided a new approach to promote its performance in the real-time application environment of embedded Linux system. At the end, several experiments are made to explain the reforming methods of the process scheduling mechanism of Linux, and the effectiveness and efficiency of our approach are demonstrated.

Keywords: Embedded Linux, Real-Time OS, Scheduling mechanism, Migration of Linux.

1. INTRODUCTION

Embedded systems have already become ubiquitous in today's world, where these systems are used in a multitude of application domains. These domains range from controllers for household appliances (each washer or microwave is equipped with some form of computer control), to automotive control systems (where systems like the anti-lock braking system, ABS, or the electronic stability program, ESP, are standard equipment in each new car), to systems that are employed for controlling large facilities (like office buildings, hotels, and airports). With an ever-increasing demand for new functionality and the number of devices that should be controlled, the complexity of these systems rises [1]. Linux is a general OS; its kernel is free for everybody. This paper mainly introduces the framework of migrating Linux kernel to embedded systems, proposes the real-time problem of Linux and a solution scenario of directly modifying the scheduling mechanism of Linux kernel.

2. THE OVERVIEW OF EMBEDDED LINUX

2.1 The Introduction of Linux

Linux is an operating system that was initially created as a hobby by a young student, Linus Torvalds, at the University of Helsinki in Finland. He began his work in 1991 when he released version 0.02 and worked steadily until 1994 when version 1.0 of the Linux Kernel was released. The kernel, at the heart of all Linux systems, is developed and released under the GNU (General Public License) and its source code is freely available to everyone. It is this kernel that forms the base around which a Linux operating system is developed. There are now literally hundreds of companies and organizations and an equal number of individuals that have released their own versions of operating systems based on the Linux kernel.

Apart from the fact that it's freely distributed, Linux's functionality, adaptability and robustness, has made it the main alternative for proprietary Unix and Microsoft operating systems. IBM, Hewlett-Packard and other giants of the computing world have embraced Linux and support its ongoing development. More than a decade after its initial release, Linux is being adopted worldwide as a server platform primarily. Its use as a home and office desktop operating system is also on the rise. The operating system can also be incorporated directly into microchips in a process called "embedding" and is increasingly being used this way in appliances and devices.

Linux has a so-called module mechanism, which can dynamically load device drivers or service programs when Linux kernel runs. Linux was initially realized on the platform of Intel-80386, it was migrated to varied CPUs now. Moreover, Linux supports different system structures, not only single-CPU structure, but also multi-CPU structure.

2.2 The Introduction of Embedded Linux

With the continuous increasing of application requirements of embedded systems and complexity of system, people have known that OS plays a great role in embedded systems as well as in PC. With the integrating Linux with embedded systems, embedded Linux arises. Embedded Linux increasingly becomes main choice for embedded OS in many embedded application development.

More and more enterprises and research institutions pay attention to embedded Linux, for it has many advantages. We list the following eight advantages:

1) Open source and abundant software resources: Linux kernel is open to everybody, customers have great freedom. Linux's resources are very abundant, every general program almost can be found in Linux.

2) Powerful kernel and multitask: Linux kernel is very powerful and stable, its high efficiency and stability is proved in varied area, especially in network server.

3) Support varied system structure: Linux has been migrated to tens of hardware platforms, such as X86, ARM, MIPS, ALPHA, SPARC, and so on. It almost supports most CPUs.

4) Satisfactory management mechanism of network communication, graphics and files: Linux is very fit for network from its origination; the function of network is very strong in Linux. Moreover, Linux also supports many kinds of file systems and graphic systems.

5) Support massive peripheral hardware devices: Drivers in Linux are more abundant, they support varied popular hardware devices and latest hardware technology. With the High modularization, it is very easy to add

components to Linux.
6) Customize the size and functions of kernel: Linux succeeds to excellent design thoughts of Unix; it is very flexible that components can be customized freely.
7) Have better development environment and more development tools: Linux has many better development tools and not a few free IDEs.
8) Get a lot of support from many developers: The free spirit of Linux attracts millions of programmer to join the process of Linux development and testing, this causes Linux to become a strong OS in a short time. At the same time, as Linux has succeeded in PC, people are convinced of good prospects of migrating Linux to embedded systems.

3. THE MIGRATION FRAMEWORK OF LINUX

Most projects of embedded Linux are developed based on the migration of Linux. Migration means that source codes on one system are transferred to other systems. Linux was initially realized on the platform of Intel-80386, and increased many migration codes on target system according to different hardware architecture [2,3]. Linux kernel has more extensive usage on different hardware, but it cannot be used in any platform. It can be used in target system through the approach of cross compile, which uses cross compile driver in development platform A to compile and link the source code into the executable code in target platform B.

This paper describes the framework as the following three parts:
1) Memory Management: In order to use memory resource efficiently, we implement and optimize the dynamic management of system memory. As there are no hardware conditions supporting virtual memory mechanism in most embedded systems, we get rid of the part of virtual memory management mechanism during the migration of Linux.
2) Multi-Process: Multi-process mechanism of embedded systems should be as possible as efficient and properly real-time. We try to break through the limit of real-time switch of process during the process scheduling of Linux, design and implement a kind of flexible preempted process scheduling.
3) Peripheral Resource Management: The kind and quantity of peripheral devices is relatively fixed on the special platform, but in order to fit in with future hardware device, we present an extended solution scheme for peripheral resource management.

During the process of concrete migration, we should modify Linux kernel, take off some unnecessary codes to adapt target systems.

4. RESEARCH OF REAL-TIME PERFORMANCE OF EMBEDDED LINUX

It is very important to research into the real-time problem in embedded systems. Embedded Linux is migrated from standard Linux kernel; it keeps the qualities of Linux so that its real-time problem is also originated from Linux [4,5]. During the process of research, we need to improve the real-time performance of embedded Linux. The quality of real-time performance is due to many functions; this paper discusses chiefly the real-time performance from the side of process scheduling mechanism.

4.1 The Overviews of Real-Time Performance of Linux

General OS is not often designed for real-time applications; it should be modified to fit in with the requirement of real-time processing. Linux is a general OS, and its kernel has several shortages in real-time processing [4,6,7]:

1) It adapts variable priority scheduling based on the fixed time slot; it is non-preempted in process scheduling. When a process is waiting for system call, it will wait until current process scheduling is over. Even if it is a process having a higher priority, it cannot get the resources of CPU. The time of system scheduling is non-predictable practically, so it cannot satisfy the requirement of real-time applications.
2) The locking and unlocking of processes having big granularity takes more time, it cannot satisfy the requirement of real-time performance.
3) Management of virtual memory becomes a bottleneck of real-time applications. There are many real-time tasks should be loaded in memory, it takes more time to go out and in memory if use the technology of virtual memory. Obviously, it cannot fit in with the requirement of real-time applications.
4) Linux adopts fixed clock interrupt, so it cannot satisfy the real-time application having high precision.
5) It often shuts system interrupt during system call and drivers executing, blocks the instant processing of interrupt having high priority.

4.2 Analysis of Current Scheduling Mechanism in Linux

There are two main concepts of scheduling mechanism, one is the preempted approach, and the other is the non-preempted approach. When a high priority task is coming under preempted scheduling mechanism, the current executing task having low priority can be preempted; the process of preempted scheduling is shown in Fig.1. However, When a high priority task is coming under non-preempted scheduling mechanism, the current task having low priority is still executing, the high priority task can be scheduled until the current one is over, the process of non-preempted scheduling is shown in Fig.2.

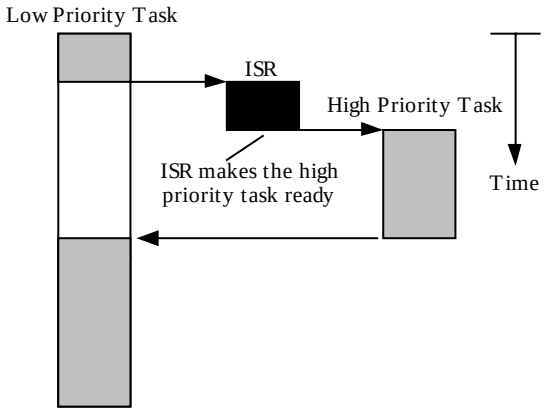


Fig. 1. The process of preempted scheduling

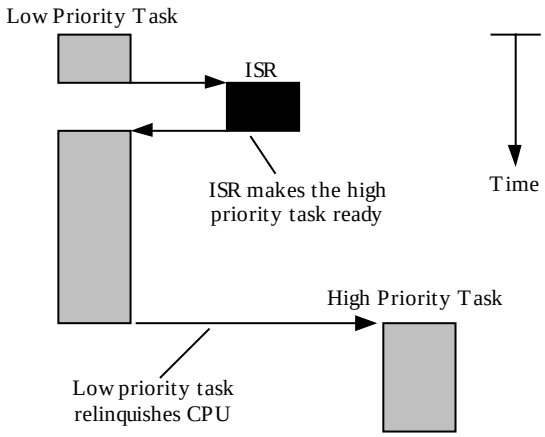


Fig. 2. The process of non-preempted scheduling

There are three questions of scheduling mechanism in Linux:

1) When can system begin to schedule?

Linux divides the executing period of system into kernel status and user status. When task A enters kernel status by system call (SYSTEM_CALL1), at the same time, SYSTEM_CALL1 is interrupted by INTR1. After the processing of INTR1 is over, it returns to SYSTEM_CALL1. At this time, if INTR1 produces more urgent task B, it cannot be immediately implemented until SYSTEM_CALL1 is over. Then system status has changed, so the requirement of calling task B cannot be satisfied, system also loses the real-time signification.

2) How to select the next task?

At present, many real-time operating systems adopt priority to describe the urgent degree of task. Linux has similar mechanism and many scheduling algorithms, such as SCHED_RR, SCHED_FIFO, and so on. So Linux can select the next task by its priority.

3) When can execute the next task?

When standard Linux goes into the end of executing scheduling function, it calls the switching function to transfer the previous process to the next process. It needn't to switch the processes if the previous process is the same as the next one. Table 1 describes the algorithm of executing next task.

Table 1. The Algorithm of Executing Next Task.

```
if (unlikely(prev == next)) {
        prev->policy &= ~SCHED_YIELD;
            goto same_process;
        }
……
        switch_to(prev, next, prev);
……
same_process:

reacquire_kernel_lock(current);
if (current->need_resched)
        goto need_resched_back;
return;
```

4.3 Design of Modification Scenarios

Through the analysis of scheduling mechanism of Linux [8,9,10], we begin to propose some modification scenarios about real-time performance from the side of scheduling mechanism. At present, there are two scenarios about the modification of Linux, one is to re-design and implement the real-time kernel bypassing Linux kernel, and installs it between hardware interrupts and Linux kernel. The other is to modify Linux kernel directly in order to satisfy the requirement of turning Linux into real-time OS fundamentally. We select the latter for some beneficial researches, because it can help us to know better the property of real-time OS.

We design the modification scenarios as following:

1) Increase a lock for preempting kernel resource.

2) Improve the management of locks when the executing process is interrupted.

3) Modify the scheduling mechanism and system call.

As the limit of the text size, we don't explain codes modification in details.

In three questions of scheduling mechanism, the question of "When can system begins to schedule?" restricts the preempted performance of Linux kernel. However, the emphasis of real-time system is whether it can be scheduled after interrupt returns to kernel status.

4.4 Experimental Methodology and Results

This experiment is main to modify process-scheduling mechanism of Linux; it aims to allow the kernel to do process scheduling after interrupt returns to kernel status. If the experiment is realized, it proves that embedded Linux can be modified to real-time system by modifying process scheduling mechanisms.

We modify dependent source file of Linux, see Table 2.

Table 2. Modification of Source Codes

Source codes:

```
ret_from_exception:
    movl EFLAGS(%esp),%eax
    movb CS(%esp),%al
    testl $(VM_MASK | 3),%eax
    jne ret_from_sys_call
    jmp restore_all
```

Modification codes:

```
ret_from_exception:
    movl EFLAGS(%esp),%eax
    movb CS(%esp),%al
    jmp ret_from_sys_call
```

There are three steps of Execution after compiling:
1) First execution:
Inputting “c”(one of the GDB Commands), kernel cannot start normally. The screen shows as following:

{
Partition check:
hda: hda1 hda2 < hda5
}
The wrong information appears:
{
Scheduling in interrupts
Kgdb assertion failed: BUG
}

If continue to inputting “c”, above wrong information appears all the time.
2) Second execution:
Inputting “c” also, kernel still cannot start normally. The screen shows as following:

{
Linux NET4.0 for Linux 2.4
Based upon Swansea University Computer Society NET3.039
Initializing RT netlink socket
}
The wrong information appears:
{
Scheduling in interrupts
Kgdb assertion failed: BUG
}

If continue to inputting “c”, above wrong information appears all the time.
3) Third execution:
After inputting “c” also, kernel still cannot start normally. The screen shows as following:

{
Partition check:
hda:
}
The wrong information appears:
{
Scheduling in interrupts
Kgdb assertion failed: BUG
}

If continue to inputting “c”, above wrong information appears all the time.

4.5 Analysis of results

The error information is the execution result of the function (asmlinkage void schedule (void)) in source file (/kernel/sched.c), main codes are shown as following:

```
{
    if(unlikely(in_interrupt())){
        printk("Scheduling in interrupt\n");
        BUG();
    }
}
```

It can be concluded that the experiment has already called the scheduling mechanism in kernel status of system. In other words, through modification we can realize the real-time calls in kernel status. But there are an exception checking in scheduling functions, it bottlenecks the formal processing of system. In fact, we can also discard the exception checking in schedulers through modifying the code in the file (/kernel/sched.c). As the limit of the text size, we don’t explain in details.

5. RELATED WORK

In this section, we review research work in the area of the real-time performance of embedded Linux. There are three kinds of different real-time Linux relevant to our research objectives, which are discussed, bellow.

5.1 RT-Linux

RT-Linux [11,12] is an operating system that allows real-time control of machinery and data from a Linux environment. RT-Linux is a hard real time operating system with guaranteed response times (up to hardware limits). Many "real-time" operating systems offer " typical" response times instead. RT-Linux was originally developed at the New Mexico Institute of Technology. Response times are close to hardware limits. On a modest, reasonably configured, x86 PC a RT-Linux interrupt handler will run under 10 microseconds from the moment the interrupt was asserted and a RT-Linux periodic task will run worst case within 30 microseconds of its scheduled time. On better hardware, these times shrink. Of course, if you insist on bad hardware, you can make things run worse. Programs are developed in a standard Linux environment with additional capability of connecting to real-time tasks. For example, it is easy to write a Perl script that displays data in X-Windows, responds to commands delivered over a network, and collects data from a real-time task.

RT-Linux has not improved the real-time performance of Linux itself; but adopts " impracticable" Linux kernel technology, realizes simple real-time kernel in the lower floor of Linux. Linux itself is regarded as the minimum task of PRI in the real-time kernel, and the PRI of all real-time tasks is higher than Linux itself. We can regard this real-time kernel of RT-Linux as a virtual machine or intermediate level between hardware and standard Linux kernel

Strictly speaking, RT-Linux is not a Linux operating system at all, its core is a real-time micro kernel of independent development, and Linux kernel is downgraded to the process in this real-time micro kernel.

5.2 RED-Linux

RED-Linux [13] is a real-time and embedded version of Linux. Its design object is on the basis of keeping standard Linux existing function, puts the real-time processing mechanism into the kernel, thus set up a real-time operating system kernel that is used in the embedded environment.
To choose one suitable system time frequency is very important in real-time system, a microsecond timer has been designed in RED-Linux, which is more accuracy than standard Linux.

In order to serve the real-time task, the kernel should be able to preemptive. But the kernel of Linux can't be preempted. RED-Linux has been adding so-called preemption point among the kernel. These preemption points divide the execution of kernel process, such as interrupt processing programs, into lighter, more manageable blocks. It will check whether there are higher PRI processes waiting for scheduling when system executes to these preemption points. If have, it stops executing the current process, then switches over to the process having higher PRI.

Another unique design of RED-Linux is a pair of module scheduler. The scheduler of RED-Linux is made up by two parts of the allocator and dispatcher. Dispatcher is resident in kernel space; it is one component of kernel. Allocator is resident in user space; it can be incorporated into the kernel space too in case of necessity. The allocator faces to application programs directly, deals with their real-time demands, and turns these demands into a group of specific parameters, such as scheduling strategies, time started, finish time, and so on, which are transmitted to the dispatcher finally. The task of the dispatcher is to insert the specific parameter of each real-time task in the task queue, and chooses the next executing task. The allocator defines the tactics of process scheduling; the dispatcher offers the concrete implementation mechanism. In this way, RED-Linux has realized the scheduling of the real-time and non-real-time task at the same time under the single kernel. The whole system can supports various kinds of real-time scenario under a uniform scheduling frame, such as the priority-driven scheduling, sharing-driven scheduling, time-driven scheduling, and so on.

The block size and the choice of insertion position of preemption points in RED-Linux have a great influence on the system execution efficiency. Kernel can be preempted to in these preemption points, but it is not solved completely that the whole kernel can be preempted. Moreover, preemption points are already fixed in the design, cannot be changed according to application needs.

5.3 KURT

KURT [14,15] is a real-time operating system with multiple kernels mechanism based on Linux, which is developed by Kansas University. It uses different kernel status to control the scheduling of real-time processes:

1) General status, the kernel does not offer real-time support, it was the scheduling mechanism of standard Linux.
2) Real-time status, the kernel only supports real-time process, all non-real-time processes are hung up.
3) Mixture status, the kernel supports real-time process and non- real-time process at the same time.

KURT supports the real-time process scheduling, at the same time, it increases the complication of system kernel. This will cause a lot of questions, for example, System switching improper among the three status most likely leads to system crash, system security questions more rely on the programmers' design and implementation.

Synthesizing above major technology of the real-time operating system based on Linux, we can see that these design methods used to reform Linux kernel are all soft real-time implementation. Can we find a kind of hard real-time implementation method to revise Linux kernel? According to research results of this domain at present, it is not an easy thing. This is a question of design other than the question of implementation.

6. CONCLUSIONS

This paper elicits the real-time problems of embedded Linux through the migration of Linux. It proposes the experiment framework for exploring the real-time scheduling of Linux kernel, and implements a kind of experiment scenario of directly modifying the scheduling mechanism of Linux kernel. Through the experiment we can make a conclusion that the non-preempted performance of Linux kernel is a design problem. There are many restrictions and protects of for kernel in source codes. When we try to break these restrictions, some unknown exceptions or mistakes appear. If make no deep apprehension on Linux kernel, we will stumble at every step. At the same time, it proves that the requirement of realizing its real-time performance through modifying process scheduling mechanism in Linux kernel is available.

7. ACKNOWLEDGMENTS

The authors wish to thank the Natural Science Foundation of Zhejiang Province, P. R. China for the National Science Fund under grant Number Y105356. We would like to thank CRMB (the Center for Research in Modern Business, Zhejiang Gongshang University) for they partially supported this research. We would also like to thank our referees for their helpful comments and suggestions.

8. REFERENCES

[1] A. Metzger, "Feature interactions in embedded control systems", Computer Networks, Vol.15, No.5, August 5, 2004, pp. 625-644.
[2] S.D. Lin, A.H. Ren, et al, "Porting Linux kernel to new hardware platform", Journal of Beijing University of Aeronautics and Astronautics, Vol.29, No.3, March 2003, pp: 197-201.
[3] T.W. Fredian, M.Greenwald, et al, "Migration of alcator C-Mod computer infrastructure to Linux", Fusion Engineering and Design, Vol.71, Issue: 1-4, June 2004, pp. 89-93.
[4] Y. Zhou, M.S. Li, "Research and implementing of real-time support of Linux kernel", Journal of computer research and development, Vol.39, No.4, April 2002, pp: 466-472.
[5] V. Vitale, C. Centioli, et al "Real-time Linux operating system for plasma control on FTU-implementation advantages and first experimental results", Fusion Engineering and Design, Vol.71, Issue: 1-4, June, 2004, pp. 71-76.
[6] G.L. Xing, H.L. Wei, et al, "Analysis and research to Linux-based real-time operating system", Mini-Micro

system, Vol.22, No.8, August 2001, pp: 1018-1021.
[7] L. Kwei-Jay, "RED-Linux: An Open Real-Time Operating System", REDSonic, Inc.
[8] H.R. Zhong, C.S. Wang, "The implementation and optimization of scheduling algorithms in embedded Linux", Computer engineering & science, Vol.24, No.3, March 2002, pp: 5-7.
[9] L. Kwei-Jay, W. Yu-Chung, "The design and implementation of real-time schedulers in RED-linux", Proceedings of the IEEE, Vol.91, No.7, 2003, pp: 1114-1130.
[10] T.W. Kuo, L.P. Chang, et al, "Efficient Online Schedulability Tests for Real-Time Systems", IEEE Trans. Software Eng., Vol.29, No.8, 2003, pp: 734-751.
[11] B. Michael, "An Introduction to RT-Linux", http://rtlinux.lzu.edu.cn/documents/ documentation/RTLManual/RTLManual-1.html, March 28, 1998.
[12] A. Romanenko, José A.A.M Castro, "An RT-Linux based control system of a pilot plant for reaction kinetics and process control studies", Computers and Chemical Engineering, Vol.24, Issue: 2-7, July 15, 2000, pp: 1063-1068.
[13] RED-Linux: for Real-Time and Embedded Systems, http://www.linuxdevices.com/links/ LK7432493143.html, January 19, 2000
[14] KURT: The KU Real-Time Linux, http://www.linuxdevices.com/links/LK5455135050.html, March 24, 2000
[15] W. Dinkel, D. Niehaus, et al. KURT-Linux user manual. http://www.ittc.ku.edu/kurt/papers/user-manual-DRAFT.pdf, March 29, 2002

Research on Aspect-Oriented Software Architecture Model for Embedded System and Its Application

A-Qun Deng, Huanjun Yu, Shangxu Hu
College of Material and Chemical Engineering, Zhejiang University
Hangzhou, Zhejiang 310027, China
Email: deng-aqun@hzfuji.com, yuhj@infotech.zju.edu.cn , sxhu@mail.hz.jz.cn
Xiaojun Li
College of Computer Science and Information Engineering, Zhejiang Gongshang University
Hangzhou, Zhejiang 310007, China
Email: lixj@mail.zjgsu.edu.cn

ABSTRACT

Aspect-Oriented Programming (AOP) is a relatively new technique, which makes it possible to overcome the weakness of Object-Oriented Programming (OOP) technique by resolving the crosscutting problem between concerns and making Multi-Dimensional concerns to separate from each other. The problems and weakness of traditional software architecture for embedded system software was analyzed. In order to overcome these problems and weakness, AOP technique was adopted and an Aspect-Oriented software architecture model for embedded system software was proposed in this paper. Its validity was illustrated by practical application in the software development of vending machine, and the advantages of this model were discussed also.

Keywords: Object-Oriented Programming, Aspect-Oriented Programming, Software Architecture, Software for Embedded System.

1. INTRODUCTION

Embedded system is an integration of software and hardware, which is designed to realize specific function by the pragrammbale controller. In recent years, with the increasing of the complexity of hardware and the significant demand for function and performance of system, particularly, in terms of more flexible and convenient connection with network, portable and electric saving wireless and powerful multi-media process capability, the software development work increased rapidly, even up to 70-80% of total work. Hence, the development of embedded software becomes the bottleneck of the development of embedded system [1]. In addition, comparing to the development of PC software, the development of software for embedded system shows low efficiency, high technology requirement, changealbe capability requirements and difficulty in maintenance of software. So, it becomes very important to increase the efficiency of software development and to raise the system stability for embedded system.

In order to make it easier to develop and maintain software, in 1968 Dijkstra introduced firstly the concept of software architecture. With the continuous development of idea, at the beginning of 1990s, a solution of software architecture was proposed, which tried to use more precise way to define software architecture. Recently, more and more researchers focus on the software architecture [2], while such researchs in China are quite few, the main reasons are: 1) Some people think that software architecture is an old topic. 2) In comparing to developed countries, we have less experiences in developing large scale and super complicated software, and do not have learned sufficient lessons on software disaster. Hence, our acknowledge regard the importance and necessity for the research on software architecture is still not enough. Moreover, the application of software architecture on embedded sytem is reported in even fewer domestic literature.

Based on the analysis upon the problems and insufficiencies of traditional Object-Oriented software architecture, by introducing Aspect-Oriented Technology (AOT) which is the most up-to-date technology of Multi-Dimensional Separation of Concern (MDSOC), an Aspect-Oriented Software Architecture Model for embedded system software was proposed in this paper, and applied to the devolpment of control software for vending machine, on which the author engaged in past several years. The efficiency and advantages of the proposed approach was proved.

2. REVIEW ON ASPECT-ORIENTED PROGRAMMING TECHNOLOGY

According to Dijkstra's point [3], many concerns exist when the software system is broken down into parts of manageable size. On the contrary, when implementing the sub-parts and composing them to get the whole program, it is impossible to realize different concerns in different basic building blocks of commonly used programming languages (procedures or classes). The reason is that it is well suited to most types of components that relate to a system's primary functionality and fails to implement special concerns like data flows in distributed systems, failure recovery, persistence, process synchronization etc. So it will result in realization of many concerns in one basic building block. The crossing realization of multi-concerns is called as CrossCutting, which will either make the program code difficult to develop, understand, and maintain, or decrease the adaptability and reusability of program code. Aspect-oriented Programming (AOP) is a new programming technique intended to solve this problem [4, 5].

As for the separation of concerns, AOP distinguishes components from aspects. Components tend to be units of the functional decomposition of system and represent the problem domain concerns as it is defined from the client perspective. Aspects tend not to be units of the functional decomposition of system, but rather to be properties that affect the performance or semantics of the components in systemic ways which represent the solution domain concerns as defined by the solution techniques such as process synchronization, failure recovery, safety and real-time control.

In order to compose the concerns and get the whole program, AOT adopts a technique called Weaving, which defines several basic concepts. One is Joint Point, which is connection point between component and aspect during the combination of concerns. The second is Pointcut which is the integration of Joint Point. The third is Advice, which defines pieces of aspect implementation executing at well-defined points in the execution of the program. Those points can be given either by named Pointcuts or by anonymous Pointcuts. The working theory of AOT is illustrated in Fig. 1.. Aspects have to be combined or woven with components in order to use the language compiler of component to produce executable programs. This weaving process is performed by a new kind of compiler called Aspect Weaver. Aspect Weaver can find the interface of Joint Point and then insert the codeof aspect into the relevant position of component and then obtain the executable applications.

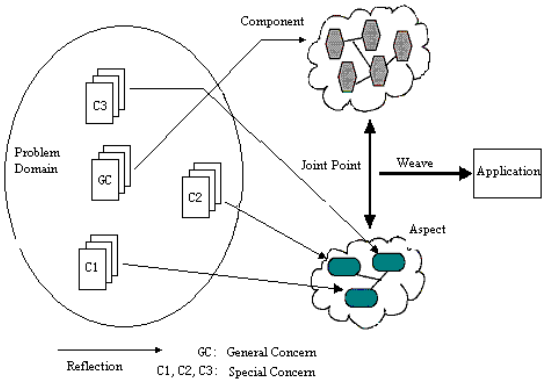


Fig. 1. Sketch map of working theory of AOT

The advantages of AOT are listed below. 1) Different concerns are able to be realized in different aspect which increases the readability of program and maintainability. 2) AOT provides a general method for expressing concerns which increase the reusability of software by introducing aspect (equalivent to the class of Object-Oriented Programming). 3) AOT also provides the mechanism for combining concerns by introducing concept such as Joint Point, which can realize the perfect combination of component and aspect. 4) AOT can realize the perfect combination with existing programming language without changing programming language of component. 5) AOT also simplifies the process of software development to certain degree. The traditional software development is too complicated since there exists complex relationship between concerns. AOT can solve the CrossCutting trouble and decrease the complexity of program development.

3. DISADVANTAGE OF TRADITIONAL SOFTWARE ARCHITECTURE FOR EMBEDDED SYSTEM AND WEAKNESS OF OOP

Embedded system software has two characters, i.e., hardware control and real-time control. In general, embedded system software includes three parts: hardware control, time control and application logic for managing these two control units. During the traditional development of software for embedded system, the software architecture illustrated in Fig. 2. is adopted.

After analysing the software architecture, it could be found that there are a lot of problems existing. The major problems are: 1) Complicated relationship between different components. Not only the relationship between model layer (application layer) and control layer is complicated, but also the inter-relationship between the model layer and the control layer is complicated according to the functional demands of software. 2) When the function of one component changed, the related components also need to be changed accordingly. Due to the close relation between components, when the function of one component changed, the related components must be changed accordingly, especially when hardware is changed or new hardware is added, the related control layer will be changed significantly and more difficulties in maintenance of software will be resulted. 3) Due to the high integration of hardware, it is hard to define the components with high reusability. 4) The sizes of reusable components are not the same.

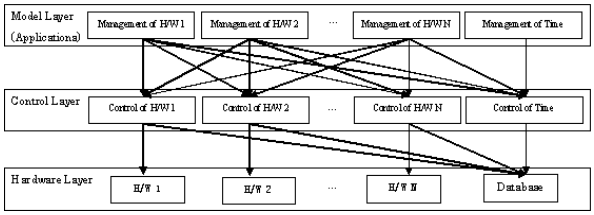


Fig. 2. Traditional software architecture for wmbedded system

In addition, when Object-Oriented programming technique was used to implement this software architecture, due to the limitation of OOP, there are two types of problems, which cannot be solved perfectly. One is the CrossCutting between concerns discussed above. During the development of software for embedded system, it is necessary to deal with something such as concurrent, synchronization, real-time and exception. All of these problems concerned with hardware control come to the CrossCutting, which cannot be modularized perfectly by OOP. This kind of CrossCutting always happens between modules. The second, for the design of software for embedded system, some objects always have different action character when they are in different places or have different roles. When OOP is employed to built the model of software, all types of action character of places or roles will be encapsulated in one module unit, that causes difficulties to understand and maintain the software due to lack of modulation and encapsulation. When analysis this problem, it will be found that the real problem is also the CrossCutting between concerns which happen in the same module unit at different place or for different roles.

AOT is aiming at solving the CrossCutting between concerns and in order to get optimal modularization of system, it is possible to realize different concerns in different module unit as its basic theory. So, it is an effective solution for solving problems mentioned above by using AOT to reconstruct the software architecture for embedded system.

4. ASPECT-ORIENTED SOFTWARE ARCHITECTURE MODEL FOR SOLUTION

According to the point of Multi-Dimensional Separation of Concern (MDSOC) [6, 7, 8], after the multi-dimensional separation, software system can be regarded as an aggregation of many aspects. So the aspects can be combined and thus realize the different concerns to get the expecting software system.

Generally, embedded software included three parts, hardware control, time control, and application logic for managing the above-mentioned two controls. When carrying on the hardware control, the frequency on hardware changing, hardware independency, as well as the synchronization for different hardware needs to be

considered. So as the general design rule the next three principles should be followed.

1) Introducing virtual hardware. Hardware control is not the direct control on hardware object, but control on the virtual hardware. When hardware changed, it is possible for only changing the virtual one, which can be solved conveniently. Many virtual hardware objects can be established and can be used multi-functionally for a kind of hardware. As for the application managing two controls, it seems that the application is dealing with many hardware objects. By introducing the virtual hardware objects, when access the hardware, message will not be sent to hardware object, but to the virtual one. It is independent for all virtual hardware objects for a kind of hardware and no message is necessary to be sent between each other.

2) Concurrent process design. Inside the embedded software, all hardware objects act independently. In order to make it close to the real hardware, all the virtual hardware objects need to be designed as concurrent objects.

3) Correspondence between hardware. For the design of embedded software, the method of managing the correspondence between hardware, generally, has two ways, managing the correspondence by state machine of every hardware object and managing the correspondence by state machine of all hardware objects.

Other than hardware control, there is time control. If hardware control is considered as central concern, a good design for time control cannot be obtained, vice versa, it is the same situation. This is CrossCutting problem need to be solved in AOT. When using AOT, time control and hardware can be separated; hence, such software architecture at least included four parts.

1) Application logic Aspect. This is for the control of all hardware, which consisted of many virtual hardware objects.

2) Hardware Aspect. This is consisted of hardware and virtual hardware. Virtual hardware is in charge of control of hardware.

3) Time Aspect. This is consisted on virtual timer and timer and virtual timer is in charge of control of timer.

4) Concurrent Process Aspect. This is to realize concurrent process for all virtual hardware objects.

So, the Aspect-Oriented Software Architecture Model can be defined as illustrated in Fig. 3.. Aspect-Oriented software is an aggregation of aspects, which are the container for all component elements, and the connection between containers is Joint Point. Aspect 1 is concurrent process aspect, and aspect 2 is application logic aspect, and aspect 3 is the time aspect, and aspect 4 is the hardware aspect.

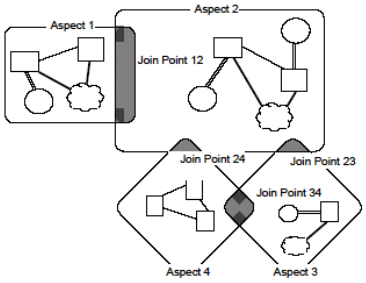


Fig. 3. Aspect-Oriented software architecture model

5. PRACTICAL APPLICATION OF ASPECT-ORIENTED SOFTWARE ARCHITECTURE MODEL

As a practical application, this software architecture model was used in the software development of vending machine. Herein, in order to describe the application in vending machine, money management system, which is a sub-system of vending machine, is given as an example. When analyzing the money object in money management system, it will be found that the action characteristic is different when money is situated in different place in vending machine.

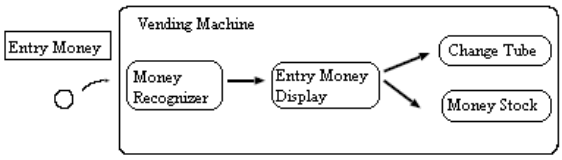


Fig. 4. Movement of money in the vending machine

As illustrated in Fig. 4., when money was put into vending machine, firstly, it must be recognized by money recognizer to distinguish it fake or not. If not fake, it will be accepted by system as substitute for buying commodity. In addition, money will be kept in the charge tube when the given charge is not sufficient. Under the condition of enough change, money will be kept in the money stock. So, there are four kind of correspondence relation between money object and all other objects. If each relation is called as a field, four fields can be obtained and their relation is illustrated in Fig. 5..

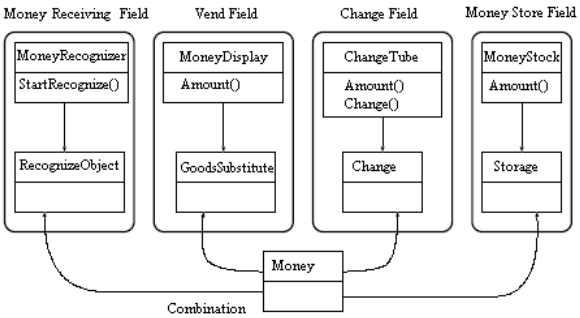


Fig. 5. Fields relative to money

1) Money Receiving Field. In money recognizer, money will be checked to distinguish whether fake or not. As for money recognizer, money is recognized object.

2) Vend Field. During the vending of commodity, money is regarded as substitute for commodity in money display.

3) Change Field. After the vending of commodity, money will be kept in change tube as change.

4) Money Store Field. When changes are sufficient, money will be kept in money tube as deposit.

The action characteristic is different when money is in different field. Being the component element of field, money is not money object while it is operated as other object. Several concerns are separated from object-oriented money management system.

1) Core concerns (hardware control). Embedded software system is designed to control hardware. So, it is nature to break down the system to modules according to the hardware objects, which is called core concern. Money management system is designed for managing multi inter-corresponds hardware objects, i.e.: money recognizer, money display, change tube, and money stock. Considering the convenience of adding new hardware and changing hardware, virtual hardware was designed for each hardware object, and corresponding controller as well. The hardware was operated through virtual one. So, it makes more convenient for changing hardware because the software change caused by hardware change will be focused on the relative virtual hardware. There is no direct communication between hardware after the introduction of controller, and

this leads to more independency among hardware objects. Money management system is equivalent to controller.

2) Concurrent process. Due to the concurrent process among all hardware objects, the realization of concurrent process for individual hardware object becomes CrossCutting of all hardware objects. As illustrated in Fig .6., the managing object and the managed object can be obtained after the module separation for concurrent process. The managed object is the content of concurrent process while the managing object is the monitor in which event is managed by queue, and each managed object get the execution priority in turn to deal with event.

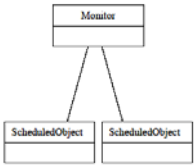


Fig. 6. Concurrent process

3) Real time process. Abnormal action for hardware is checked by accounting time. So, it is necessary for real time process. The model illustrated below in Fig.7 was obtained after analyzing real time process. Real time process is realized by the communication between timer and the object, which use timer (TimedObject). TimedObject need to start timer before process, and also need to be informed when timer is up.

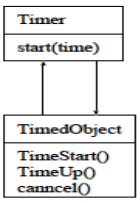


Fig. 7. Real time process

Combining all the above mentioned concerns, the software architecture for money management system can be obtained as illustrated in Fig. 8..

The code of core concern when attaching message with Joint Point, and the call order of attached message are indicated in Fig. 8.. In order to simplify the description, the terms such as before, after and around were adopted as defined in AspectJ [9]. The term before means before the code of core concern, while the term after means after the code of core concern. The term around means that core concern code is replaced by attached message. All virtual hardware and controllers are concurrent objects, and message will be sent to monitor when communicating with these objects. Each virtual hardware object gets execution priority from monitor and then takes action. After ScheduledObject gets execution priority from monitor, message will be send to virtual hardware object. And then each virtual hardware object will send message to the relative money component elements (hardware object). Money recognizer starts to count time when distinguishing money, and system will define it is abnormity when money can not be distinguished in a certain period. So, money recognizer also has the action characteristic of TimedObject (real time process object) and need to send message to TimedObject.

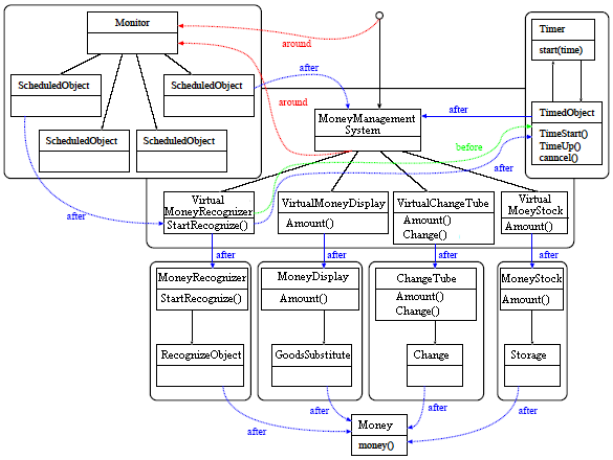


Fig. 8. Aspect-Oriented doftware srchitecture of money management system

6. ANALYSIS AND DISCUSSION

According to the analysis in chapter 5, there are following characteristics for software architecture by using Aspect-Oriented Embedded Software Architecture Model.

1) The separation of concerns from CrossCutting is realized. For example, in the money management system, the concern separation between concurrent process and real time process were realized. The concurrent process and real time process were also modularized individually and Joint Point for aspect became the connector of these two processes.

2) The perfect modularization for the objects, which action characteristic changed with different places or roles, was realized. Like the above mentioned money management system, money is the super class, while the money objects at different places can be implemented as derived class. Integrate the derived class with other related object class to one aspect, and the system will be realized. So, the different objects at different places or roles can be separated perfectly by means of aspect.

Therefore, software architecture obtained by using this model can overcome the limitation of OOP, in the meantime, the following advantages were observed.

1) It simplifies the relation among all components of software. The connection among components changed from previous direct function call to joint point, and the coupling relation between them became weak.

2) It makes software easy to be maintained. In the above mentioned money management system, if a hardware needs to be added for returning money, it does not need to modify original software but can be realized by adding Return Money Field and a virtual money return object. Basically, it does not need to change original code and only need to add some component.

3) Due to the reducing of coupling relation between components, it makes the reusability of component increased, and also makes it possible to obtain the same size of reusable components.

4) Software architecture not only can reflect the model of software, but also can determine the software development process. In this way, software system consists of many aspects, and aspect is the aggregation of several component elements. So, when the layer and relation of all components are clear, it is easy to derive the software development process from software architecture.

7. CONCLUSIONS

Based on the analysis of the existing problem in the traditional software architecture for embedded system, by introducing the most up-to-date technology of MDSOC, an Aspect-Oriented Software Architecture Model was proposed in this paper. The model was strictly tested by practical software development of vending machine, and very good result was obtained. This model will be applied to software development of other embedded system and try to prove the universality of this model in future.

8. REFERENCES

[1] Q. Zhou, Z.Z. Kou, et al, "M ethods to build the unified framework for embedded software", Electronic Product World, No.9, 2002, 217-18.

[2] Z.C. Li, "Actuality and Development of Software Architecture", Journal of Mudanjing Teachers College (Natural Sciences Edition), No.4, 2003, 11-13.

[3] E.W. Dijkstra, "A Discipline of Programming. Prentice-Hall, Englewood Cliffs ", N.J., 1946

[4] A.Q. Deng, X.J. Li, et al, "Issues of a New Technology for Software Design", Systems Engineering and Electronics, 2004, 26(7). 970-975.

[5] D.G. Cao, H. Mei, "A new programming method: Aspect-Oriented Programming", Computer Science, 2003, 30(9). 5-10.

[6] H. Ossher, P. Tarr, "Multi-Dimensional Separation of Concerns and the Hyperspace Approach. In Proceedings of the Symposium on Software Architectures and Component Technology", The State of the Art in Software Development, Kluwer, 2001.

[7] P. Tarr, H. Ossher, et al, "N-Degrees of Separation Multi-Dimensional Separation of Concerns", In Proceedings of the 21st International Conference on Software Engineering, May.1999.

[8] H. Ossher, P. Tarr, "Using Multidimensional Separation of Concerns to (Re) Shape Evolving Software", Communications of the ACM, 2001, 144(10). 43-50.

[9] G. Kiczales, E. Hilsdale, et al, "Getting Started With AspectJ", Communications of the ACM, 2001, 44(10).59-65.

Research of the Rapid Embedded Quantitative Test System For Antibiotics

Qian Xu
College of Information Engineering, China Jiliang University,
Hang Zhou, Zhe Jiang, China
Email: lotus_abright@yahoo.com.cn
Xiaozhu Chen
College of Information Engineering, China Jiliang University,
Hang Zhou, Zhe Jiang, China
Email: chenxiaozhu@cjlu.edu.cn
Wenbin Zheng
College of Information Engineering, China Jiliang University,
Hang Zhou, Zhe Jiang, China
Email: zhengwb@cjlu.edu.cn

ABSTRACT

This system tests for antibiotic residues quantitatively. First, it used a gold-colloid strip to detect the antibiotic on which two lines would appear after about 12 minutes. One is standard line and the other is test line. The color of the test line would be different as the antibiotic concentration was changed. Then it captured the two lines information as a bit map by CMOS image sensor and sent it to S3C2410 processor based on ARM 920T core. At last it called the application to analyze the data and gave the exact antibiotic concentration value. The process takes less than 15 minutes. And the precision can be 0.5ppb.

Keyword: Embedded System, CMOS Image Sensor, Antibiotics Test, Quantitative.

1. INTRODUCTION

Nowadays the major international instrument for antibiotic residues test uses the physical-chemical test method which utilizes the special response engendered by atomic mission of antibiotics molecule, but the test procedure is complex, needs long time and costs high. In recent years the gold-colloid strips for antibiotic detection with the advantages of visual, rapid, handy operation and low price is widely used. It uses the gold-colloid test technology, which should first prepare antibody by gold-colloid and then test the mini quantity unusual substance by means of ELISA (Enzyme-Linked Immunosorbent Assay) [4]. Antibiotics of different concentration will cause different amount of gold particles absorbed, then the chroma of test line (T line) on the gold-colloid strip will change correspondingly [3].

Insert the gold-colloid strip into the specimen with indicator in it, and wait for about 12 minutes. Then we could make a qualitative judgment by eyeballing way: it is negative when the chroma of test line (T line) is deeper than (or similar with) that of standard line (C line), and the other way round it is positive. Thus we think of getting the lines information of the tested gold-colloid strip and send the data to an embedded system to be analyzed so as to get the antibiotics concentration value rapidly and dependably.

2. STRUCTURES AND PRINCIPLE OF CMOS IMAGING DEVICE

CMOS image sensor is mainly made up of image sensitivity unit arrays and MOS (Metal Oxide Semiconductor) Field Effect Transistor (MOSFET) integrated circuit. Image sensitivity unit array is actually photodiode array that is core component of CMOS image sensor. The image sensitivity unit is constituted with photodiode and amplification circuit. MOSFET simulates the switch to control the out-put of signal.

The principle of photodiode bases on photovoltaic volts effect. When the incidence radiation with the flux of $\Phi_{e,\lambda}$ acts on the semiconductor PN knot that causes the levy absorption, photovoltaic cavity and photovoltaic electron are separated and move in opposite direction on the effect of PN knot's build-up electric field. So it engenders photovoltaic electric current I_Φ . viz.:

$$I_\Phi = \eta q / hv[(1-e^{-ad})\Phi_{e,\lambda}]$$

In the formula: η stands for radiation efficiency; h stands for Planck constant ($6.62\times10^{-34} J\cdot s$); v stands for optical shock frequency (s^{-1}); The current generated by diode is: $I = I_\Phi + I_D$. I_D is the underset and usually ignored for minuteness. So current generated by photodiode and incidence radiation are linear relationship:

$$I = \eta q / hv[(1-e^{-ad})\Phi_{e,\lambda}]$$

MOS switches are mainly used to control the addressing and reading out [1]. Photodiode array converts the optical image imaging in itself to an electric image engendered by electric charge density distribution. In accordance with the driving pulse the electric image could be changed into video signal and putted out row by row under certain rules.

As shown in Fig. 1., signal will be sent to the line through simulation switches ($S_{i,j}$) which have already been connected in order under control of Y direction address decoder. Then it will be transmitted to the amplifier under control of X direction address decoder. Because of having set both switch of row and line the strobe of which is controlled by the data added on the X and Y direction address decoder, it can use two directions shift register of X,Y as the working form to output signal by scanning mode or single scanning mode. Analog signals outputted from the amplifier are sent to the A/D conversion and then they are outputted as digital signals after being processed by the pre-processing interface circuit. Scheduling signal generator

in the figure offers all kinds of working plus that can be controlled by the synchronous control signal made in the interface circuits for the entire CMOS Image sensors [7].

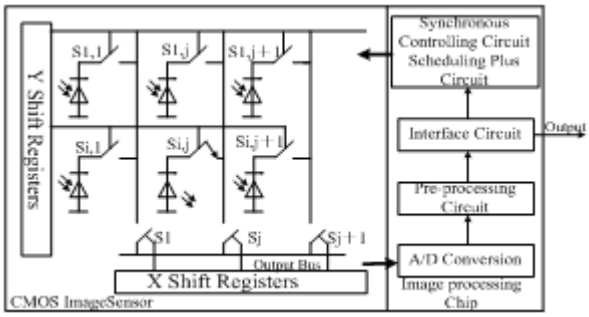


Fig1 Principle sketch map of CMOS imaging device

3. SYSTEM PLANTFORM

The rapid embedded quantitative test system for antibiotics mainly consists of two parts: embedded processing system and data collection system.

The embedded processing system is the core of the entire system using the Samsung S3C2410 processor based on the ARM920T core. The ARM920T core has five lines, Harvard structure, independent MMU, 16K instruction Cache, 16K data Cache, and AMBA bus interface, providing 1.1MIPS/MHz performance. So it is a high-performance and low loss hard macro module. On the base of rich resources the system expands a number of functional modules, such as the 64M NAND Flash, 16M memory, LCD controller, touch screen interface, two USB interface controller, and an expanding network interface by the Ethernet control chip DM9000E to connect with other quantitative detection equipment. The embedded processing system is in charge of sampling equipment-driven, image data storage and analysis, LCD touch screen driven and the communication between people and the system [2].

HY7131E (CMOS image sensor) and 301P (high-performance cameras to USB interface control chip) constitute the data collection system. HY7131E is used to collect the information of the color of the gold-colloid strip and its work patterns are: output resolution of 640 x 480, 24 bits RGB image data format; 301P is responsible of the storage, compression, transmission of the image [6]. Following diagram is the overall structure of the system:

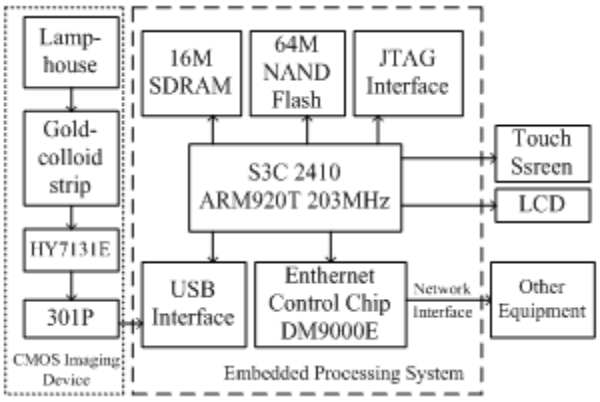


Fig 2 Diagram of the overall structure of the system

The process is that firstly 301P receives control signals through USB interface from S3C2410 processor and sets the work pattern of HY7131E by SCCB (Serial Camera Control Bus) interface; secondly, receiving control signal from S3C2410 HY7131E collects the color information of the gold-colloid strip; thirdly, 301P processes the sampled image signals by the A/D conversion, image compressor, processing circuit. And then it sends the image data which is saved as a file finally to inner buffer of the S3C2410 through the within USB controller and the external USB transceiver. At last the image processing program transplanted in this system platform will processes and analyzes the data to make a quantitative judgment.

4. EMBEDDED LINUX SYSTEM

This system is developed in Linux. So it needs to transplant the Linux system into ARM first. Video4Linux, which is compiled in the system core as mode of build-in or module is a kind of supporting document of embedded Linux for video image equipment such as image sensor. It contains a series of API functions and images, voice card-driven. It uses ioctrl () to send command to the driver and modify the parameter sent in to set equipment.

Programming by Video4Linux API functions the following operations are needed to collect the image data [2]:

1) Open and set the equipment:

```
fd = open (szDevName, O_RDWR);
```

2) Enquire the information of image buffer and then set it:

```
if (ioctl (fd, VIDIOCGMBUF, &vid_mbuf) == -1) {
    struct video_window vid_win; m_MapSize = 0 ;}
 else {m_MapSize = vid_mbuf.size; m_BuffSize =
  m_MapSize ;}
  // VIDIOCSFBUF is an API function of
  Video4Linux that sets the parameters of image buffer
  for the collection card.
```

3) Enquire the information of image intercepted window and then set it:

```
if (ioctl (fd, VIDIOCGWIN, &vid_win) != -1) {
    vid_win.width = width; vid_win.height= height;
  ioctl (fd, VIDIOCSWIN, &vid_win) ;}
 // VIDIOCGWIN obtains current information and
    VIDIOCSWIN sets new values.
```

Image interception sets mainly the size, brightness, color, contrast gradient, chroma, and depth and so on.

4) Enquire the channel information and set:

```
if (ioctl (fd, VIDIOCGCHAN, &vid_chnl) == -1)
   {perror ("ioctl (VIDIOCGCHAN)") ;}
else    {vid_chnl.channel = channel;
     if (ioctl (fd, VIDIOCSCHAN, &vid_chnl) == -1)
       {perror ("ioctl (VIDIOCSCHAN)") ;}
  // Many video-equipments can capture data from one
  or more channels; each channel can be enquired by
  parameter VIDIOCGCHAN and set by parameter
  VIDIOCSCHAN.
```

5) Obtain the image and save it in the buffer:

Steps presented above are series of setting for getting an image, and the image can be saved through mmap interface, which allows mapping equipment memory to the address space of user program directly. To use the mmap to mapping, firstly it is necessary to set the information of frame that need to be mapped, including the size of a frame, the maximum amount of supported frame, the excursion relative to the base address of each frame; And then using the function VIDIOCGMBUF to send the size of buffer and the excursion of each image to mmap. Transferring mmap could make the memory mapping coming true, which means

mapping the appointed file or part of an object to the memory. In this way it is possible to access the file or object directly by accessing the appointed memory area.
The following are some of the procedures code about MMAP:

```
vid_mmap.format = VIDEO_PALETTE_RGB565;
vid_mmap.width = m_Width;
vid_mmap.height= m_Height;
if (ioctl (fdVideo, VIDIOCMCAPTURE, &vid_mmap) == -1)
 {perror ("ioctl (VIDIOCMCAPTURE)");return (0);}
 if (ioctl (fdVideo, VIDIOCSYNC, &vid_mmap) == -1)
   {perror ("ioctl (VIDIOCSYNC)"); return (0) ;}
```

5. IMAGE PROCESSING

Image processing procedures of this system adopted following methods:

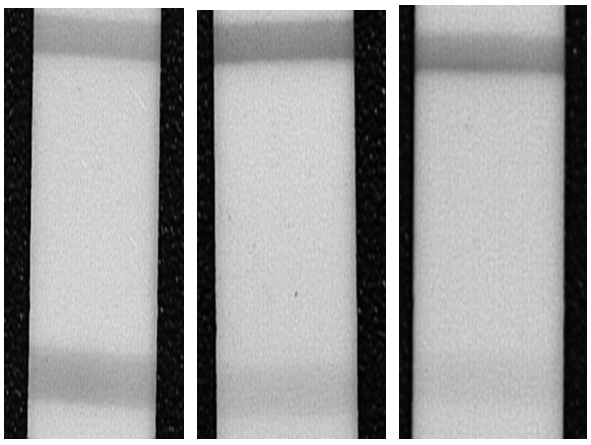
First, the procedure uses mean filter to get rid of the image noise protecting the edge information in the mean time.

Second, it gets the color information about T line and C line. For the chroma of T line becomes lighter as the concentration being stronger, it'll be harder to get the information of T line. Firstly we use the edge detection to find the location of C line and then decide the place of T line by pulsing the offset. Finally it calculates the gray value of T line and C line.

Finally, it ascertains the exact content of antibiotic residues by calculating the ratio of T line to C line and comparing the result to the value saved in the standard database.

6. INITIAL EXPERIMENT

Chloramphenicol is a widely used antibiotic beforetime. And now it has been forbidden for its serious side-effect. We choose it as the specimen, and confect some standard solution of different concentration [5], such as: 0ppb, 0.1ppb, 0.4ppb, 1.0ppb, 1.5ppb, 2.0ppb, 2.5ppb, 3.0ppb, 3.5ppb, and 4.0ppb. And then put the gold-colloid strip made by Dalian Practical Company into the solution for about 12 minutes, then put it into the embedded quantitative test system to ascertain the content of chloramphenicol. Fig 3 shows some of the images captured by CMOS imaging device. Table 1 shows the data got from the Image processing procedures:



0.1ppb 1.5ppb 2.5ppb

Fig. 3. Images collected by CMOS imaging device

Table 1 Data got from the image processing procedures

Concentration（ppb）	0	0.1	0.4	1.0	1.5
Ratio of T line to C-line（T/C）	0.735	0.825	0.912	0.989	1.072

Concentration（ppb）	2.0	2.5	3.0	3.5	4.0
Ratio of T line to C-line（T/C）	1.105	1.164	1.197	1.200	1.220

7. CONCLUSION

The rapid embedded quantitative test system for antibiotics uses the gold-colloid testing technology to detect mini quantity unusual substances and collecting data by the CMOS image sensor, then analyzes the data through the embedded system, which uses hardware and software integrated approach to enhanced system stability and reliability. It lowers the costs at the same time. So the system provides a high practical value.

8. References：

[1] Q.Y. Wang, Image Sensor Applied Technology, Beijing Publishing House of Electronics Industry, 2003.

[2] S3C2410 32-Bit RISC Microprocessor [Electronic Literature], SumSung Electronic, 2002.

[3] Y.P. Chen, M. Du, "Research of Strip Quantitative Test Based on Photoelectric Technique", Apparatus and Instrument Transaction, 2002.

[4] D.Y. Deng, S.B. Yu et al, "Testing of Antibiotic Residue in Milk", Food Security and Testing, 2004.

[5] M.F. Shen, W.Y. Zhao, et al, "Testing the Chloramphenicol Residues of the Aquatic Products by Meilian Immune Adsorption Analysis Technology", JOURNAL OF NANJING NORMAL NIVERSITY (ENGINEERING AND TECHNOLOGY), 2003, 3(2).

[6] H.W. Ma, J.H. Zhang, et al, "Real-Time Image Collecting System Based on 32 Bits Embedded System", Application of Electronic Technology, 2003(4).

[7] M.Q. Zhou, "The Interface Technology of CMOS Image Sensor in Embedded System", Development of Microcomputer, 2005, (10).

Body Pressure Measurement System Simulated By LabVIEW

Jian-xiong Lu, Fu-chang Zhang
School of Design, Southern Yangtze University, Wuxi Jiangsu, P. R. China 214122
Li-ming Shen, Yu-kun Hua
School of Wood Technology, Nanjing Forestry University, Nanjing Jiangsu, P. R. China 210037

Abstract

In this paper, a measurement system for pressure distribution is introduced, which is composed of hardware and software. The hardware includes sensors matrix containing 256 pressure-sensor, interface circuit, data acquisition board and computer system. The software consists of modules for data collection, storage, display and analysis. The system can measure pressure distribution between the chair and its occupant. The results simulated by LabVIEW platform can be displayed directly on vision after the data captured.

1. Introduction

A well-designed seat can make the user comfortable and efficient in working. Therefore a reasonable design for the seat and a favorable scheme of sitting posture is very important in Human-Machine Engineering. Tekscan Corp. has produced pressure measurement system for seat design, but it isn't affordable for most of the users. In this paper, based on our previous work, we propose a Body Pressure Measurement System (BPMS) using Virtual Instrument (VI). VI is a technique that the user can define and design the measurement function of instruments in accord with his own demand. Essentially, the VI is integration of the traditional instrument hardware and the emerging software technology that can generalize the function of the traditional instrument. Compared to the traditional instrument, VI has many advantages due to its intelligence, process ability, ratio of performance-price etc. The proposed system provides some methodologies for designing various seats and medical studying.

2. The Function of the BPMS

The functions of the BPMS are as follows:
Record the signal. Collect and record the signal from the press sensors and form a record file. Display the signal wave pattern and signal amplitude. Real-time analyze the signal. Print the output including wave pattern and analytical results. Manage the system. Manage the user information and record file. Check the outcome files.

3. Simulation of BMPS using LabVIEW

The method of LabVIEW programming is very different from the traditional programming method in that it is free from perplex aroused by the linear structure of traditional programming language. The LabVIEW is implemented according to the data flow instead of the sequence of the code, and therefore we can design the flow chart of multi-program. We can select the necessary control and data display object from the control module as we are designing the human-machine interface. The virtual instrument in LabVIEW is made up of front panel and flow chart. The front panel corresponds to the control board of electronic instrument, which is the graphic user interface and used to set the input and display the output. The front panel implements the interaction between man and machine, making the user only face the external forms of the computational instruments. The flow chart, the source code executing the program, in which we can set the input and output of signal data to implement collecting the signal from the virtual instrument, analyze and process the operation and control of the function.

4. Data Acquisition Board

Data acquisition is the processing of multiple electrical or electronic inputs from devices such as sensors, timers, relays, and solid-state circuits for the purpose of monitoring, analyzing and/or controlling systems and processes. Device performance, analog outputs, form factor, computer bus, connection to host, and environmental parameters are all important specifications to consider when searching for I/O modules, data acquisition. In our system, PCI-8319 Photoelectric Isolation Analog Input Interface Board is used. PCI-8319 applies to the PC computers that provide Bus Connector and its PCI-bus Plug and Play feature allows a user to install the data acquisition board without making manual system configurations. PCI-8319 Data Acquisition Board supports many operation systems, such Windows series and Unix, as well as data acquisition software package such as TestPoint or LabVIEW. It can be installed in any PCI-bus slot to play with the cable connected into the machine box.

PCI-8319 adopts triple-bus photoelectric isolation technique, which isolates the measured signal system from the electrical equipment in the computer. Match directly with various sensors, PCI-8319 can be operated at gain of 100 to amplify the signal. With its self-provided DC/DC power supply module, the board needs no external power. Each PCI-8319 board provides 32 analog input terminals. These terminals are configurable by software either as single-ended inputs or, in pairs, as differential inputs. Each single-ended or differential input is commonly referred to as an input channel. Because the signal of the pressure on seat is very small, our system adopt 16 input terminals with input signal -5~+5 and gain 1000. When the PCI board working in pair ended or differential input mode, the conversed digital of 12 bits is binary shift code with the highest bit the sign bit, where "0" denotes negative and "1" positive. The analog voltage can be given by: Analog Voltage=digital 10/4096 (V), which is 2.44mV.

The installation of the driver follows the installation of PCI-8319 board. PCI8KP.dll is a dynamic link library running on Windows 95/98/2000/NT, which is provided for the PCI8000 Series Data Acquisition Board. The function sealed by PCI8KP.dll can be called directly by other applications.

5. The Architecture of the BPMS

The architecture of the BPMS can be divided into four layers as shown in Figure 1: Test Management Layer, Test Program Layer, Instrument Driving Layer and I/O interface Layer. To use the LabVIEW platform, we can implement the four-layered software much more easily.

The Test Management Layer is an instrument-independent test management environment that produces the operation system interface, manages and executes the testing task. In BPMS, the Test Management Layer provides the human-machine interface. The function option items of the system and the control component responses to the function option and process control can be made by the user. The parameter input column can acquire the parameter input by the user. The output display window displays the signal wave pattern and the test results. Moreover, the Test Management Layer also controls the order, the branch and the loop of the test, which is implemented by the components defined by the LabVIEW. The layer can also realize the file storage, password management and multi-level access. The layer can also realize the file storage, password management and multi-level access.

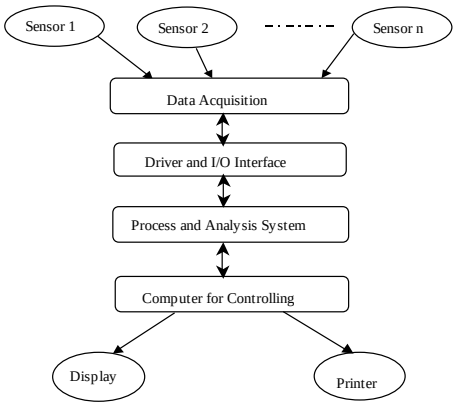


Fig. 1. Architecture of the BPMS

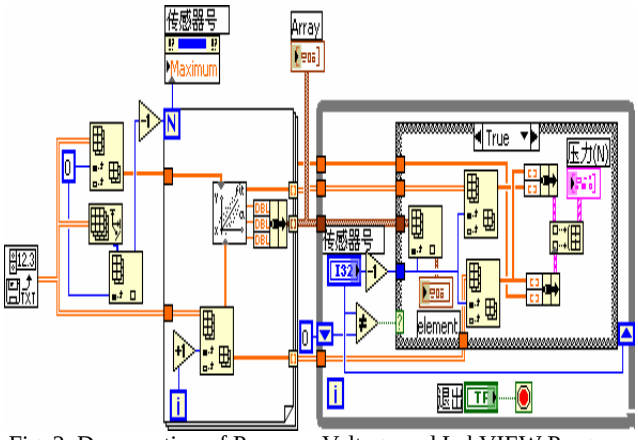


Fig. 2. Demarcation of Pressure-Voltage and LabVIEW Program for "Zero-level" deviation

In the design of the system for BPMS data collecting, in order to acquire stable data, we skipped over the first 100 data samples and began the record at the 101st value. The process of data collecting carries through until the "confirmation" button is pressed. Before the data was processed using Least-Square method, we fit the data labeled by each sensor, i.e., work out the average value of data measure by each senor and its circumjacent colleagues. Therefore we can obtain the relationship between pressure of sensor and the voltage, as well as "zero-level" deviation (see Fig. 1 and Fig.2).

The Test Program Layer implement the various test functions. LabVIEW provides many standard instrument function modules and data process modules of general purpose and instrument independence. The library functions of data collection, expression analysis and data storage is utilized in testing, and these functions are the advanced modules that construct the system. In practical measurement, it must be taken into account that the "zero-level" of the measured value will alter due to the influence by temperature and other environmental elements. However, in a relatively short time span, the change of the value is so small that is can be neglected. Thus, we can set the value to zero in the way of zero bus collection.

Inputting the disposal graph, sampling and analyzing the value from the sensor in each channel, we can obtain the results that display in the screen of the computer (see fig.4). The proposed data will be store in the database for further utility.

The Instrument Drive Layer drives and controls hardware instrument directly. This layer is the most difficult to develop in the system in past. Presently, since the drive program of the instrument is released to the user in module and in the way of instrument-independence, it can be utilized conveniently and the source code can be revised to satisfy the specific necessity. In our BPMS, the Instrument Drive Layer is mainly employed to control the Data Acquisition Board; set the sampling ratio, gain and trigger condition; separate frequency and count; set the buffer mode and buffer length; control the output. The data collection board used to develop the system by National Instruments Corporation on LabVIEW is unaffordable to domestic corporation，we must develop driver program when use 8319 data acquisition card, which is made in our own country. Our system employed LabVIEW with the dynamic linkage library transferred.

I/O interface is the communication protocol of low level that is used to process the connection between the instruments and the hardware.

6. Results and Discussions

The BPMS can help the producer of seat determine how the joint, mat, prop, spume stiffness and support structure of the seat have impacts on the pressure distribution and the comfort of the rider. It can help to optimize the option of seat material, analyze the pressure variation when the rider, optimize the posture of the rider, record the change of the rider's posture. The figure 5 shows a practical application of the BMPS.

The measured data is analyzed to determine the distribution of human-seat interface pressure under static seating environments. The results are compared with those derived from measurements performed on a 400×405mm rigid seat with flat pan and backrest inclination of zero degree. The height of the rigid seat used in the reported study was identical to that of the soft seat (420mm). The measured pressure distribution acquired under different postures are evaluated in terms of static

Fig. 3. The front panel of the pressure sensors parameter

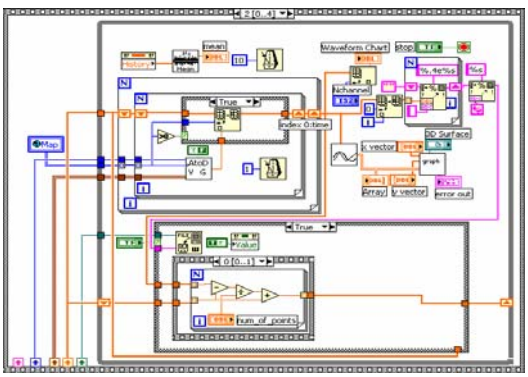


Fig. 4. The display LabVIEW program of the testing results

pressure distribution contours, maximum ischium pressure, elective contact area, and longitudinal contact force distribution curve.

Fig. 5 illustrates the typical surface plots and contour maps of the interface pressure measured at the surfaces of the rigid and the soft seats under static seating conditions (the subject is height 1.75m and weight 750N). It should be noted that the above data is derived from measurements performed with subject assuming an EBS posture, while both seats were adjusted to an identical height of 420mm. The results show that sitting on a rigid surface yields dominant pressure distribution within the ischium region with peak pressure occurring in the vicinity of the ischial tuberosities. The high interface pressure peaks observed with the rigid seat are expected to cause fatigue and discomfort over prolonged sitting. The resulting elective human-seat contact area is thus quite small and the magnitude of pressure under the thighs is relatively negligible. The interface pressure distribution obtained for the soft seat, however, is considerably different, as shown in the figure. The seated weight is more or less uniformly distributed over a considerably larger sitting area and the peak interface pressure is significantly lower than that encountered on a rigid seat. The maximum static pressure, however, appears in the vicinity of ischial tuberosities. The soft seat with low-level peak interface pressure can thus be expected to reduce driver fatigue and discomfort caused by the local concentration of interface pressure.

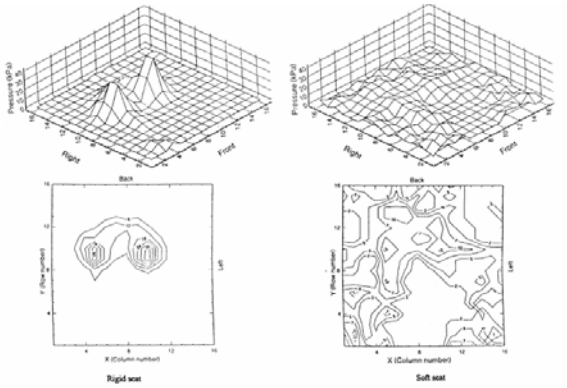


Fig. 5. Static interface pressure distribution measured on the rigid and soft seat surface

7. Conclusion

In this paper, propose a design methodology of a Body Pressure Measurement System (BMPS). The system consists of four layers and use cheap hardware such as PCI-8319 data acquisition board. The BPMS is simulated using Virtual instrument-LabVIEW software platform, which is easy to operation and has a friendly interface as well as promote efficiency and precision. The applications of virtual instrument, particularly the application of the simple programming environment of the LabVIEW, promote the development of applied systems and simplify the development process considerably. The LabVIEW can support various measurement systems, realize the functions of many instrument and has good prospect of application.

8. References

[1] LabVIEW User Manual, National Instruments Corporation, January 1998 Edition.

[2] LabVIEW Data Acquisition Basics Manual, National Instruments Corporation, January 1998 Edition.

[3] Sensing solutions for the 21st century, Tekscan, America, 2000.

[4] J.X. Lu, Body pressure measurement system based on virtual instrument and data acquisition board. The International Conference of Impulsive Dynamical and Application Proceeding 3(2005) 1906-1913.

[5] R. Padmanabhan, N.M. Zarroug, et al, A novel adaptive control system for a custom-built tension-torsion machine. Adavanced in Engineering Software. 36(2005), pp.137-146.

[6] K. N. Whitley, L.R Novick, et al, Visual Programming in the Wild: A survey of LabVIEW programmers. Journal of Visual Languages & Computing. Volume 12, Issue 4, August, 2001, pp. 435-472.

[7] K.N. Whitley, L.R Novick, et al, Evidence in favor of visual representation for the dataflow paradigm: An experiment testing LabVIEW's comprehensibility. International Journal of Human-Computer Studies. Volume 64, Issue 4, April, 2006, pp. 281-303.

[8] Jackson, E. Mark, et al, Numerical simulation of nonlinear feedback model of saccade generation circuit implemented in the LabVIEW graphical programming language. Journal of Neuroscience Methods. Volume 87, Issue 2, March 1, 1999, pp. 137-145.

SC: An Approach to Ensure Software Reliability and Reusability

Feiqiao Mao[1], Deyu Qi[2], Weiwei Lin[3]
[1, 2, 3]Department of Computer Science, South China University of Technology
Guangzhou, 510640, China
Email: [1]maofeiqiao@163.com [2]qideyu@scut.edu.cn [3]linweiwei2000@yahoo.com

ABSTRACT

Problems of software reusability and reliability deeply faze developers. Object-oriented technology does not solve the problems satisfactorily, though it can take some effect. The results from recent research in component-based development depend too much on developers' experience and commercial practice. Thus, more effective approaches are eager to be found. In this paper, first a model named Smart-Component (SC) with good reusability is proposed and its five core parts are defined, and then a formal method which contains four steps is introduced to ensure the reliability of application system integrated with SCs. Finally, an example of SCs integration is provided to show how to ensue the reliability of SC-based application system using this formal method.

Keywords: reliability, reusability, component, formal method, formal synthesis.

1. INTRODUCTION

The problems of software reusability and reliability deeply faze developers. Object-oriented technology can take effect in a certain extent, but it cannot solve the problems satisfactorily. Thus, more effective approaches are eager to be found in this domain. To solve the problems, recent researchers especially focus on component-based development technologies combined with agent, specification or other technologies.

Reference [1] discussed a structured data object-based method for simplifying the complex information system development and evolution, proposed a structured data object-based, agent component-oriented framework. In its application instance, the core structured data object-based agent component is a Delphi TDBGrid control encapsulated by the published interfaces.

In addition, M. Brian Blank introduced a rule-driven coordination agent [2], which is a middle agent and encapsulates the interaction policy rules, the aspects of communication, data management, and policy execution. So it could realize asynchronous communication and ensure autonomy. In its typical electronic commerce domain application, the operational details revealed that the rule-driven coordination agent implemented the management of java-based component flexibly and dynamically.

And another group of researchers, Michelle Lasagna and Margaret Lyell in reference [3] compared two component frameworks: the FIPA-compliant multi-agent system and the web-centric J2EE platform. The two frameworks are based on specification and they mandate platform responsibilities towards their respective components. The multi-agent system is more open and flexible than the J2EE platform. However, the J2EE platform has better performance, more reusability and greater reliability than the multi-agent system because there are a number of commercial implementations of the J2EE specifications and the J2EE framework specifications offer more support to the application developer, especially one attempting to reuse components.

There is also other research of component-based development using component specification techniques, such as references [4,5,6,7,8], in which different topics about specification techniques are discussed.

Generally, the aforementioned research results make some improvement or offer some recommendations on component reuse or software reliability. However, their common pitfalls are that they depend so much on developers' experience and commercial practice when they are used for successful software development. Accordingly, a model of SC with good reusability, and a formal method for SCs integration which should ensure software reliability in higher extent are presented in this paper. After the definition of SC and the description of the prototype of SC model, a formal method and its example for SCs integration are described in detail.

2. SC MODEL

2.1 SC definition

SC is what we introduce to solve the problem of software reusability on the level of component. Our research of reuse is different with other research of reuse which is on the level of software architecture. Here is the definition of SC.

Definition 1 (SC) SC is a smart software component. SC comprises five parts: sensor, mandatory, ability, decision-making and action. SC can be tested and deployed independently. SC runs autonomously. Customized SCs can cooperate and aggregate to be integrated as an application system automatically.

The following section 2.2 will deeply explain each part of SC in its definition.

2.2 The SC model

Fig. 1 shows the prototype of a SC model. The five parts of a SC are defined as follows.

Sensor Sensor is the entrance of a SC. It senses and receives the four kinds of information elements from the outside of SC: *events*, *messages*, *conditions* and *data*. The explanation of each information element is described as follows.

> Note: *Events* element, *messages* element, *conditions* element and *data* element are contained in a package, which is sent from one SC to another.
>
> *Events* element describes the state of the occurrences that are activated by the outside

behavior, such as the location of the mouse, the state of the mouse buttons and the state of the keyboard keys. The attributes of *events* element include name, id, type, value and modifiable. *Events* element can be described with XML. For example:

<event name="mouseLocation" id="SC003e01" type="mouseEvent" value.x="100" value.y="50" modifiable="1"/>

It describes an event which is activated when mouse moves at (100, 50) on the screen.

Messages element describes the basic information of requests or responses. The attributes of *Messages* element include sender, receiver, id, name and type. *Messages* element is also described with XML. For example:

<message sender="SC001" receiver="SC112" id="SC001m000" name="priceQuery" type="SOAP"/>

It describes the basic information of a SOAP message which is sent from SC001 to SC002.

Conditions element describes the state of condition attached to an action. The attributes of *conditions* element include name, id, action, state and modifiable. *Condition* element is also described with XML. For example:

<condition name="fileFind" id="SC002c001" action="deleteFile" state="1" modifiable="0"/>

It describes the state of the condition attached to an action to delete a file. The state of "fileFind" condition means that the file has been found and "deleteFile" action can be executed.

Data element describes the basic data information sent from one SC to another. The attributes of *data* element include name, class, type, value and modifiable. *Data* element is also described with XML. For example:

<data name="ticketPrice" class="simple" type="float" value="550.00" modifiable="0"/>

It describes the basic data information of ticket price.

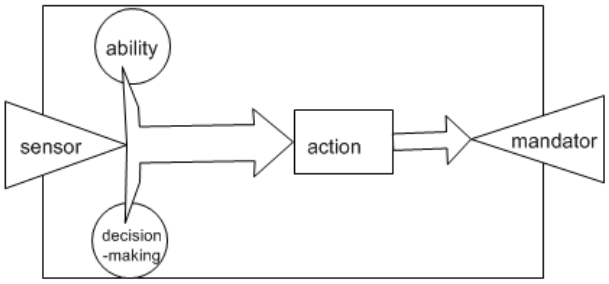


Fig. 1. The prototype of SC model

Mandatory Mandatory is the exit of a SC. It sends the four kinds of information elements produced by or passing by SC: *events*, *messages*, *conditions* and *data*. All of the results produced by or information passing by SC should be sent out from it.

Ability Ability is the repository of a SC. It stores the functions which SC can perform. It also stores relevant conditions which must be provided to perform a function. Ability assists decision-making to make decision.

Decision-making Decision-making is to decide which action should be taken according to the information from sensor and the knowledge stored in ability.

Action Action is the execution part of a SC. Relative actions which are decided by decision-making to perform a function should be done in action using the information from sensor.

2.3 Distinctions among object, component, agent, web service and SC

We address that SC has advantages over object and component in supporting software reuse; SC also provides the basis for our formal method for their integration into reliable software. Maybe, just from the definition of SC, it is hard to recognize its advantages. In order to more clearly support our argument it is necessary to compare SC with object and component in the following aspects.

Before starting, we first discriminate SC from agent, because from the definition of SC, it seems to be agent. SC is like agent as we adopt some technologies used for agent constructing to construct SC, such as decision-making and function rules, and some concepts from agent when we describe SC, such as sensor part and decision-making part. But SC is not so intelligent and complex as agent in AI because it has no special part such as satisfaction grade part, predicting part and past-environment-information reservation part, which are owned by agent in AI. SC is also different with agent in agent-based development, because those agents are agent-based component which commonly encapsulate ActiveX controls or components with self-defining wrap, and they have no decision making part, sensor part and mandatory part in them, and therefore there must be policy component in agent-based development to make decision. But special policy SC is not needed in SC-based development

Abstraction In object, data representations of entity and their associated primitive operations are abstracted. While component abstracts not only what object abstracts but also interface, context, contract and method. However, SC abstracts even policy and function rules, which object and component do not abstract, as decision-making part and ability part in it. That makes SC more flexible than object and component in reuse.

Encapsulation Object encapsulates its state and action, mainly like variables, data structures and operating functions, as class, which cannot be deployed independently. And its state is visible from outside. While component encapsulates its state and action, mainly like process and unchangeable primitive objects as well as classes, as a box which can be deployed independently. And its state is not visible from outside. SC encapsulates not only its state and action but also policy and knowledge of functions, which can be performed. And SC's state is also not visible from outside. That makes SC more adaptive, and reusable as well as changeable.

Separation of Policy and Implementation Object and component both separate their policy and implementation, while SC does not. Policy of object is distributed in many classes in a frame. Policy of component is centralized in some special components as policy components. SC contains policy and implementation; and separates them into two parts in itself. That makes it more flexible.

Coupling and Cohesion In coupling grade, SC is low, component is high, and object is higher. Because of inheritance there are complex relationships between

objects, but there is no inheritance in component and SC. SC has lower coupling than component, because SC can run autonomously but component cannot do and SC has no remote method invoke from each other but cooperatively performs task through communication. In cohesion grade, object is low, component is high, and SC is higher. Object's state is visible from outside. And some of its encapsulated public elements such as variables and functions can be directly visited from other objects. Those make object has low cohesion. Component's state is invisible from outside. Its encapsulated elements such as method and process only can be indirectly visited via its interface. Those make component has higher cohesion than object. SC's state is also invisible form outside. SC has the best information hiding compared with object and component. Nothing of its encapsulated elements such as states and actions as well as policy can be visited directly or indirectly from other SC, but through sensor and mandatory to communicate. Those make SC has higher cohesion than component.

Separation of Concerns Different or unrelated responsibilities should be separated in software system. In component-based development, they can be put into different components. But in anfractuous object-based system it is difficult to separate concerns. We adopt separation of concerns technology constructing SC, related or close responsibilities are implemented as different functions in a SC, but different or unrelated responsibilities are implemented separately in different SCs.

We choose the above aspects when object, component and SC are compared, because they are closely related and they are important technologies in building software system, especially in enhancing reusability and changeability.

At last, we would like to refer that SC is also different with web service, which is much more distributed than SC and depends on Internet a lot. In addition, our SC model is different with the object model based on UML in literature [9], which is also proposed for reuse and formal synthesis as well as formal verification.

3. THE FORMAL METHOD AND EXAMPLE

Based on SC defined above, we propose a formal method for the process of SCs integration to ensure the reliability of an application system, which is integrated with SCs.

3.1 The main idea of the formal method

An SC-based application system is an assembly of SCs in essence. Software development based on SCs is result oriented and it depends little on code, developers' experience and commercial practice. Developers only need to concentrate on the *key result points* passing by each SC. An SC-based application system is considered as a sequence of many *key result points* logically. Not only simple SC-based application systems but also complex SC-based systems can be abstracted and formalized as a sequence. So the operations and deductions on sequence in the formal method can provide an effective way for developers to develop reliable application systems. In order to describe this formal method, some necessary definitions are provided below.

Definition 2 (***key result point***) The aforementioned information elements including *events*, *messages*, *conditions* and *data* existing when an SC-based application system runs can be classified as the incoming information elements (into a SC), the outgoing information elements (out of a SC) and the inside information elements (in a SC). *Key result point* (denoted by *Xi*) denotes the incoming information elements or the outgoing information elements of a SC.

The definition of the concept of *key result point* is very important because it is the basis of our formal method.

Definition 3(***sequence result***) Sequence result (denoted by *S*) denotes a sequence of *key result points*, which is formed according to generating sequence of *key result points*. Logically and formally, sequence result stands for an application system integrated with SCs. Because the whole *key result points* in an SC-based application system can be considered as a serial sequence or a parallel sequence (both denoted by *S*) according to their generating sequence.

Definition 4 (***result-oriented***) the concept of *result-oriented* denotes that the development based on SC is result oriented, and its comparative concepts are *object-oriented* and *component-oriented*. Based on *key result points*, a SC is a producer or consumer of *key result point*, application system is the assembly of the SCs and it is also the sequence of the *key result points*. That is result oriented.

Based on the above definitions, a SC-based application system can be formalized as follows.

(1) Finding and formalizing *key result points.*

This step is very important because in this step we provide two formal forms of each *key result points.* The first formal form is the basis of our formal method, which guides developers to integrate SCs, while the second formal form helps developers to validate software.

The first formal form: The concentrated information elements of each *key result point* should be formalized like $X_i : \langle X_i \rangle$.

The second formal form: Each *key result point* should be further formalized like

$$X_i : (event_{xi}, data_{xi}, message_{xi}, condition_{xi}).$$

The purpose of the further formalization of each key result point is to help developers to check the correctness of the output information elements of a SC and the correctness of a system integrated with SCs. Because the formal expression form can help developers to instance the key result point of a SC. If a *key result point* in a sequence is instanced, the instances of other *key result points* in the sequence can be deduced using the theorems and corollaries in our formal method.

(2) Using generated sequence of *key result points* to formalize SCs integration process for a SC-based application system.

A SC-based application system which implemented some specific functions using SCs, then the application system S should be formalized like this:

$S = \langle X_i \rangle^{\wedge} \langle X_{i+1} \rangle = \langle X_i, X_{i+1} \rangle$, that is $S : \langle X_i, X_{i+1} \rangle$.

(Note: '^' is an operator, $\langle X_i \rangle^{\wedge} \langle X_{i+1} \rangle$ denotes the concatenation of $\langle X_i \rangle$ and $\langle X_{i+1} \rangle$.

The expressions of $S : \langle X_i, X_{i+1} \rangle$ or $S = \langle X_i, X_{i+1} \rangle$ denotes the sequence *S* containing *key result points* X_i and X_{i+1}. And it also means that the

application system *S* contains $SC_{i \text{ and } SC_{i+1}}$.)
The purpose of this formal step is to logically keep the gradual integration from two to three; four… up to the whole of SCs of a system correct.
(3) Simplification and deduction.
Here, we introduce two operators: Head and Tail. If $S=\langle X_1, X_2, X_3\rangle$, then $HeadS=\langle X_1\rangle$, $TailS=\langle X_2, X_3\rangle$. They mean that in the sequence *S* X_1 happens beforehand, X_2, X_3 happens subsequently.

Theorem 1 If $HeadS=\langle X_1\rangle$ and $TailS=\langle X_2, X_3,\ldots, X_n\rangle$, then $S = HeadS\hat{}TailS=\langle X_1, X_2, X_3,\ldots, X_n\rangle$.

Explanation: Theorem1 is used for deduction when guiding dependable integration. It guides how to correctly integrate SC_1 (which produces X_1) into a system (which has been integrated with $SC_2, SC_3,\ldots, SC_n$ which separately produce $X_2, X_3\ldots, X_n$.), logically compared with which, SC_1 always generate result beforehand. And it is also useful in the process of validating. For example, a given built system $S=\langle X_1, X_2, X_3,\ldots, X_n\rangle$, if SC_8 (which produces X_8) has been found something wrong with it and the system *S* works abnormally, we need to find whether only SC_8 is abnormal. After the test of the subsystem $S_1=\langle X_9, X_{10}, X_{11},\ldots, X_n\rangle$ and the subsystem $S_2=\langle X_1, X_2,\ldots, X_7\rangle$, it is found that both of them are normal, we can figure that only SC_8 is abnormal.
Proof of theorem 1:
Because $HeadS=\langle X_1\rangle$ and $TailS=\langle X_2, X_3,\ldots, X_n\rangle$, according to the definition of the operator '^', it can be deduced as $\langle X_1\rangle\hat{}\langle X_2, X_3,\ldots, X_n\rangle=\langle X_1, X_2, X_3,\ldots, X_n\rangle$, that is $S = HeadS\hat{}TailS=\langle X_1, X_2, X_3,\ldots, X_n\rangle$.

Corollary 1 Suppose *S* is finite, and $S=\langle X_1, X_2,\ldots, X_n\rangle$, then $Tail^{n-1}S=\langle X_n\rangle$. Notice that $Tail^{n-1}S=\underbrace{TailTail\ldots TailS}_{n-1}$.

Proof of corollary 1: The method of induction principle can be used to prove corollary 1, so the proof process is omitted here.

Theorem 2 If the variable *T* is introduced to denote the actual time when X_i happens, note as (T_i, X_i), then $S=\langle(T_1, X_1),(T_2, X_2),\ldots,(T_n, X_n)\rangle$.

Explanation: We provide theorem 2 because it is necessary to introduce the time when *key result point* is produced as a variable element, only sequence without time element it is difficult to discriminate parallel and intercurrent when our method is used to guide integration.
Proof of theorem 2:
In the sequence $S=\langle X_1, X_2,\ldots, X_n\rangle$, if T_i denotes the actual time when *key result point* happens, the element X_i in the sequence should be expressed as (T_i, X_i), so the sequence *S* comes into a new sequence like $S=\langle(T_1, X_1),(T_2, X_2),\ldots,(T_n, X_n)\rangle$. That is obvious.

Theorem 3 If $\begin{cases} HeadSi = HeadSj \\ Tail^m Si = Tail^n Sj \end{cases}$, then $Si / HeadSi / Tail^m Si$ and $Sj / HeadSj / Tail^n Sj$ are allowed to occur in parallel.
(Notice: "/" is an operator, Si / Sj denotes that if $Sj=\langle(T_1, X_1),(T_2, X_2)\rangle$, $Si=\langle(T_1, X_1),(T_2, X_2),(T_3, X_3),(T_4, X_4),(T_5, X_5)\rangle$ and $Sj \subseteq Si$ is necessary, then $Si / Sj=\langle(T_3, X_3),(T_4, X_4),(T_5, X_5)\rangle$.)
Explanation: Sequence is good at express serial relationship, but it is difficult to directly express parallel relationships with sequence. Theorem 3 is provided for software designers to find parallel relationships in SCs' behavior, so it guides reliable integration from another point of view.
Proof of theorem 3: Here, we give an example to prove the conclusion. For example, suppose a system *S* has two subsystems (*S1* and *S2*), which are separately formalized as $S_1=\langle(T_1^1, X_1^1),(T_2^1, X_2^1),(T_3^1, X_3^1)\rangle$ and $S_2=\langle(T_1^2, X_1^2),(T_2^2, X_2^2),(T_3^2, X_3^2),(T_4^2, X_4^2)\rangle$, and $\begin{cases} HeadS_1 = HeadS_2 \\ Tail^2 S_1 = Tail^3 S_2 \end{cases}$. Then, we can deduce it as $\begin{cases}(T_1^1, X_1^1)=(T_1^2, X_1^2) \\ (T_3^1, X_3^1)=(T_4^2, X_4^2)\end{cases}$. It means that X_1^1 and X_1^2 occurs at the same time when system *S* is running, and their four information elements are the same (That is $(event_{X_1^1}, data_{X_1^1}, message_{X_1^1}, condition_{X_1^1})$ and $(event_{X_1^2}, data_{X_1^2}, message_{X_1^2}, condition_{X_1^2})$ are the same.) , and X_3^1 and X_4^2 also occurs at the same time, and their four information elements are also the same (That is $(event_{X_3^1}, data_{X_3^1}, message_{X_3^1}, condition_{X_3^1})$ and $(event_{X_4^2}, data_{X_4^2}, message_{X_4^2}, condition_{X_4^2})$ are the same.). So we can figure that the rest (which is X_2^1)of *key result points* in S_1 and the rest (which are X_2^2, X_3^2)of *key result points* in S_2 occur in the interval of a same short duration, and the time of T_2^1 and $T_2^2 - T_3^2$ is within the interval. So $S_1 / HeadS_1 / Tail^2 S1$ and $S_2 / HeadS_2 / Tail^3 S_2$ occur in the interval of a same short duration. So we conclude that they are allowed to occur in parallel.
The above proof of theorem 3 just using an example seems not rigid and formal as well as persuadable, a more rigid proof should be provided. That is one part of our future work.

3.2 Example

In order to show how to use the above formal method to develop reliable application systems based on SCs, an example is provided as follows.
Fig. 2 shows an example of SCs integration for a simple

application system. In the example, four SCs: SCa, SCb, SCc and SCd in the dash rectangle frame are integrated to a simple application system. The arrows direct branches and flow of information elements passing through each SCs. The direction of arrow also implies the producing order of information elements. *C, R, X1, X2, Y1* and *Y2* are vectors of information elements. They denote respectively four kinds of information elements: *events, messages, conditions* and *data*, which pass between different SCs. *C* represents information elements which flow into SCa and it is also the entrance of the simple application system. *R* represents information elements which flow out of SCd and it is also the exit of the simple application system. *X1* represents information elements, which flow out of SCa, and into SCc. *X2* represents information elements which flow out of SCc and into SCd. *Y1* represents information elements, which flow out of SCa, and into SCb. *Y2* represents information elements which flow out of SCb and into SCd.

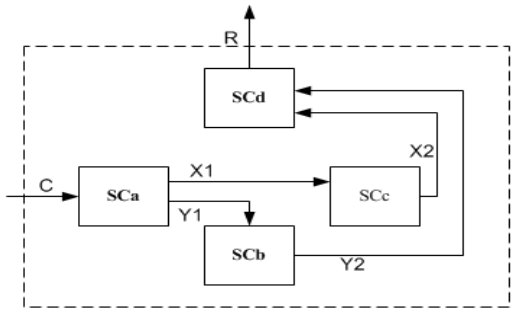


Fig. 2. SCs integration for a simple application system

The following steps show how to use the above formal method to formalize the four SCs integration for a simple application system.

(1) Finding and formalizing *key result points*.

Six *key result points* are found in the simple application system. They are *C, X1, X2, R, Y1* and *Y2*. According to the first step in the formal method, they are formalized as follows:

$$\langle C\rangle, \langle X1\rangle, \langle X2\rangle, \langle R\rangle, \langle Y1\rangle, \ \langle Y2\rangle \qquad (1)$$

Key result point C is further formalized like this:

$$(event_c, data_c, message_c, condition_c) \qquad (2).$$

This further formal form will be used for validating.

The further formalizing of the other *key result points* is omitted here.

(2) Using generated sequence of *key result points* to formalize the four SCs integration process for the simple SC-based application system.

According to the second step in the formal method, the simple application system S is formalized like this:

$$S:\begin{cases} S1 = \langle C\rangle^{\wedge}\langle X1\rangle^{\wedge}\langle X2\rangle^{\wedge}\langle R\rangle = \langle C, X1, X2, R\rangle \\ S2 = \langle C\rangle^{\wedge}\langle Y1\rangle^{\wedge}\langle Y2\rangle^{\wedge}\langle R\rangle = \langle C, Y1, Y2, R\rangle \end{cases} \qquad (3)$$

There are mainly two flows of information elements in the aggregation process of SCs in the example, so the simple application system *S* is formalized originally with two sequences: *S1* and *S2*.

(3) Simplification and deduction on sequence *S*.

According to theorems and corollaries of the formal method in the third step, the sequence *S*, comprising *S1* and *S2*, is deduced as follows:

$$S:\begin{cases} S1 = \langle (T_1^1, C), (T_2^1, X1), (T_3^1, X2), (T_4^1, R)\rangle \\ S2 = \langle (T_1^2, C), (T_2^2, Y1), (T_3^2, Y2), (T_4^2, R)\rangle \end{cases} \ldots The.2 \ (4)$$

$$\begin{cases} HeadS1 = \langle T_1^1, C\rangle, HeadS2 = (T_1^2, C) \\ Tail^3S1 = (T_4^2, R), Tail^3S2 = \langle T_4^2, R\rangle \end{cases} \ldots The.1, Cor.1 \ (5)$$

$$\begin{cases} S1/HeadS1/Tail^3S1 = \langle (T_2^1, X1), (T_3^1, X2)\rangle \\ S2/HeadS2/Tail^3S2 = \langle (T_2^2, Y1), (T_3^2, Y2)\rangle \end{cases} \ldots The.3 \ (6)$$

In the example illustrated in Fig. 2, it is obvious that

$$\begin{cases} HeadS1 = HeadS2 \\ Tail^3S1 = Tail^3S2 \end{cases} \qquad (7),$$

so from theorem3 it can be deduced that $\langle (T_2^1, X1), (T_3^1, X2)\rangle$ and $\langle (T_2^2, Y1), (T_3^2, Y2)\rangle$ are able to occur in parallel, that is, SCb and SCc can run in parallel.

(4) Useful conclusions for developers.

According to the formal process of the integration of SCs for the simple application system, some useful conclusions below are obtained and they can guide developers to develop this simple SC-based application system reliably.

SCb and SCc can be developed to work in parallel.

In order to ensure the correctness of the simple application system, process of the four SCs must keep the occurrence sequence of the six *key result points*.

Using the further formal form of *key result point* in step (1) in the example, an instance of k*ey result point C* will be deduced to an instance of *key result point R* after being functioned by SCb, SCc and SCd. Developers can compare the state of the information elements of *R*'s instance with the requirement of the simple application system to evaluate whether the result produced by the system accord with the requirement in a extent from another point of view.

3.3 Advantages of the formal method

We conclude that the formal method has some advantages.

(1) The form of formal expressions is consistent. Because after formalized with the formal method, a *key result point*, a SC, a sub-system and a system all are in the form of sequences. This make that the method has the characteristic of good expansibility as well as convenience for guiding reliable integration.

(2) The method is easy to understand and master. The core formal form in the method is sequence, so it is distinguish with other formal method, such as B method, VDM method and Z method, which have complex and various formal form and expressions.

(3) The method has good ability to express relationships. Because serial relations and intercurrent relations of *key result points*, sub-system and system are directly expressed in the sequence, and their parallel relations also can be deduced using theorems and corollaries in the method.

(4) The method is convenient to be used to analyze the correctness of SC-based system precisely, because each *key result point* can be instanced when software developers try to validate or test a system, using the

second formal form of each *key result point* in step 1 of the method.

4. CONCLUSIONS

This paper proposes a model of SC with better reusability compared with object and component even agent and it also proposes a formal method for reliable software development based on SC, especially for reliable integration and validating.

From the above formal example, it is found that if all *key result points* are correct(Here, correctness of a *key result point* means that the state of the four information elements of a *key result point* produced by a SC (That is *key result point* out of a SC in definition2.)accord with what they should be, or the state of the four information elements of a *key result point* into a SC accord with what they should be.), the above formalizing process can ensure the sequence of *key result points* are correct, that is, the integration of SCs for the simple application system can be ensured to be correct. There are also the deductions in the formalizing process, which help to ensure the correctness of the system. However, *key result point* is produced by SC, so its correctness depends on the reliability of SC. SCs are smart components, they can be tested, deployed independently and they can run autonomously. Hence, it is easy to ensure their correctness successfully. Therefore, the formalizing process is an effective approach to ensure the reliability of software, which ensures the reliability on a rigorous abstract level and reduce risk in a higher extent for software development based on SC.

On the other hand, all SCs have the same sensor, mandator and expression format of their *key result points*, so they can be integrated in different ways to form different SC-based application systems without any change. Therefore, SCs have better reusability.

5. FUTURE DIRECTIONS

We would like to pursue further research in several directions: (1) Completing the formalizing system to be more rigid and more expressive, including increasing definitions, theorems and corollaries etc. (2) Constructing an SC protocol stack for SCs communication. (3) Specifying the information elements format of *key result point*, including *messages* expression format, *events* expression format, *data* expression format and *conditions* expression format. (4) Building a prototype software development environment based on SC. (5) Developing an application case using the formal method and SCs with the platform.

ACKNOWLEDGEMENT

Thanks to Jim Woodcock and Jim Davies, because the authors get many sparks from their work [10] when the formal method for SCs integration is proposed. In addition, thanks to ZhengPing Qian who has provided lots of good suggestions for the revision of this paper.

REFERENCES

[1] Y.T. Qu, B.Y. Jin, et al., A Structured Data Object Based, Agent Component Oriented Approach to Software Development. Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on Volume 1, 18-21 Aug. 2005, 270 – 275.

[2] M.B. Blake, Rule-driven coordination agents: self-configurable agent architecture for distributed control, Autonomous Decentralized Systems. 2001. Proceedings of 5th International Symposium on 26-28 March 2001,271 – 277.

[3] M. Casagni, M. Lyell, Comparison of two component frameworks: the FIPA-compliant multi-agent system and the web-centric J2EE platform. Software Engineering, 2003. Proceedings of 25th International Conference on 3-10 May 2003, 341 – 351.

[4] P. Brada, Component revision identification based on idl/adl component specification. Proceedings of the 8th European software engineering conference held jointly with 9th ACM international symposium on Foundations of software engineering, pages 297-298, New York, USA, 2001. ACM Press.

[5] J. Han, An approach to software component specification. In Proceedings of 1999 International Workshop on Component Based Software Engineering, May 1999.

[6] Y. Liu, H. Conrad Cunningham, Software component specification using design by contract. In Proceedings of the South-East Asia Software Engineering Conference. National Defense Industry Association, Tennessee Valley Chapter, April 2002.

[7] F. Plasil, Enhancing component specification by behavior description - the SOFA experience. In Proceedings of the 4th Symposium on Information and Communication Technologies, January 2005, 185-190.

[8] C.J. Michael Geisterfer, Sudipto Ghosh, Software Component Specification: A Study in Perspective of Component Selection and Reuse. Commercial-off-the-Shelf (COTS)-Based Software Systems 2006 Fifth International Conference on 13-16 Feb. 2006,100 – 108.

[9] P.A. Sun, S.W. Lin, et al., VERTAF: An Application Framework for the Design and Verification of Embedded Real-Time Software. IEEE Transactions on Software Engineering. 2004, vol. 30(10): 656-674.

[10] W. Jim, D. Jim, Using Z, Specification, Refinement, and Proof. University of Oxford.

Stability of Delayed Dynamical Congestion Control Algorithm*

Hongyong Yang
School of information science and Engineering, Northeastern University
Shenyang, Liaoning 110006, P.R. China
Or
School of Computer Science and Technology, Ludong University
Yantai, Shandong 264025, P.R. China
Email: hyyang@yeah.net
Fusheng Wang
Information Technology Office of Library, Ludong University
Yantai, Shandong 264025, P.R. China
Email: fswang@yeah.net
Shu-yan Bai
YanTai Vocational College
Yantai, Shandong 264000, P.R. China

ABSTRACT

An Internet congestion control algorithm is a complicated dynamical nonlinear feedback system with the communication delays. Based on the queue length of the buffers in link node, a more general TCP congestion control algorithm is analyzed. The delayed dynamical system with discrete-time is discussed, and many criteria on the stability are obtained. Finally, the computer simulations show that the stability criteria are validity.

Keywords: Internet, dynamical, congestion control algorithm, communication delays, discrete-time, stability criteria.

1. INTRODUCTION

When the requirements in Internet exceed the network resources (e.g., link bandwidths and the buffers etc.), the network congestion will take place. If the network congestion would not be controlled in time, the network was broken down at last. The network congestion control algorithms have two kinds: one is TCP rate control algorithms applied at the sources to adjust the system rates based on the network congestion marking (e.g., TCP Reno [1] and TCP Vegas [2]); Another is AQM control algorithms used at the link nodes to throw away the arriving packets when the network is overload (e.g., Drop Tail[1] and RED[3]).

The predominant TCP implementations in present Internet are TCP Reno [1] that has four functions mainly: slow-start, congestion avoidance, fast retransmit and fast recover. At the beginning of the network connection, TCP Reno starts with a small congestion window size (commonly one packet) to send out packets. The source increases its windows by one when it receives an acknowledgment, so the congestion window sizes will enlarge one times every round-trip time (RTT). This phase of doubling the window every round-trip time is called slow-start. When the congestion window sizes reach a threshold, TCP Reno enters the congestion avoidance phase, where it increases its window by one every round-trip time. On detecting a loss by timeout, TCP Reno sets the slow-start threshold to half of the current window size, retransmits the lost packet, and re-enters slow-start by resetting its windows to one. This congestion avoidance scheme is called additive-increase multiplicative-decrease (AIMD) algorithm. On detecting a loss through duplicate acknowledgments, TCP Reno adopts fast retransmit/fast recover method to adjust TCP parameters.

The new TCP rate control algorithm, TCP Vegas algorithm, improves upon TCP Reno through several main techniques. TCP Vegas increases the congestion window sizes with a more prudent way in the slow-start phase, where the window sizes are enlarged one times every two round-trip times. Entering the congestion avoidance phase, TCP Vegas adopts a new congestion avoidance mechanism. TCP Vegas sources compute the expected throughput every round-trip time, and obtain an estimated error by comparing the practice throughput with the expected throughput. The congestion window is adjusted by the estimated error. The window size is unchanged if the estimated error is between two number α and β which are set by the TCP Vegas. The source increases its window by one if the error is less than α, or decreases by one if the error greater than β. Another study of the TCP Vegas congestion avoidance is given in [11], in which a TCP Vegas congestion avoidance algorithm is proposed based on the queueing delay in the link. There are a number of studies for the feature of TCP Reno and TCP Vegas (e.g., [4,10,13]).

The delays are the inherent physical characteristic of the network system. It is necessary to analyze the complicated dynamical feedback systems with the delays for studying the stability of the congestion control algorithms. There are mainly two types of the studying methods. The first is to study the continuum-time delayed network communication systems, where the delayed differential equations were regarded as the models of the network to study the stability of the network systems [5, 6, 7, 9]. The second is to analyze the discrete time delayed network systems, where the features of the characteristic roots of the system equations were studied for the stability of the network systems [7, 8].

* This work was supported by the National Natural Science Foundation of China (under the grant 60574007), the Natural Science Foundation of Shandong Province (under the grant Y2006G01), the Science Foundation of LuDong University (under the grant 22320301) and the Postdoctoral Science Foundation of Northeastern University (under the grant 20060110), P R China.

In this paper, based on the TCP Vegas congestion control scheme in [11], we discuss a more general TCP Vegas congestion control algorithm. Considering the feedback delays in the algorithm, we study the stability of the delayed TCP Vegas congestion control algorithm in the discrete time system.

The rest of this paper is organized as follows. A delayed dynamical network model with different coupling strength is presented and some preliminaries are introduced in Section 2. In Section 3, the stability of TCP congestion control algorithm is analyzed. In Section 4, a network with time delay is used as an example to illustrate the theoretical results. Finally, conclusions are presented in Section 5.

2. DESCRIPTION OF THE QUESTION

In this section, we consider a network communication system with multiple-sources and multiple-links, $L=\{1,...,L_0\}$ denotes the links set, $N=\{1,...,N_0\}$ denotes the sources set. For every $l\in L$, we suppose the congestion rate of the link l as p_l (the link price), the network loads of the link l as y_l, the service capacity of the link l as $q_l(p_l(t))$, the source set using the link l as $N(l)\subseteq N$. For every $i\in N$, we suppose the congestion window of the source i as w_i, the transmission rate of the source i as x_i, the link set using the source i as $L(i)\subseteq L$. We denote a routing matrix R with $L_0\times N_0$ dimensions as

$$R_{li}=\begin{cases}1, & l\in N(l)\\ 0, & otherwise\end{cases}.$$

In this paper, supposing matrix R is of full low rank.

Supposing that every link can compute the total transmission rates of the sources using this link, and every source can receive the total feedback congestion rates from the links. In every transmission route, the source i adjust the transmission rate x_i based on the feedback congestion rate, then uses this new rate to transmit the data. The link l updates the congestion rate p_l by the network loads, then sends the feedback mark to each source. So the data are transmitted circulatively[3,12].

After analyzing the congestion avoidance mechanism of the large scale network with multiple-sources and multiple-links and discussing the equilibrium characteristic of TCP Vegas congestion control algorithm at the optimal network utilities, a TCP Vegas congestion control algorithm is proposed in [11]

$$p_l(t+1)=[\,p_l(t)+(\frac{y_l(t)}{c_l}-1)]^+,$$

where $p_l(t)$ denotes the queueing delay in the link, c_l as the service capacity of the link l, $[z]^+=\max\{0,z\}$.

Here, we discuss a more general algorithm than this one, and enter the feedback delays in the algorithm. We analyze the general TCP Vegas congestion control algorithm with the feedback delays, and study the asymptotically stability at the equilibrium point of the algorithm. The delayed TCP Vegas congestion control algorithm in discrete time is discussed in this paper as following (for all $l\in L, i\in N$):

$$p_l(t+1)=p_l(t)+\kappa_l(y_l(t)-c_l) \quad (1)$$

$$x_i(t)=f_i(q_i(t)),\ i\in N. \quad (2)$$

where f_i (the rate of the source i) is a strict monotone decrease function of the queuing delays $q_i=\sum_{l\in L}R_{li}p_l(t-d_2(l,i))$, $y_l=\sum_{i\in N}R_{li}x_i(t-d_1(l,i))$ is the network loadings, the control gain κ_l is a sufficiently small positive number. $d_1(l,i)$ Denotes the forward delay of the packets from the source i to the link l, $d_2(l,i)$ denotes the return delay of the feedback signal from the link l to the source i. There is a propagation delay on every route, denoting D_i, and $d_1(l,i)+d_2(l,i)=D_i$, for $i\in N$.

The system (1)-(2) can be explained as following: the link l receives the aggregate rate from the sources before $d_1(l,i)$ time, and compares with the service capacity, then adjusts its link delay (the congestion rate) $p_l(t)$. The source i updates the rate based on the aggregate queueing delays from the links at the time $d_2(l,i)$ ago.

This paper discusses the local stability of the system (1)-(2) with the communication delays. Supposing the equilibrium point as (x_i^*, p_l^*), we can obtain easily that the system (1)-(2) satisfies, for all $l\in L$

$$y_l^*=\sum_{i\in N}R_{li}x_i^*=\sum_{i\in N}f_i(q_i^*)=c_l\ . \quad (3)$$

Where $q_i^*=\sum_{l\in L(i)}R_{li}p_l^*$.

3. STABILITY OF DELAYED DYNAMICAL CONGESTION CONTROL ALGORITHM

In this section, we consider the delayed congestion control algorithm (1)-(2). Supposing the equilibrium point of the system is p_l^*, $f_i=f_i(\sum_{l\in L(i)}R_{li}p_l^*)$ is the equilibrium rate of the source i, f_i' the derivative of f_i on the equilibrium delays $\sum_{l\in L(i)}R_{li}p_l^*$. Define $p_l(t)=p_l^*+\hat{p}_l(t)$, the linearization of (1)-(2) yields neglecting higher order terms, for $l\in L$, $i\in N$

$$\hat{p}_l(t+1)=\hat{p}_l(t)+\kappa_l\hat{y}_l(t),$$

$$\hat{x}_i(t)=f_i'\hat{q}_i(t).$$

Where

$$\hat{y}_l(t)=\sum_{i\in N}R_{li}\hat{x}_i(t-d_1(l,i)),$$

$$\hat{q}_i(t)=\sum_{l\in L(i)} R_{li}\hat{p}_l(t-d_2(l,i))\cdot$$

Applying the Z transforms, we obtain

$$\hat{x}_i(z)=f_i'\sum_l R_{li}z^{-d_2(l,i)}\frac{k_l}{z-1}\sum_j R_{lj}\hat{x}_j(z)z^{-d_1(l,j)} \quad (4)$$

Where $X_i(z)=\mathrm{Z}(\hat{x}_i(t))$, and $X_i(0)=0$. Defining:

$R(z)=(R_{li}z^{d_2(l,i)}, l\in L, i\in N)$,

$R(-z)=(R_{li}z^{-d_2(l,i)}, l\in L, i\in N)$,

$F=diag(-f_i', i\in N)$,

$K=diag(\kappa_l, l\in L)$。

From (4) we have

$$X(z)=-diag(\frac{z^{-D_i}}{z-1}, i\in N)F^{1/2}R^T(-z)KR(z)F^{1/2}X(z) \quad (5)$$

The characteristic equation of the closed loop transfer function of the discrete time system (1)-(2) is

$$\det(1+diag(\frac{z^{-D_i}}{z-1}, i\in N)F^{1/2}R^T(-z)KR(z)F^{1/2})=0 \quad (6)$$

The discrete time system (1)-(2) is asymptotically stable, if and only if all poles of the closed loop transfer function (6) have modules less than unity. For studying easily, we take transform $z=e^s$, obtaining

$$\det(1+G(s))=0, \quad (7)$$

Where

$G(s)=diag(\frac{e^{-sD_i}}{e^s-1}, i\in N)F^{1/2}R^T(-s)KR(s)F^{1/2}$,

$R(s)=(R_{li}e^{sd_2(l,i)}, l\in L, i\in N)\cdot$

closed loop transfer function (6) has modules less than unity, i.e., all poles of (7) have negative real part. The discrete time system (1)-(2) is asymptotically stable, if and only if all poles of the closed loop transfer function (7) have negative real parts. The following theorem can ensure the stability at equilibrium point of the discrete time network system

Theorem 1: Supposing the equilibrium point of (1)-(2) is (x_i^*, p_l^*), the system (1)-(2) is asymptotically stable if existing a matrix $H=diag(h_l>0, l\in L)$, for $l\in L$,

$$\kappa_l\sum_{i\in N}R_{li}\frac{-f_i}{\sin(\frac{\pi}{2(2D_i+1)})}\sum_{r\in L}\frac{h_r}{h_i}R_{ri}<2$$

where κ_l is the element l in the congestion control gain matrix K.

Proof: It is easily to know that the opened loop transfer function of (7) is

$$G(s)=diag(\frac{e^{-sD_i}}{e^s-1}, i\in N)F^{1/2}R^T(-s)KR(s)F^{1/2}$$

Since the open-loop system is stable, for $s=j\omega$, obtain

$$G(j\omega)=diag(\frac{e^{-j\omega D_i}}{e^{j\omega}-1}, i\in N)F^{1/2}R^T(-j\omega)KR(j\omega)F^{1/2}\cdot$$

Firstly, we discuss the Nyquist image of the function $(\frac{e^{-j\omega D_l}}{e^{j\omega}-1})$, for $l\in L$,

$$\frac{e^{-j\omega D_l}}{e^{j\omega}-1}=\frac{1}{e^{j(D_l+1)\omega}-e^{jD_l\omega}}=\frac{e^{-j(\frac{\pi}{2}+\frac{(2D_l+1)\omega}{2})}}{2\sin(\frac{\omega}{2})}\cdot$$

Knowing easily that the Nyquist image of the function $(\frac{e^{-j\omega D_l}}{e^{j\omega}-1})$ crosses the real axis at $-\frac{1}{2\sin(\frac{\pi}{2(2D_l+1)})}+0j\cdot$

Secondly, we consider the matrix

$T(j\omega)=diag(M_iE_i(\omega), i\in N)$,

$Q(j\omega)=F^{1/2}R^T(-j\omega)KR(j\omega)F^{1/2}$,

$M=diag(1/M_i, i\in N)$,

$M_i=2\sin(\frac{\pi}{2(2D_i+1)})\cdot$

Where $E_i(\omega)=\frac{e^{-j(\frac{\pi}{2}+\frac{2D_i+1}{2}\omega)}}{2\sin\frac{\omega}{2}}$. we have

$$G(j\omega)=M^{1/2}T(j\omega)M^{1/2}Q(j\omega)$$

Since

$$\begin{aligned}&\rho(M^{1/2}Q(j\omega)M^{1/2})=\\&=\rho(M^{1/2}F^{1/2}R^T(-j\omega)KR(j\omega)F^{1/2}M^{1/2})\\&=\rho(R^T(-j\omega)KR(j\omega)FM)\\&=\rho(KR(j\omega)FMR^T(-j\omega))\\&=\rho(HKR(j\omega)FMR^T(-j\omega)H^{-1})\end{aligned}$$

where matrix $H=diag(h_l>0, l\in L)$, for $l\in L$, and $\rho(\cdot)$ denotes the matrix spectral radius. So

$$\begin{aligned}&\rho(M^{1/2}Q(j\omega)M^{1/2})\\&\le\max\{\kappa_l(\sum_{i\in N}R_{li}\frac{-f_i'}{2\sin(\frac{\pi}{2(2D_i+1)})})\sum_{r\in L}\frac{h_r}{h_i}R_{ri}\}\\&<1.\end{aligned}$$

Supposing λ as the eigenvalue of matrix $G(j\omega)$, there exists the eigenvector v, $\|v\|=1$, satisfying $G(j\omega)v=\lambda v$, obtaining

$$\begin{aligned}\lambda&=v^HG(j\omega)v\\&=v^H(M^{1/2}Q^{1/2}(j\omega))^Hdiag(M_iE_i(\omega))M^{1/2}Q^{1/2}(j\omega)v\end{aligned}$$

Defining $u=(\hat{G}^H(j\omega)/\sqrt{\rho})v=[u_1,\mathrm{K},u_{L_0}]^T$, then $u^Hu=1$. We have

$$\begin{aligned}\lambda&=u^Hdiag(\rho M_iE_i(\omega))u\\&=\sum_{l=1}^{N_0}(|u_i|^2\rho M_iE_i(\omega))\end{aligned}$$

and yield

$\lambda\in Co(0\cup\{\rho M_iE_i(\omega)\})\cdot$

Where $Co(\cdot)$ denotes the convex hull of a point set.

Because the Nyquist image of the function $(M_iE_i(\omega))$ crosses the real axis at $-1+0j$ point and the spectral radius of matrix $\rho<1$, we yield that the closed loop system is asymptotically stable by the generalized Nyquist criterion[14]. End of the proof of Theorem 1.

Supposing the matrix H as unity matrix in Theorem

1, we can obtain the following corollary directly.

Corollary 1: Suppose the equilibrium point of the system (1)-(2) is (x_i^*, p_l^*). The system (1)-(2) is asymptotically stable, if, for al $l \in L$,

$$k_l(\sum_{i\in N} R_{li}(\frac{-f_i'}{\sin(\frac{\pi}{2(2D_i+1)})})\hat{L}(i)) < 2.$$

Where $\hat{L}(i)$ denotes the link number used by the source i.

Corollary 2: Suppose the equilibrium point of the system (1)-(2) is (x_i^*, p_l^*), The system (1)-(2) is asymptotically stable, if, for all $l \in L$,

$$\kappa_l L_0 \sum_{i\in N} R_{li}(\frac{-f_i}{\sin(\frac{\pi}{2(2D_i+1)})}) < 2.$$

If the network system (1)-(2) have the same communication delays $D_i = D$, for $i \in N$, we can yield the following criteria for the stability of the system.

Corollary 3: Suppose the network system (1)-(2) with the same communication delays $D_i = D$, for $i \in N$. The system (1)-(2) is locally asymptotically stable, if one of the next conditions satisfying, for all $l \in L$,

1) $\kappa_l(\sum_{i\in N} R_{li}(-f_i')\hat{L}(i)) < 2\sin(\frac{\pi}{2(2D+1)})$.

2) $\kappa_l L_0 \sum_{i\in N} R_{li}(-f_i') < 2\sin(\frac{\pi}{2(2D+1)})$.

Where $\hat{L}(i)$ denotes the link number used by the source i.

4. SIMULATIONS

In order to validate the stable criteria obtained in this paper, we study the computer simulations as following. Supposing there are two links and five sources in the network, the rate of the sources is individually（x_i）:0.0025，0.005，0.0075，0.01，0.0125（Kb/ms），the window size of the sources（$w_i, i = 1,2,3,4$）：0.0025, $w_i = 0.005$; the service capacity of the links: $q_1 = 1(Kb/ms)$，$q_2 = 0.7(Kb/ms)$.The sources of x_1, x_2, x_3, x_4, x_5 share the first link, the sources of x_2, x_3, x_4, x_5 share the second link, so we have $R = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{pmatrix}$. The system adjusts the rate $\max(\min(w_i/(D+p(t)),1), x_i)$ [12] based on the queuing delay $p(t)$ from the links, where $p(t) = p_1(t - d_2(1,1))$, for $i = 1$; $p(t) = p_1(t - d_2(1,i)) + p_2(t - d_2(2,i))$, for $i \geq 2$; and $d_1 = d_2 = D/2$. We can obtain the equilibrium point (0.0125，0.003125). Supposing the diagonal matrix $H = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, delays $D_1 = 40ms$ and $D_2 = 80ms$, control gain $\kappa_1 = 0.000423$ and $\kappa_2 = 0.000128$, we study the stability and convergence of the TCP Vegas congestion control algorithm (Fig 1, Fig 2). Fig 1 draws the network loads in the link, and Fig 2 shows the queueing delays in the link. We can find that the transfer capacities in the links reach the maximum and the queueing delay is asymptotically stable, which is consistent with our conclusions.

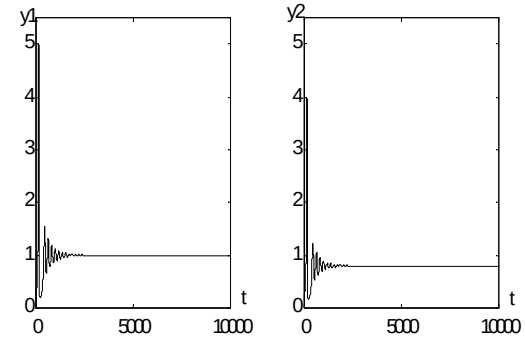


Fig. 1. The network loads in the links

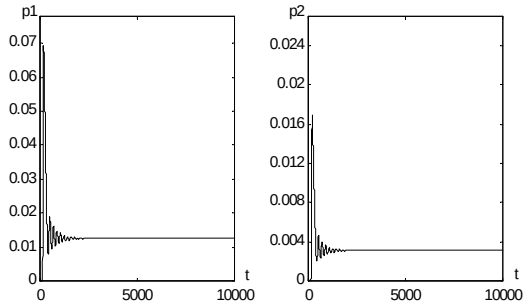


Fig. 2. The queueing delays in the links.

5. CONCLUSIONS

In the paper, we study the TCP Vegas congestion control algorithm with the feedback delays, and obtain the stability criteria on the equilibrium point in the network system. In lastly, the computer simulations show that the criteria is validity and practicability. These conclusions provide the firm theories foundation for designing the network scheme, ensuring the network stability and avoiding network congestion.

6. REFERENCES

[1] V. Jacobson, “Congestion avoidance and control,” Proceedings of ACM SIGCOMM’88, Stanford, CA, 1988, pp.314–329.

[2] S. Brakmo, L. Peterson, “TCP Vegas: end to end congestion avoidance on a global Internet,” IEEE Journal on Selected Area in Communications, 13(8), 1995, pp.1456-1480.

[3] S. Floyd, V. Jacobson, “Random early detection gateways for congestion avoidance,” IEEE/ACM Transactions on Networking, Vol.1, August.1993, pp.97-413.

[4] S. Low, F. Paganini, et al, “Internet congestion control,” IEEE Control Systems Magazine, 22(1), 2002, pp.28-43.

[5] S. Chong, R. Nagarajan, et al, “Designing stable ABR flow control with rate feedback and open loop control: First-order control case,” Performance Evaluation,

Vol.34, 1998, pp.189-206.
[6] S. Chong, R. Nagarajan, et al, “First-order rate-based flow control with dynamic queue threshold for high-speed wide-area ATM networks,” Computer Networks ISDN System, Vol.29, 1998, pp.2201-2212.
[7] R. Johari, D. Tan, “End-to-end congestion control for the internet: delays and stability,” IEEE/ACM Transactions on Networking, 9(6), 2001, pp.818-832.
[8] F. Kelly, A. Maulloo, et al , “Rate control for communication networks: shadow prices, proportional fairness, and stability,” Journal of the Operational Research Society, Vol.49, 1998 pp.237-252.
[9] Y.P. Tian, H.Y. Yang, “Stability of the Internet congestion control with diverse delays,” Automatica, 40(9), 2004, pp.1533-1541.
[10] W. Feng, S. Vanichpun, “Enabling compatibility between Tcp Reno and Tcp Vegas,” Proceedings of 2003 Symposium on Applications and Internet, 2003, pp.301-308.
[11] S. Low, L. Larry, et al, “Understanding Vegas: a duality model,” Journal of the ACM, 49(2), 2002, pp.207-235.
[12] S. Low, D. Lapsley, “Optimization flow control –I: basic algorithm and convergence,” IEEE/ACM Transactions on Networking, Vol.7, 1999, pp.861-874.
[13] O. Ait-Hellal, E. Altman, “Analysis of TCP Vegas and TCP Reno,” IEEE International Conference on Communication, Vol.1,1997, pp.495 –499.
[14] C. Desoer, Y. Yang, “On the generalized Nyquist stability criterion,” IEEE Transaction on Automatic Control, Vol.25, 1980, pp.187-196.

An Improvement on Parameter Recovery Technique Supporting Software Pipelining Optimization

Miao Wang[1], Yuqin Wang[5]
Dept. of Electrical Information Engineering, Institute of information science and Technology
Zhengzhou, Henan 450002, China
Email: zzwangmiao@sina.com
Guoming Cai[2]
Institute of Electronic Technology, Information Engineering University
Zhengzhou, Henan 450004, China
Rongcai Zhao[3], Jianmin Pang[4]
National Digital Switching System Engineering &Technological R&D Center
Zhengzhou, Henan 450002, China

ABSTRACT

Software pipelining is a well-known loop optimization technique, which can be very effective in exploiting instruction level parallelism by overlapping the execution of several consecutive iterations. However, it also profoundly restructures the low level code of programs and complicates the task of software systems that statically analyze or modify executable programs, e.g., reverse engineering systems, static binary translators. One problem we encountered in static binary translation was that using traditional parameter recovery technology brings in incorrect recovery in terms of binary code optimized by software pipelining due to the presence of register rotation. In order to solve this problem we present an improved parameter recovery algorithm. This algorithm has been implemented in our static binary translator and the experimental results are satisfactory.

Keywords: Parameter Recovery, Software Pipelining, Register Rotation, Live-variable Analysis.

1. INTRODUCTION

Parameter recovery is one of the key techniques in static binary translation. It serves to analyze which registers or parts of memory stack are used to pass parameters for a procedure call in the binary code, which is the foundation for translating a procedure into a function in high level languages.

Parameter recovery includes formal parameter recovery and actual parameter recovery (in the view of cal lee and caller), user procedure parameter recovery and library procedure parameter recovery (in the view of procedure type), register parameter recovery and memory stack parameter recovery (in the view of parameter location), and so on. In this paper, we are only concerned about the register formal parameter recovery of user procedures.

The traditional way to recovery parameter is: first determining where to pass the parameters (either on the registers or memory stack) based on specific operating system application binary interface (ABI) conventions; then determining the set of live-variables for a procedure by live-variable analysis; finally, determining which registers or stack locations are used to pass parameters with above information.

However, using traditional parameter recovery technology sometimes causes incorrect recovery in terms of IA-64 binary code particularly optimized by software pipelining due to the presence of register rotation.

The contribution of this paper is to present our improved parameter recovery algorithm which eliminates the affection of software pipelining optimization.

2. SOFTWARE PIPELINING

This section gives background information on the software pipelining and IA-64 loop support features. A more detailed description is in [1].

Software pipelining [2] is an important instruction scheduling technique for efficiently overlapping successive iterations of loops and executing them in parallel. It has been widely implemented in modern optimizing compilers for VLIW and superscalar processors [3,4] such as IA-64, Texas Instruments' C6X and StarCore's SC140 DSP that support instruction level parallelism.

Software pipelining has demonstrated benefits for loop-intensive application programs. However, the code expansion that results from software pipelining on traditional architectures can increase the number of instruction cache misses, thus reducing overall performance. IA-64 provides extensive support for software-pipelined loops, including register rotation, special loop branches, and application registers. When combined with predication and support for speculation, these features help to reduce code expansion, path length, and branch mispredictions for loops that can be software pipelined.

2.1. Register rotation

Register rotation provides a renaming mechanism that reduces the need for loop unrolling and software renaming of registers. Registers are renamed by adding the register number to the value of a register rename base register (rrb) contained in the current frame marker. The rrb register is decremented when certain special software pipelined loop branches are executed at the end of each kernel iteration. Decrementing the rrb register makes the value in register X appear to move to register X+1. If X is the highest numbered rotating register, its value wraps the lowest numbered rotating register.

Below is an example of register rotation. The swp branch pseudo-instruction represents a software pipelined loop branch:

```
L1:  ld4    r35 = [r4], 4
     st4    [r5] = r37, 4
      swp_branch L1;;
```

The value that the *load* writes to r35 is read by the *store* two-kernel iterations (and two rotations) later as r37. In the meantime, two more instances of the *load* are executed. Because of register rotation, those instances write their results to different registers and do not modify the value needed by the store.

2.2. Predication

Predication is the conditional execution of an instruction based on a qualifying predicate. A qualifying predicate is a predicate register whose value determines whether the processor commits the results computed by an instruction. IA-64 has 64 predicate registers, among those predicated registers p16 through p63 are rotated.

The rotation of predication registers serves two purposes. The first is to avoid overwriting a predicate value that is still needed. The second purpose is to control the filling and draining of the pipeline. To do this, a programmer assigns a predicate to each stage of the software pipelining to control the execution of the instructions in that stage. The predicate is called the stage predicate. For counted loops, p16 is architecturally defined to be the predicate for the first stage; p17 is defined to be the predicate for the second stage, etc.

A register rotation takes place at the end of each stage (when the software-pipelined loop branch is in the prolog and kernel loop). Thus a 1 written to p16 enables the first stage and then is rotated to p17 at the end of the first stage to enable the second stage for the same source iteration. Each 1 written to p16 sequentially enables all the stages for new source iteration. This behavior is used to enable or disable the execution of the stages of the pipelined loop during the prolog, kernel, and epilog phases.

3. HOW DO THE SOFTWARE PIPELINING OPTIMIZATIONS AFFECT THE PARAMETER RECOVERY

3.1. IA-64 parameter recovery technique

Our IA-64 parameter recovery technique is based on a combination of machine specifications, binary-interface specifications and the procedure recovery technology described in [5,6]. There are 3 steps when analyzing parameters of a single procedure:

1) Determine where to pass the parameters based on specific operating system ABI conventions

In IA-64, parameters may be passed in a combination of general registers, floating-point registers, and memory [7]. The first eight parameters are always passed in registers, while the remaining parameters are always passed on the memory stack, beginning at the caller's stack pointer plus 16 bytes.

2) Determine Livein (entry) by live-variable analysis

Live-variable analysis is performed to determine the set of livein variables at the entry of the procedure. A register that is used before its caller must define definition in a subroutine. To do it, following data-flow equations are set up for each basic block of the procedure. Iteratively compute equations until the fix point is presented.

$$IN[B] = (OUT[B] - DEF[B]) \cup USE[B]$$

$$OUT[B] = \bigcup_{S \in Succ[B]} IN[S]$$

Here, IN[B] represents the set of variables live at start point of block B; OUT[B] represents the set of variables live at end point of block B; USE[B] represents the set of variables with upwards exposed uses in block B; DEF[B] represents the set of variables defined in block B;

Step 3: Get formal parameters of the procedure

Finally, we compute the intersection of Livein(entry) and PossibleParameterSet to get formal parameters of the procedure.

3.2. an example

In order to explain how the software pipeline optimizations affect parameter recovery, we give an example as follows:

Example 1 The C code:

```
int mystery(unsigned);
main()
{
   unsigned x;
   printf("Enter an integer: ");
   scanf("%u", &x);
   printf("The result is %d\n", mystery(x));
   return 0;
}
int mystery(unsigned bits)
{
   unsigned i, mask = 1 << 15, total = 0;
   for (i = 1; i <= 16; i++, bits <<= 1)
      if ((bits & mask) == mask)
         ++total;
    return total % 2 == 0 ? 1 : 0;
}
```

This example was compiled on Itanium using icc 8.0 at optimization level –O2 running Redhat Linux 2.4.21-4.EL environment. Fig.1 shows Itanium code corresponds to Example 1 (Here we only illustrate the outline).

In this example, the steps of parameter recovery are as follows:

1) Determine possible parameter set

PossibleParameterSet={r32,r33,r34,r35,r36,r37,r38,r39}.

2) Determine Livein(entry)

Livein (entry) ={ r32, r35}.

3)

PossibleParameterSet∩Livein (entry)={r32,r35}.

So register r32 and r35 are the parameters of the procedure mystery. The result is different from the source codes.

```
4000000000001100 <mystery>:
4000000000001100:  alloc r9=ar.pfs,8,8,8
4000000000001106:  mov r3=0
400000000000110c:  mov r10=pr
4000000000001110:  mov r2=32768
4000000000001116:  mov r8=r32;;
400000000000111c:  mov pr.rot=0x10000
4000000000001120:  nop.m 0x0
4000000000001126:  mov r33=r32;;
      …                ...
4000000000001140: (p20) adds r3=1,r3
4000000000001146: (p16) and r34=r33,r2
400000000000114c: (p17) cmp4.eq.unc p19,p0=r35,r2
4000000000001150: (p16) shladd r32=r33,1,r0
      …                                ...
400000000000115c:  br.ctop.sptk.few 4000000000001140
<mystery+0x40>;;
      …              ...
400000000000118c:  br.ret.sptk.many b0;;
```

Fig. 1. Itanium code corresponds to Example 1

3.3. Problem analysis

The reason of incorrect recovery is the presence of register rotation. When the software-pipelined loop branch is executed in the prolog and kernel phase, a register rotation takes place at the end of each stage.
In this example:
L1: (p20) adds r3 = 1, r3;
L2: (p16) and r34 = r33, r2
L3: (p17) cmp4.eq.unc p19, p0= *r35*,r2
L4: (p16) shladd r32= r33, 1, r0

The value that the *and* instruction writes to r34 (L2) is read by the *cmp* instruction one iteration (and one rotation) later as r35 (L3). So the value of r35 is actually the value of r34, iteration before. And r34 has a definition before it is used as r35. So r35 is not a live-in variable at the entry of the procedure. The correct result is:

correct_Livein(entry) = {r32}

And the correct formal parameter is only in register r32.

4. SOLUTION

From above analysis, we can see that register rotation affects the analysis of the live variable and finally causes the incorrectness of the recovery of parameters. Considering the presence of rotated registers, we must remove those registers, which are rotated but have already been defined before rotation from Livein set. In this section, we present an algorithm for finding rot-related registers and an improved parameter recovery algorithm for IA-64.

4.1. The algorithm for finding Rot-related register

To know whether a rotated register has been defined, first we must find before rotation, which registers it, is related to, then we should judge whether the related register is in the defined set. For example, in Fig. 1 r34 is the rot-related register of r35, and r34 is in the defined set, so r35 is also defined before rotation.

To handle this problem we introduce the definitions of rot-related register and rot-interval.

DEFINITION 4.1. Register r1 is said to be rot-related register of r2 if the value of r2 is rotated from r1 by one or more times of register rotation.

DEFINITION 4.2. If register r1 is the rot-related register of r2 and the value of r2 is rotated from r1 by n times rotation. We call n the rot-interval of r1 and r2.

We can see that one register may have more than one rot-related registers, but for a specific rot-interval, it has only one rot-related register.

Our algorithm for finding rot-related register involves the following steps:

1) Judge whether it is a rotatable register

In IA-64, a fixed-size area of the predicate and floating-point register files (p16-p63 and f32-f127), and a programmable-sized area of the general register file are defined to rotate. The size of the rotating register in the general register file is determined by an immediate in the *alloc* instruction and must be either zero or a multiple of 8, up to a maximum of 96 registers. The lowest numbered rotating register in the general register file is r32.

2) Permute instructions within software-pipelined loop

We permute each instruction in the ascending sequence of its stage in the software pipelining. From the contents of section 2.2, we can know that software-pipelining stage can be computed by the stage predicate. For example, in a ctop loop, if the stage predicate is p16, it is in the first stage; if the stage predicate is p17, it is in the second stage, and so on. There are two particulars: the predicate register of the instruction maybe not specified or the predicate register of the instruction maybe not the stage predicate. For those unguarded instructions, they are in the first stage. For those instructions whose predicate register are not the stage predicates but are set by other compare instructions, they are in the next stage of the corresponding compare instructions.

3) Compute rot-interval between two instructions

Since a register rotation takes place at the end of each stage, rot-interval is equal to S1-S2 (Here S1 and S2 separately represent the stage number which two instructions are in).

4) Compute rot-related register

Assume N is a rotated register, n is the rot-related register of N with the rot-interval of d, s is the num of rotatable register.

A formula as follows can be used to compute rot-related register num:

$$n = \begin{cases} N - d & ((N-d) >= 32) \\ N - d + s & ((N-d) < 32) \end{cases}$$

4.2. Improved live-variable analysis

Here, we define a new set named DEFROT.

DEFROT [B]= {x | x is a rotated variable in block B and one of the rot-related registers of x is in the DEF [B]}

We use two new data-flow equations described below to compute live variable.

$$IN[B] = (OUT[B] - DEF[B]) \cup USE[B] - DEFROT[B]$$

$$OUT[B] = \bigcup_{S \in Suss[B]} IN[S]$$

Our improved live-variable analysis algorithm is shown in Fig. 2.

```
Input: a control flow graph G
Output: in[B] and out[B] of every basic block B in G
Algorithm:
for every basic block B do  in[B]:=Ø;
change :=true;
while change do begin
    change :=false;
    for every basic block B do begin
     out[B] = ⋃ in[S]   (S∈Suss[B])
     in[B] = (out[B] − def[B]) ∪ use[B] − defrot[B]
    if in[B]≠old in[B] then change :=ture;
   end
end
```

Fig. 2. An improved live-variable analysis algorithm

4.3. Improved IA-64 parameter recovery algorithm and comparison

Our improved IA-64 parameter recovery algorithm still consists of 3 steps.

1) Determine possible parameter set

2) Determine Livein(entry) by improved live-variable analysis

3) Compute the intersection of Livein(entry) and PossibleParameterSet

Among those three steps, the step 1 and step 3 are the same as before. The distinct difference between two algorithms is the live-variable analysis. In improved algorithm, we use improved live-variable analysis described in section 4.2. When a procedure has no rotated registers, the

set of DEFROT is always null. And the revised data-flow equations are

$$IN[B] = (OUT[B] - DEF[B]) \bigcup USE[B] - \Phi$$

$$OUT[B] = \bigcup_{S \in Suss[B]} IN[S]$$

It is obvious that the new data-flow equations are the same with the traditional ones if there is no rotated register. When a procedure has some rotated registers, the result of the new data-flow equations is different from that of traditional ones. It excludes the rotated register whose rot-related registers have already been defined from the set of Livein set.

In example 1:

The result of new algorithm is: Livein (entry) = {r32}

The result of original algorithm is:

Livein (entry) = {r32, r35}

In conclusion, our improved IA-64 parameter recovery algorithm is different from the traditional one if there are some rotated registers. In other conditions, it is same with the traditional one.

4.4. Experimental results

The improved parameter recovery technology has been implemented in a static binary translator I2A, which can automatically translate the binary code on IA-64 to a C program first and to an ALPHA binary code next. We conducted experiment on 82 c programs belonging in different applications with different loop length and various situations of procedure calling. First, we compiled those c source programs on the source machine (IA-64) using icc 8.0 with -O2 options. Then we use our static binary translator I2A to convert obtained IA-64 executables to ALPHA binary codes. Finally, we compare the time to execute the converted executables on the target machine with the time needed for the same source programs compiled by a native gcc compiler on that target machine. All computation results show our improved parameter recovery algorithm is valid for these programs. Table 1 shows the results of four programs from the SPECCint-2000 benchmark suite with the input of test. The results were obtained on an ALPHA DS20E, 666MHz (*2) machine running the Redhat Linux 7.2 for ALPHA environment.

Table 1. The experimental results of four programs

Benchmark	mcf	bzip2	gzip	Crafty
Number of functions	24	62	89	104
Number of functions including rotate registers	11	13	23	26
Number of functions incorrectly recovered with traditional algorithm	6	10	12	13
Number of functions incorrectly recovered with improved algorithm	0	0	0	0
native code size(bytes)	200263	368517	482797	578731
native code execute time(sec)	0.94	9.71	8.05	0.12
translated code size(bytes)	375242	560304	851376	4599942
translated code execute time(sec)	4.35	195.77	58.98	1.47
code expansion ratio	1.8	1.5	1.7	7.9
execute time ratio	4.4	2.4	3.2	12.8

5. CONCLUSION

In this paper, we discussed the need for improving parameter recovery technique by giving a typical example of code optimized by software pipelining. It showed that using traditional parameter recovery algorithm sometimes would lead to incorrect results due to the presence of the register rotation. It was the key issue to eliminate the affection of register rotation on live-variable analysis. We proposed an improved parameter recovery algorithm that could be used to recovery parameter from codes optimized or not optimized by software pipelining.

The distinct difference between our algorithm and traditional algorithm was the data-flow equations. We defined a new set named DEFROT and revised the data-flow equations. The revised data-flow equations were equal to the unrevised data-flow equations in absence of register rotation. The improved parameter recovery algorithm had been implemented in a static binary translator I2A. Preliminary results indicated that using our improved algorithm could correctly recovery parameters from codes optimized by software pipelining.

6. REFERENCES

[1] Intel Corporation, “Intel IA-64 Architecture Software Developer's Manual”, IA-64 Application Architecture, 245319-001, Vol.1, January 2000.

[2] V.H. Allen, R.B. Jones, et al, “Software pipelining”, ACM Computing Surveys, 1995, 27(3): 367-432.

[3] J.C. Dehnert, R.A. Towle, “Compiling for the Cydra 5”, Journal of Supercomputing, 1993,7(1-2): 181-228.

[4] M, Lam, “Software pipeling: An effective schedule technique for VLIW machines”, Proceedings of the SIGPLAN 1988 Conference on Programming Language Design and Implementation. New York: ACM Press, 1988,318-328.

[5] C. Cristina, S. Doug, “Procedure Abstraction Recovery from Binary Code”, Technical Report 448, Department of Computer Science and Electrical Engineering, the University of Queensland, 1999.

[6] C. Cristina, Mike van Emmerik et al, “The University of Queensland Binary Translator (UQBT) Framework”, February 2002.

[7] Intel Corporation, “IA-64 Software Conventions and Runtime Architecture Guide”, January 2000.

Modeling Software Project Scheduling Based on Team Productivity And Its Simulations

Yujia Ge
Dept. of Computer Science and Technology, Zhejiang Gongshang University
Hangzhou, Zhejiang 310035, China
Email: yge@mail.zjgsu.edu.cn

And

Carl K. Chang, Hsin-yi Jiang
Dept. of Computer Science, Iowa State University
Ames, IA, 50011, USA
Email: chang@cs.iastate.edu, hsinyij@cs.iastate.edu

ABSTRACT

Experienced software project managers understand the difficulty intrinsic to managing a complex software project where a large team of engineers and developers work together in a dynamic environment. This paper proposes a model to support software project scheduling based on team capability, in which dynamic elements are included. The model and its simulation experiments are reported. The results show that this model provides useful decision support under realistic conditions.

Keywords: Project Scheduling, Genetic Algorithm, Team Productivity, Optimization, Software Project Management.

1. INTRODUCTION

Software engineering is neither a classical engineering discipline as mechanical or electrical engineering, nor does it dictate a standardized software development process. These characteristics make software project management (SPM) very difficult to satisfy budget and schedule constraints that are set by an organization and its stakeholders. It is obvious that significant productivity differences do exist among individual software developers. However, current resource-constrained scheduling techniques mostly focus on the availability of resources instead of the capability of resources [1]. This is not reasonable and incomplete especially in software project management where the capability of human resources can determine whether a project fails or succeeds.

Based on our previous work, this paper addresses the aforementioned problem and proposes a novel model for software project scheduling by modeling software development process at the micro level using system dynamics simulations. GAs, already popular in many domains of optimization research, provide better solution to the complex project management problems than some other methods [2]. To the best of our knowledge, [3] represented the first attempt to develop a rich representation for SPM, which is capable of supporting a genetic algorithm based optimization approach. Improvements to the original model were described in the task-based model [2]. We continue to use GA as the major optimization approach in this work.

2. THE MODEL

A framework illustrated in Fig. 1 is introduced to do scheduling based on skill match and the capability of resources, and help capture the dynamic nature of software management, such as learning, overworking and schedule pressure with the introduction of hard deadline.

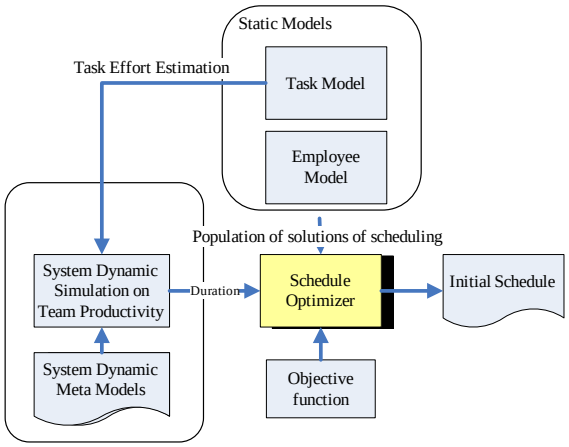


Fig. 1. A Model to Support Software Project Scheduling with Dynamic Factors

The framework has three main parts:
1) Schedule optimizer. Central to automatic scheduling and optimization in our work is a genetic algorithm. With the information from the static model of tasks and employees (part 3), the GA can generate populations of possible solutions, which are evaluated by an objective function through system dynamics simulation (part 2). Finally, a near-optimal schedule can be obtained.
2) Dynamic models and simulation. Capability-based system dynamics simulation is employed to calculate task durations according to different human resource assignments. It provides the task durations to the schedule optimizer (part 1) as a part of the input to evaluate a schedule.
3) Static models. It includes the static part of task models (estimated effort of task, task penalty model and required task skill lists, etc.) and employee models, such as the employee payment.

2.1 Static Models

In the employee model, an employee is represented by his ID and several properties, such as hourly salary, overwork salary, skill list with proficiency level, learning factor, experience factor, max work load and available time. Employee payment model, skill proficiency-obtaining model,

and experience-obtaining model are defined in a way similar to the timeline-based methodology reported earlier [4].

In the task model, a task is represented by its ID and a set of properties, such as ideal finished time, hard deadline, required skill list, ancestors list, required effort, penalty rate. Same as most other scheduling models, tasks are based on activity networks.

2.2 Team Productivity Model

Realizing that it is difficult to model a complex system with hundreds of interrelated factors, a high-level Meta model is introduced to help simplify the issue. In software development process, team productivity, which determines the overall project performance, is the key component affected by a complex set of factors as illustrated in Fig.2:

1) Individual productivity, communication overhead, and other factors, which are different among individuals yet contribute to the team productivity;

2) Experience, learning, overworking, and schedule pressure factors, which affect the individual productivity.

Although other factors such as motivation are also critical to individual productivity, for the purpose of demonstrating only key concepts we will not model more factors in this paper. Further noted that the quality of project management largely depends on the ability of managers, the determination of factor influences can therefore be controlled by project managers through a control parameter. Control parameter takes value 0 or 1 to turn the factors on or off.

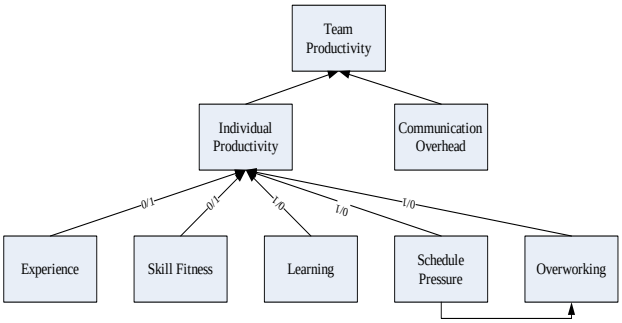


Fig. 2. Team Productivity Model

Team productivity (P_{team}) describes the relationship between individual productivity (InP_i) and related factors, such as communication overhead (*ComOverhead(N)*).

$$P_{team} = \sum_{i=1}^{N} InP_i \times \left(1 - ComOverhead\left(N\right)\%\right) \qquad (1)$$

Here N is the number of employees assigned to the task. The productivity of the software development group can be stated as the psychological model of group productivity by Ivan Steiner [5]. In his model:

Actual Productivity = Potential Productivity - Losses Due to Faulty Process

where losses due to faulty process refer basically to communication and motivation losses, as illustrated in Eq (1).

Communication is an essential component in software development, but it is also an overhead. Within a bigger team, communication overhead among team members is obviously increased. Abdel-Hamid [5] lists several researches to show that it is widely held that communication overhead increases in proportion to n^2, where n is the size of the team, which can be expressed as Eq. (2).

$$ComOverhead(n) = \begin{cases} 100 & \\ 0.05 \times (n-1)^2 & 2 \le n \le 45 \\ 0 & n = 1 \end{cases} \qquad (2)$$

Individual productivity is affected by different types of factors, which can be represented using Eq (3).

$$InP = nomP \times f_{skill} \times f_{experience} \times f_{learning} \times f_{schedulel} \times f_{overworking} \qquad (3)$$

Where *InP* corresponds to the individual productivity; *nomP* corresponds to the nominal productivity; f_{skill} corresponds to the skill fitness factor; $f_{experience}$ corresponds to the experience factor; $f_{learning}$ corresponds to the learning factor; $f_{overworking}$ corresponds to the overtime factor; $f_{schedulel}$ corresponds to the schedule pressure factor.

Skill fitness is defined as:

$$f_{skill} = \frac{0.1 \times \sum_{i=1}^{s} S_i}{s} \qquad (4)$$

Eq.(4) is to evaluate the skill fitness of an employee to certain task given *s* required skills. f_{skill} can be computed off-line and considered as a constant before optimizing the schedule. $f_{experience}$ represents the experience of an employee for a given task, and it is given beforehand and is also a constant for our simulation.

$f_{learning}$ represents the improvement of understanding the task along with the progress of task itself. Mathematically, it is defined as an s-curve equation [5].

$f_{schedulel}$ can be obtained from *schedule pressure* using a lookup table from the Vensim documents [6]. S*chedule pressure*} is a function of the current time ($T_{current}$), which is defined as Eq. (5) derived from the research by Abdel-Hamid [5].

$$schedule \quad pressure = \begin{cases} 5 & T_{current} > T_{idealfinishtime} \\ \min\{\dfrac{workflow_{required}}{workflow_{normal}}, 5\} & T_{current} \le T_{idealfinishtime} \end{cases}$$

where

$$workflow_{required} = \frac{effort_{remaining}}{T_{idealfinishtime} - T_{current}}$$

$$workflow_{normal} = P_{team} \qquad (5)$$

Overworking ($f_{overworking}$) is defined as Eq (6).

$$f_{overworking} = overtime \times fatigue \quad productivity \qquad (6)$$

We adopt the relationship of *overtime* and *fatigue productivity* given in Vensim [6], which uses the lookup table to obtain the value of *overtime* given the value of *schedule pressure*; it uses the lookup table to obtain the value of *fatigue productivity* given the value of *overtime* obtained above. These two-associated lookup table is listed as follows:

(1) Overtime lookup table: (schedule pressure, overtime) = (1, 1), (1.2, 1.2), (1.5, 1.4), (2, 1.45), (5, 1.5);

(2) Fatigue productivity lookup table: (overtime, fatigue productivity) = (1, 1), (1.2, 0.9), (1.4, 0.75), (1.45,0.6), (1.5, 0.3).

2.3 The Objective Function

Several assumptions have been made for our scheduling problem:

1) Each task has to be finished in a continuous duration;
2) Different people can work on different tasks at the same time, but cannot work over their maximum overwork level;
3) Every employee assigned to the task needs to do the work in the whole duration for a certain task in each schedule.

Although other possible objectives, such as minimum time to finish the project and maximum return on investment, may also be expected to achieve, we focus on the overall minimum cost of a project at this stage.

Simulation of any candidate scheduling solution for the schedule optimizer can be done to produce task duration. Fig. 3 shows an example of system dynamics simulation.

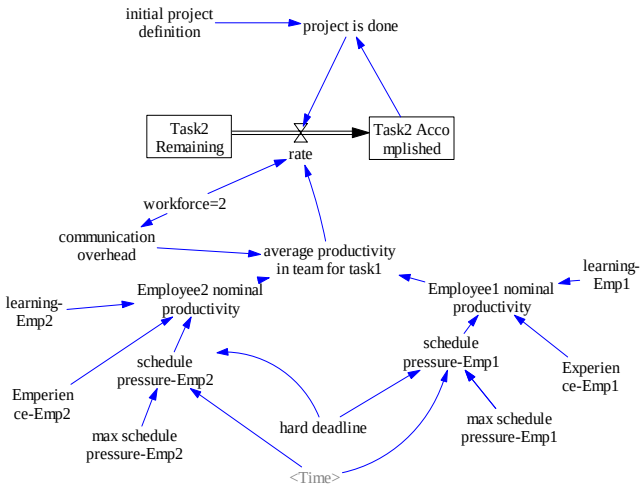


Fig. 3. An Example of System Dynamics Simulation

3. SIMULATION EXPERIMENTS

The principle that using whatever encoding is most natural to solving your problem, and devising a GA with that encoding, has been widely accepted unless there is more theoretical progress on GAs [7]. Our solution representation *SchedGenome* is written as a composite genome with two independent structures. It overcomes many of the complexities inherent in searches by not generating invalid solutions and causing no loss in expressiveness.

A schedule is represented as a candidate solution $S = \{A, L\}$ as shown.

- A is a 1D task-employee array that stores the information of task-employee assignments;
- L is the priority list by which a certain topological-sort vector representing the execution order of the tasks in the schedule can be derived.

Genetic operators, such as initialization, selection, crossover and mutation operators are also defined to manipulate the two structures and also account for the validity of the solutions in the solution space.

Besides several small experiments to validate the method, larger experiments were also conducted, such as the one consisting of 21 tasks with the properties listed in Table 2 and 10 employees. The employees who are all available from Jan 1, 2005 to Nov 1, 2005, in turn, each possessed 5 skills to a greater or less extent as listed in Table 1. Each of the 5 skills was needed by at least one task and many tasks required multiple skills. It is not immediately obvious, even the most experienced software manager is entrusted to do this, what the optimal assignments would be in this case.

Table 1. Employee Properties with Skill Proficiency

ID	Hourly& Overwork salary ($)	Max Workload (%)	Initial Exp	Learning	Skill 1	2	3	4	5
1	38, 48	150	5	1.5	4.5	4	0	5	5
2	33, 33	100	4	1.3	0	0	2	0	5
3	30, 35	125	3	1.3	4	4	2	0	5
4	35, 35	75	4	1.1	0	4.7	3	3	0
5	35, 40	125	4	1.1	4.5	4	4	0	5
6	36, 46	125	4	1.4	0	4.5	4	4	0
7	36, 36	100	4	1.5	4.5	5	5	0	0
8	35, 35	100	4	1.1	0	0	3	4	5
9	30, 35	125	3	1.2	0	0	4	5	0
10	36, 46	150	5	1.3	5	4	0	3	4

Table 2. Task Properties

Task ID	Effort	Soft Deadline	Hard Deadline	Penalty Per Day	Skills Required
1	0.25	02-01	03-01	1000	1 3
2	0.5	03-01	04-01	1000	2 3 5
3	0.8	03-01	04-01	2000	1 5
4	0.25	03-01	04-01	2000	1 2
5	0.6	11-01	11-01	0	1 3
6	0.5	11-01	11-01	0	3 4
7	0.3	11-01	11-01	0	2 4 5
8	0.4	11-01	11-01	0	2 3
9	0.25	11-01	11-01	0	2 4
10	0.5	11-01	11-01	0	4 5
11	0.5	11-01	11-01	0	1 3 5
12	0.25	07-01	08-01	1000	2 4
13	0.8	11-01	11-01	0	3 4
14	0.5	09-01	10-01	3000	1 2 3
15	1	11-01	11-01	0	2 3 4
16	0.5	11-01	11-01	0	4 5
17	0.5	11-01	11-01	0	1, 3, 5
18	0.25	11-01	11-01	0	2 4
19	0.8	11-01	11-01	0	3, 4
20	0.5	11-01	11-01	0	1, 2, 5
21	1	10-01	11-01	5000	2, 3, 5

After setting all the parameters, the best result from the search algorithm is shown in Table 3, and the cost of the

schedule is 127448, by considering all the factors. Table 4 shows the calculated costs with certain factors excluded. Without the learning factors or communication overhead, the overall cost increases as expected. Without schedule pressure, no solution is found in this highly constrained case. Although these results are intuitive in this simple scenario, in more complicated situation data can help analyze the influence of certain factors and assist managers to make more sensible decisions.

Table 3. A Near Optimal Schedule

Task ID	Days	Start & End Date	Prec	Resource
1	11	Jan 1-Jan 11		Employee7
2	18	Jan 12- Jan 29	1	Employee3 [50%], Employee5
3	24	Jan 30-Feb 22	1	Employee1, Employee3
4	16	Jan 12-Jan 27	1	Employee1 [50%]
5	21	Feb23-Mar 15	3,4	Employee3 [25%], Employee7 [75%]
6	15	Mar17-Mar 31	2,5	Employee6
7	22	Feb 23- Mar 16	2	Employee1 [50%]
8	11	Apr 3- Apr 13	6	Employee7 [75%]
9	5	Mar 17-Mar 21	4	Employee1
10	12	Mar 22 - Apr 2	9	Employee1
11	18	Mar 17 - Apr 3	7	Employee3
12	10	Apr 3 - Apr 12	9,10	Employee1
13	29	Apr 4 - May 2	11	Employee6, Employee9 [75%]
14	19	May 3- May 21	7, 8, 9, 12, 13	Employee7
15	31	May 3 - Jun 2	13	Employee6
16	12	May 22 - Jun 2	14	Employee1
17	23	Jun 3 - Jun 25	12	Employee3 [75%]
18	9	Jun 3 - Jun 11	15	Employee1 [75%]
19	33	Jun 3 - Jul 5	15, 16	Employee9
20	14	Jun 26 - Jul 9	17	Employee1
21	83	Jul 10 - Sep 30	17, 20, 19, 18	Employee5

Table 4. Results without Considering Some Factors

Cost (dollars)	Best	Worst	Average
With all the factors	127448	135356	131104
Without learning factor	149426	152072	151174
Without communication overhead factor	125222	127478	126017

4. CONCLUSION

This paper proposed a software project scheduling model to support initial schedule generation. Improvements to our previous models have been made especially in integrating dynamic factor modeling with a new optimization method. Preliminary case studies demonstrated that it has the ability to support good schedule generation and prevent the kind of project failures made by skill mismatches.

There are however a few constraints. There is still a dearth of software and process metrics that would support this type of research. Care must be taken to correctly construct dynamic models. Despite these constraints, our experience with capability-based scheduling indicates that it provides a meaningful approach to get better scheduling by considering dynamic factors.

5. REFERENCES

[1] V. Plekhanova, "On Project Management Scheduling where Human Resource is a Critical Variable", Proc. 6th European Workshop on Software Process Technology (EWSPT-6), in Lecture Notes in Computer Science series, Springer-Verlag, London, UK, 1998, 116-121.

[2] C. Chang, M. Christensen, et al, "Genetic Algorithms for Project Management", Annals of Software Eng., vol.11, Kluwer Academic Publishers, 2001, 107-139.

[3] C. Chao, "Software Project Management Net: a New Methodology on Software Management", PhD dissertation, University of Illinois at Chicago, Department of Electrical Engineering and Computer Science, 1994.

[4] Y. Di, "Timeline Based Model For Job Scheduling with Genetic Algorithms", master's thesis, Univ. Illinois at Chicago, 2001.

[5] T. A. Hamid, S.E. Madnick, Software Project Dynamics: An Integrated Approach, Englewood Cliffs, NJ, Prentice Hall, 1991.

[6] Vensim Software Tool, Ventana Systems, Inc., 1996-2004, http://www.vensim.com/.

[7] M. Mitchell, An Introduction to Genetic Algorithm, Cambridge MA, MIT Press, 1996.

IF-AntNet: A New Load Balanced Routing Algorithm for Direct Interconnection Networks

Jiyun Niu[1], Huaxi Gu[2], Donglei Yao[3], Changshan Wang[1], Xiaohui Hao[1]
[1.] School of Computer Xidian University Xi'an China 710071
Email: njy_001@163.com
[2.] State key lab of ISN Xidian University Xi'an China 710071
Email: hxgu@xidian.edu.cn
[3.] Information Technology Centre of Hebei Provincial Finance Bureau 050051)

ABSTRACT

This paper introduces a novel algorithm IF-AntNet (Improved Forward-only AntNet) for direct interconnection networks. IF-AntNet only uses forward ants to update the routing information on each visited node during their trip. A forward ant carries along the estimated trip time from the current node to its source node. And a precise model, which contains transmission delay, propagation delay and queuing delay, is used to figure out the information. The routing strategy in IF-AntNet balances between current network condition and historical routing information recorded in the routing table. Data packets share the same routing strategy as forward ants to exploit the maximum link utilization. The simulation results show that the new algorithm outperforms traditional algorithms such as dimension order routing (DOR), Duato's algorithm, and GAL.

Keywords: AntNet, load balanced, direct interconnection network, forward routing.

1. INTRODUCTION

Direct interconnection networks have found wide applications, from processor memory interconnect [5], I/O interconnect [6], to switch and router fabrics [7]. Typical networks include Mesh, Torus and Hypercube [9].

Many algorithms are proposed to solve the routing problem in Interconnection networks, such as Dimension order routing (DOR) [9], Duato's algorithm [10], GAL [11], and so on. Unfortunately, those algorithms consider only local information and lack a good knowledge of global network condition, resulting in the failure to balance the network loads to a satisfactory degree.

Nowadays, there arises a study of routing based on swarm intelligence. Swarm intelligence is an evolving field yielding methods for distributed systems optimization [1]. As far as the network routing is concerned, the paradigms of ants are considered to be most promising. Those remarkable insects can discover the shortest path from their colony to the food source and share it with other ants through stigmergy [2]. Stigmergy is a form of indirect communication through which ants coordinate their problem-solving activities. Pheromone [3], which is laid by ants during their trip and can be sensed by other ants, is used for stigmergy. Besides, pheromone diffuses with time. An ant is more likely to choose a path with more pheromone. In this way, the shortest path will be left finally.

There already exit many successful adaptations of ant behavior to network control and network routing, with the AntNet being the most prominent [4]. Two of the most remarkable contributions of AntNet are its stochastic decision to route a packet to result in an automatic network load balance and the acquirement of global network information with low cost by launching ants towards the network regularly. In AntNet, ants fall into two categories: forward ants and backward ants. A forward ant is used to collect routing information and get an acyclic path between a source and destination node, while a backward ant is responsible for updating the routing tables during its return trip. However, a Forward ant-only algorithm for telecommunications networks was proposed early in [16]. Unfortunately, the algorithm only applies to cost-symmetric networks and lacks preciseness in its routing model. An overview of ant-based algorithms can be found in [8].

Inspired by [4] and [16], we propose a new algorithm called IF-AntNet (Improved Forward-only AntNet) for indirect interconnection networks.

To provide low latency on local traffic as well as livelock freedom, minimal routing is adopted in IF-AntNet for both ants and data packets. A livelock happens when a packet goes round and round but can not reach its destination forever. Deadlocks, caused by packets' circular waiting for occupied resources, are detected and resolved by a software-based detection and recovery mechanism [17]. When a packet can not be sent out within a given period, then it is supposed to encounter a deadlock and submitted to the local node with its resources freed. After a short period, the submitted packet will be injected into the network again.

As far as load balance is concerned, the solution is borrowed from [4] and [16]. Like [16], only forward ants, which are responsible for collecting routing information and updating the routing table during their trip, are used in the algorithm for the purpose of reducing the cost brought by ants. Unlike [16], forward ants in IF-AntNet make an estimation of the trip time from current node to their source node with a mathematic model. What is more, the model used by forward ants is rather precise for its consideration of transmission delay, propagation delay as well as queuing delay. Here, queuing delay is a statistical value calculated from a model borrowed from [4]. In addition, data packets are routed in the same way as forward ants to exploit the maximum link utilization. Similar to [4], packets and ants are routed randomly according to the goodness of neighbors. Different from [4], the goodness of a neighbor is measured by not only the routing information recorded in the routing table but also the current link condition, which makes the routing decision more favorable.

The paper is organized as follows. Section 2 gives a detailed description of IF-AntNet. We measure the performance of IF-AntNet in section 3, and compare its performance with several existing algorithms through simulation. Finally we arrive at some conclusions and give the future work in section 4.

2. IF-ANTNET ALGORITHM

In AntNet, both forward ants and backward ants are used. Forward ants use the same queues as data packets, while backward ants use high-priority queues to update the routing tables as quickly as possible with the information collected by forward ants. Though it has achieved a good performance, the traffic of ants may become a burden of the network. Under low traffic, ants may increase the ETE delay of data packets for the reason that they compete with data packets for the same link resources; Under High traffic, ants may take a long time to reach their destination and thus be unable to update the routing tables timely.

So In IF-AntNet model, only forward ants are used to prevent those unfortunate things to happen. In the algorithm, each node in the network needs to maintain three routing structures:

1) The trip-time table T, which are records of the estimated trip time from current node across each of its neighbors to other nodes.

2) The statistical model S of the average queuing delay for a packet to be sent out, which is borrowed from AntNet:

$$t_{new} = t_{old} + \gamma * (t_w - t_{old}) \quad (1)$$

Where t_{new} and t_{old} is respectively the new and the old statistic value, t_w is the newly detected value of the queuing delay, γ is a justifying factor, ranging from 0 to 1.

3) The pheromone table P, which records the goodness of choosing a neighbor as the next node to reach a destination.

Informally, the main characteristics of IF-AntNet can be summarized as follows:

1) Forward ants are launched concurrently with the data traffic by all nodes towards random destinations at regular intervals. Destinations are locally selected according to the data traffic patterns: if the current node id is *k*, f_k^d represents the total number of packets coming across the current node *k* towards destination node *d*, then the probability of choose *d* as destination is defined as follows:

$$p_d = \frac{f_k^d}{\sum_{i=1}^{N} f_k^i} \quad (2)$$

Where *N* is the total number of the nodes in the network.

2) Forward ants use the same queues as data packets.

3)As it moves step by step towards its destination, a forward ant carries along the estimated trip time from the current node to its source node, which is calculated by a precise mathematic model described later. With this newly calculated information, it directly updates the information of its source node in the trip-time table on each immediate node.

4) Once a packet is sent out to a certain port, its queuing delay is recorded to update the statistics of the average queuing delay for the port.

5) The goodness of a neighbor for a destination is dependent not only on historical information recorded in the trip-time table T but also on the current status of queues for outputting.

6) Both forward ants and data packets are routed randomly according to the pheromone table.

7) Minimal routing is adopted for both ants and data packets to provide low latency on local traffic as well as livelock freedom.

8)A software-based detection and recovery mechanism is applied to detect and resolve deadlocks.

The models used in the algorithm are described in detail in the following paragraphs.

When a forward ant reaches an intermediate node *k*, it first estimates the trip time a packet takes to return to the previous node *p* it has come from with the following formula:

$$t_k^p = \frac{ql_k^p + pksize}{br_k^p} + txd_k^p + pd_k^p + wt_k^p \quad (3)$$

Where ql_k^p is the length of the queue for neighbor *p* at node *k*, $pksize$ is the average size of a packet, br_k^p is the bandwidth of the link between node *k* and node *p*, txd_k^p and pd_k^p are respectively transmission delay, propagation delay of the link between node *k* and node *p*, wt_k^p is the statistical queuing delay of the port at node *k* for its neighbor *p*.

Then it updates the trip time from current node to its source node:

$$t_k^s = t_k^p + t_p^s \quad (4)$$

Where t_p^s is the trip time a packet takes to go from previous node *p* to the source node *s* that is carried along.

Finally it directly writes this information into the corresponding item of the trip-time table on the current node.

The trip-time tables built by the random movements of ants provide an estimate of the condition of the network. However, this information may be more or less outdated due to rapid changes in the network traffic. Therefore, it is quite necessary to take the present network conditions into account. The model used to figure out the pheromone table is defined as follows:

$$g_n = \alpha * \frac{1}{T_{kn}^d} + \beta * \frac{1}{\frac{ql_k^n}{br_k^n}} \quad (5)$$

Where g_n is the goodness of a neighbor *n* for current node *k*, T_{kn}^d is the estimated trip time a packet takes to reach its destination *d* through neighbor *n*, which is recorded in the trip-time table on current node *k*. ql_k^n is the size of the queue for a neighbor *n* at node *k*, br_k^n is the bandwidth of the link between *k* and *n*. α and β are weighting factors for historical records and current network status, both ranging from 0 to 1. And β is usually defined as (1- α). If α is zero, then the goodness of a neighbor is completely dependent on the current states of links; if β is zero, only information recorded in a trip-time table is used to determine how good a neighbor is; if α is more than 0 but less than 1, then the algorithm is a real tradeoff between history and today.

Then an ant or a data packet is sent out randomly using the following formula:

$$P_n = \frac{g_n}{\sum_{i \in N_k} g_i} \quad (6)$$

Where P_n, the probability of choosing neighbor n as the next node for the destination d, ranges from 0 to 1.

3. SIMULATION STUDY

3.1 Evaluation Methodology

In this section, we compare IF-AntNet with other popular algorithms including DOR, Duato's algorithm, and GAL. The simulations have been carried out on 8-ary 2-cube network as well as 16-ary 2-cube network by OPNET simulation software [13]. Due to space limitations, we only present the results of 8-ary 2-cube network. The results obtained for 16-ary 2-cube network are of similar trends.

Only virtual cut through switching mechanism [14] is used in the simulations, and a head flit is responsible for the job of a forward ant. So the completed algorithm also applies to the wormhole switching [18].

Packet length distribution is a specific distribution SP (Size and Percent) based on the IP (Internet Protocol) packet size and percentages sampled over a two-week period [15], shown in the following table.

Table. 1. Packet length distribution

Packet Size (bytes)	Percent (%)
40	56
52	4.5
576	16.5
1500	23

The following traffic patterns are used here:

1) The uniform traffic pattern: each node sends packets to all the nodes in the network with the same probability.

2) The hotspot traffic pattern: one node is chosen randomly to be the hotspot node, which receives 10% more traffic than all other nodes.

3) The tornado traffic pattern: the node (i, j) only sends packets to node $(i, (j + [k/2] - 1) \bmod k)$, with k being the radix of the network.

The nodes operate asynchronously. They generate packets at the time intervals following the negative exponential distribution. And packets arriving at a destination node are consumed immediately.

Four virtual channels are used for each algorithm for a fair comparison. The value 0.3 for γ, 0.2 for α, 0.8 for β, which is experimentally tuned to be the best, is used in the simulations.

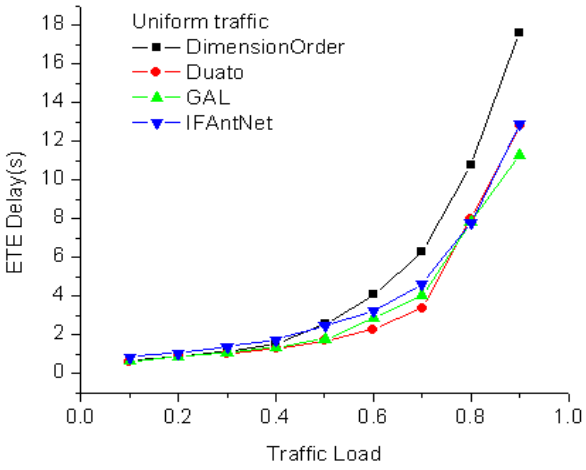
The performance of a routing algorithm is evaluated in terms of two main metrics described below:

1) Average ETE (End-To-End) delay: the mean time from the time a packet is generated to the time it reaches its destination

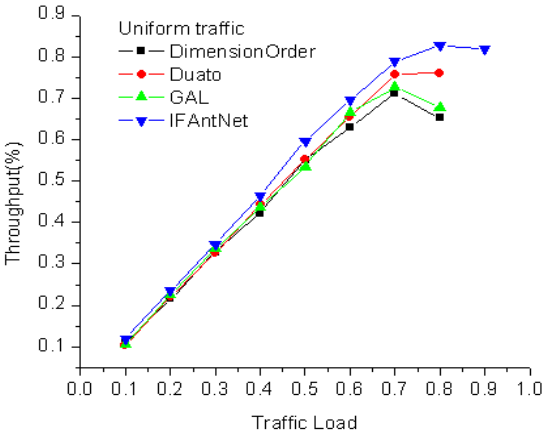
2) Normalized network throughput: the current network throughput divided by network capacity.

3.2 Simulation results

Fig. 1 shows the performances of the four algorithms under uniform traffic. The four algorithms yield almost identical performances under low traffic loads. However, the gap between them broadens as traffic loads increase. DOR, a deterministic minimal routing algorithm, reaches its saturation point firstly. GAL and Duato are both adaptive algorithms following after DOR. IF-AntNet achieves a good performance in throughput, due to the fact that it tries to build many optimal paths between a source-destination pair. However, it has a slightly larger ETE delay compared with GAL and Duato, for it needs time to converge to optimal paths.



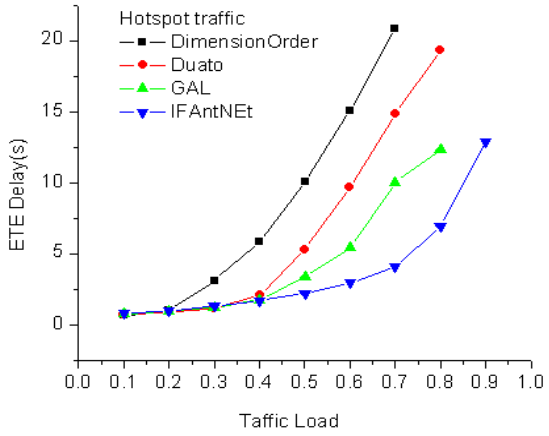
(a) ETE Delay



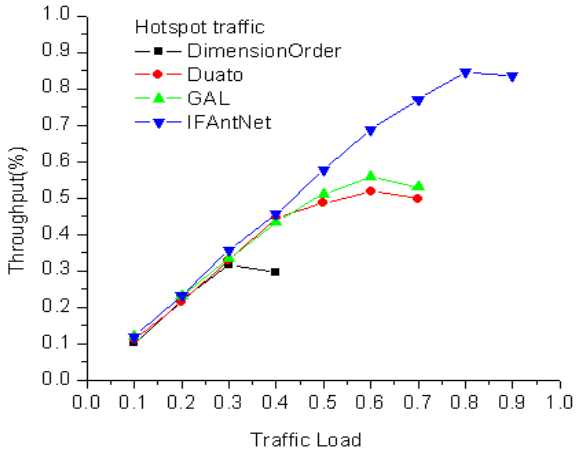
(b) Throughput

Fig. 1. Performances of the four algorithms under uniform traffic

What Fig. 2 shows is the comparison of the four algorithms under hotspot traffic. DOR gives a poor performance under the hotspot traffic, mainly because it always uses the same path between a source and destination node regardless of network conditions. Duato is the second with its saturation point of 0.4, due to its minimal adaptive routing method to bypass the hotspots. GAL has its saturation point 0.1 lower than Duato, because it can turn to a non-minimal path to avoid congested areas when it detects a congested point. IF-AntNet turns out to be the best one, with the lowest latency and the highest throughput. The reason lies in its stochastic routing decision concerning both the information in the trip-time table and that of the current network condition.



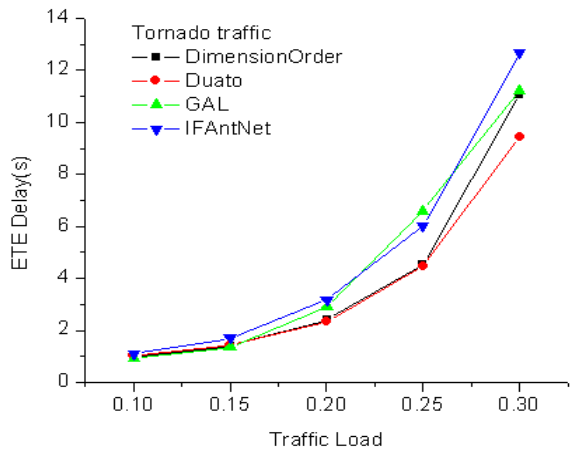
(a) ETE Delay



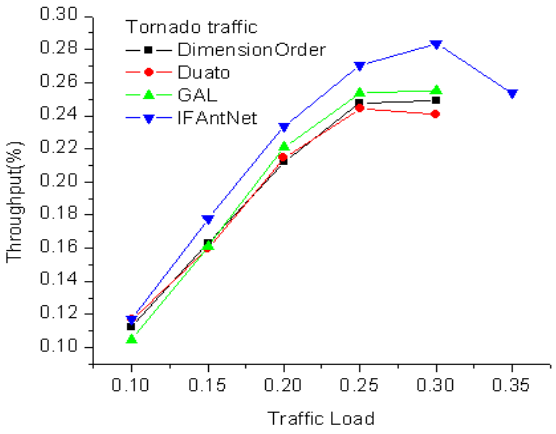
(b) Throughput

Fig. 2. Performances of the four algorithms under hotspot traffic

The performances of the four algorithms under tornado traffic are shown in Fig. 3.. As can be seen, DOR and Duato perform poorly under the tornado traffic for their consideration of only minimal routing. GAL achieves a good performance under high traffic load, due to its adoption of a non-minimal path to relieve the network congestion. IF-AntNet, also a minimal algorithm, though bad in ETE Delay, dramatically outperforms others in throughput, due to its efforts to try to build many optimal paths between a source-destination pair.



(a) ETE Delay



(b) Throughput

Fig. 3. Performances of the four algorithms under tornado traffic

4. CONCLUSIONS

This paper proposes a new load-balanced algorithm IF-AntNet for Direct interconnection networks. In this algorithm, only forward ants are used to collect network information and update the routing information of the source node on each intermediate node during their trip. Here, the routing information refers to the trip time from current node to the source node, which is estimated by a precise model including all possible delays from transmission delay, propagation delay to queuing delay. Both ants and data packets are routed based on the information of the trip-time table as well as the current status of queues for outputting. The simulation results show that the algorithm is rather competitive with traditional algorithms such as DOR, Duato's algorithm and GAL.

However, minimal routing is used in this algorithm, which is destined to suffer from some adversarial traffic such as tornado. In near future, improvements will be made to address this problem.

5. REFERENCES

[1] E. Bonabeau, M. Dorigo, et al, "Swarm intelligence: from natural to artificial systems", Oxford University Press, 1999.

[2] M. Dorigo, G. D. Caro, et al, "Ant algorithms for discrete optimization," Artificial Life, 1999 5(2):137-172.

[3] J. L. Deneubourg, S. Aron, et al, "The self-organizing exploratory pattern of the argentine ant," Journal of Insect Behavior, 1990, 3:159-168.

[4] G. D. Caro, M. Dorigo, "AntNet: Distributed stigmergetic control for communications networks," Journal of Artificial Intelligence Research, 1998, 9:317-365.

[5] S. Scott, G. Thorson, "The crayt3e network: adaptive routing in a high performance 3d torus," In Proceedings of Hot Interconnects Symposium IV, 1996.

[6] G. Pfister, "An Introduction to the InfiniBand Architecture (http://www.infinibadta.org)", IEEE Press, 2001.

[7] W. J. Dally, P. Carvey, et al, "Architecture of the Avici terabit switch/router," In Proceedings of Hot

Interconnects Symposium VI, 1998, 41-50.

[8] K.M. Sim, W.H. Sun, "Ant colony optimization for routing and load balancing: Survey and new directions", IEEE Transactions on Systems, Man, and Cybernetics. Part A, 2003, 33(5): 560-572.

[9] W. Dally, B. Towles, "Principles and Practices of Interconnection Networks", Morgan-Kaufmann Press, 2004.

[10] H. Sullivan, T. R. Bashkow, "A large scale, homogeneous, fully distributed parallel machine, I," in Proceedings of the International Symposium on Computer Architecture, 1977, 105-117.

[11] J. Duato, "A New Theory of Deadlock-Free Adaptive Routing in Wormhole Networks," IEEE Transactions on Parallel and Distributed Systems, 1993, 4:1320-1331.

[12] A. Singh, W. J. Dally, et al, "Globally adaptive load-balanced routing on tori," Computer Architecture Letters, 2004, 3.

[13] OPNET Modeler documentation, OPNET Technologies, Inc. http://www.opnet.com/, 2004.

[14] P. Kermani, L. Kleinrock, "Virtual Cut through: A new computer communication switching technique," Computer Networks, 1979, 3(34).

[15] Newman, "Internet Core Router Test", On the Web at www.lightreading.com, March 6, 2001.

[16] R. Schoonderwoerd, O. Holland, et al, "Ant-like agents for load balancing in telecommunications networks," In Proceedings of the First International Conference on Autonomous Agents, ACM Press, 1997, 209-216.

[17] J.M. Mart´ınez, J. Duato, et al, "Software-Based Deadlock Recovery Technique for True Fully Adaptive Routing in Wormhole Networks," 1997 International Conference Parallel Processing, 1997.

[18] L. Ni, P. McKinley, "A Survey of Wormhole Routing Techniques in Directed Networks," Computer, 1993, 26:62-76.

Direction-Field Based Fast Block Matching Motion Estimation Technology*

Yigang Wang[1,2] , Huiming Tang[2] , Jialin Cui[1] , Bin Xu[1]
[1] Dept. of Information Science & Engineering, Ningbo Institute of Technology, Zhejiang University, Ningbo, Zhejiang 315000, China
[2] Dept. of information science & Electronic Engineering, Zhejiang University, Hangzhou, Zhejiang 310027, China
Email: wyg@nit.zju.edu.cn
Email: tanghm@zju.edu.cn
Email: cuijl_jx@163.com
Email: xbin@nit.zju.edu.cn

ABSTRACT

In a video encoder, the most time-consuming component is the motion estimation (ME) algorithm. To reduce the computational complexity of the ME algorithm, a direction field based fast search algorithm is proposed. The search range was reduced by using the previous motion vector (MV) information. The search times were also decreased by adding the initial predictor MV candidate. Thus, a large number of searching points are skipped according to the information of current direction field. The experimental results demonstrate that the performance of the proposed algorithm is comparable with the other algorithms such as FS, APDZS, MVFAST, and PMVFAST, in terms of both PSNR-Y and bit stream length, with less searching computational complexity.

Keywords: Video Compression, Motion Estimation, Direction Field, Block Matching, Quick Search.

1. INTRODUCTION

Block matching motion estimation and compensation is an important technique in video-coding standards, such as H.263/264 and MPEG-2/4, since it can effectively eliminate the temporal redundancy which exists in video sequences, thus achieving compression without reducing the visual quality of each video frame.

In order to acquire accurate estimation, the ideal method, named full search (FS), can be adopted to compare the current micro-block (MB) with each possible MB in the reference frame. The MB in the reference frame, which gets the minimum sum of absolute difference (SAD) with the current MB, is considered as the best MB. Unfortunately, it was reported that the FS algorithm can approximately consume up to 80% of the total computational time of the encoder [1].

Several techniques have been previously proposed in an effort to reduce the encoding complexity of the FS algorithm. These include Three-Step Search (TSS), New Three-Step Search (NTSS), and Diamond Search (DS) [2]. Although these algorithms help the search run faster, they usually incur a significant loss in image quality. As a comparison, the proposed algorithm in this paper is very fast, while maintaining the image quality.

* This work was supported by the Key Sciences Research Program of Zhejiang Province, China, under Grant No. 2005C21006.

Advanced Diamond Zonal Search (ADZS) [3] is another more effective algorithm. It adopts three techniques: early termination of the search, predictor motion vector (MV) candidate, and diamond zone searching.

Advanced Predictive Diamond Zonal Search (APDZS) [4] is an enhanced version of ADZS, by introducing two new features: the multiple initial predictor candidates, and the adaptive threshold calculation.

Motion Vector Field Adaptive Searching technique (MVFAST) introduced in [5] achieves a satisfactory performance by using some previously-proposed techniques, such as early termination at MV (0, 0), multiple motion vector candidate and modified DS searching around the point indicated by the best MV.

Predictive Motion Vector Field Adaptive Searching Technique (PMVFAST) [6] can provide an even better performance by combining the following two mechanisms: the adaptive threshold calculation of APDZS and the big and small diamond searching of MVFAST. The PMVFAST algorithm has been incorporated into the MPEG-4 optimization model.

In this paper, an algorithm that makes a further improvement on APDZS is proposed. In the new algorithm, the number of search point is reduced, thus, the encoding speed is improved. Also, both the achieved image quality and obtained length of the output bit stream are closed to those of the FS algorithm, by adding the candidate vector and relaxing the search range.

The rest of this paper is organized as follows. In Sec. 2, the APDZS algorithm is introduced. In Sec. 3, after analyzing the full search process of APDZS, the techniques used to improve ADPZS is presented. Finally, in Sec. 4 some simulation results and comparisons are presented.

2. THE APDZS ALGRITHM

The APDZS algorithm originally presented in [4] is derived from the Advanced Diamond Zonal Search (ADZS) algorithm. The working mechanism of APDZS is described as follows.

In the first step, the self-adaptive threshold *T1* and *T2* are determined and used as the following two criteria in early termination: 1)Criterion 1 -- if $MinSAD<T1$, then stop searching; 2)Criterion 2 -- if $MinSAD\geq T1$ and $MinSAD<T2$, then go to next iteration, and then terminated whether or not the better motion vector should be searched. The threshold

T1 is equal to Min(SAD_{left}, SAD_{top}, $SAD_{top\text{-}right}$), and the parameters SAD_{left}, SAD_{top}, and $SAD_{top\text{-}right}$ are the SAD of left MB, top MB, and top-right MB in current frame, respectively. $T2 = T1 + 256$.

In the second step, the best predictive motion vector is set. The predictive motion vector $MV_{predicted}$ is equal to Median (MV_{left}, MV_{top}, $MV_{top\text{-}right}$), where MV_{left}, MV_{top}, and $MV_{top\text{-}right}$ are the motion vector of left MB, top MB, and top-right MB in current frame, respectively.

In the third step, the search center is found. In the algorithm, it computes the corresponding SAD of $MV_{predicted}$ and takes $MV_{predicted}$ as the best MV. It stops if the particular stopping criteria are satisfied. If the criteria are not satisfied, it chooses the best vector as $MV_{current}$ within MV_{left}, MV_{top}, $MV_{top\text{-}right}$, and $MV_{collocated}$, and stops if the criteria are satisfied, otherwise, it starts to search from the center which is indicated by MV..

The fourth step is performing search. The algorithm adopts three steps to find the best MV. The first one is to search from the $MV_{current}$ which is acquired from previous discussion. The second one is to search from MV (0, 0). The third one is to search from the latest $MV_{current}$. The search will stop immediately if one of the two criteria is satisfied. The search pattern employed for the APDZS is shown in Fig. 1.

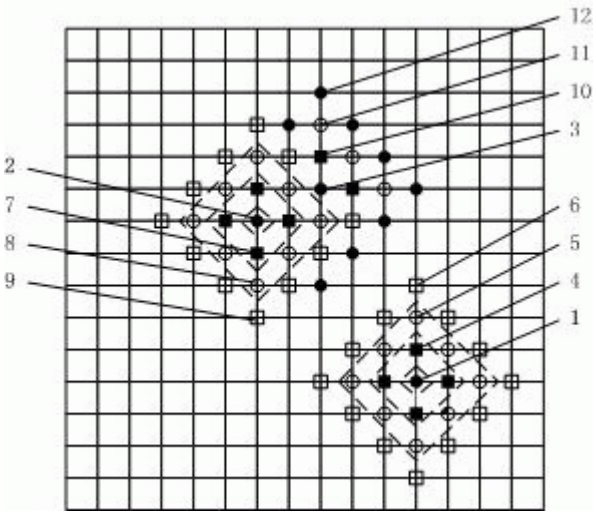


Fig. 1. APDZS search process

In Fig. 1., point 1 is the first search center which is acquired according to $MV_{current}$. Similarly, point 2 is the second search center, namely, MV (0, 0). Point 3 is the third search center acquired according to the latest best MV. Zone 4, 7 and 10 are the first search zones. Zone 5, 8 and 11 are the second search zones. Zone 6, 9, 12 are the third search zones.

3. PRINCIPLE AND REALIZATION OF IMPROVED APDZS ALGORITHM

The APDZS algorithm manages an improved performance versus the ADZS in terms of speed and PSNR-Y, because of the addition of the initial MV candidate and the adopting of the early termination thresholds *T1* and *T2*.

In this section, we will find that there is more redundancy in search process by analyzing the APDZS algorithm. Then, we will present an improved algorithm.

3.1 Definition of the Pattern of Search Result

For convenience purpose, we define some patterns of the searching results and the corresponding conditions according to the different results.

Pattern I is shown in Fig. 2.. We can not find any better point (MV) during the first search in the zone around point A, All the four points in the zone B are not better than point A. If we do not stop searching at this point, the second searching result will cover the following three conditions:

- Condition 1: one of the four points in zone C, shown in Fig. 2..
- Condition 2: one of the four points in zone D.
- Condition 3: no better point.

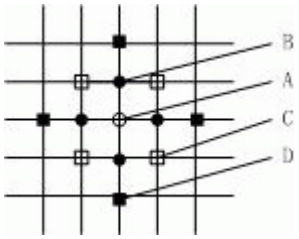


Fig. 2. Pattern I

Pattern II is shown in Fig. 3.. We find a better point during the first search in the zone B and E around point A. The current best MV is one of the four points around point A. We assume that it is the point B shown in Fig. 3.. If we do not stop searching at this point, the consequent second searching result will cover the following three conditions:

- Condition 1: one of the three points closest to point B, namely, zone C shown in Fig. 3.. If point B is in zone E, then, zone C will change accordingly.
- Condition 2: one of the five points near zone C, namely, zone D.
- Condition 3: no better point.

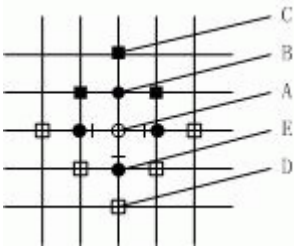


Fig. 3. Pattern II

Pattern III is shown in Fig. 4.. We find a better point during the second search in the zone near point B, and the best MV is located at one of the four points in zone C, i.e., top, bottom, left, and right relative to the results of first search in zone B. If we do not stop searching at this point, the third searching result will cover the following three conditions:

- Condition 1: one of the three points closest to the current best MV. The point is in zone E. As described earlier, zone E will change along with the changing of zone C.
- Condition 2: one of the nine points near three points in zone E.
- Condition 3: no better point.

Pattern IV is also shown in Fig. 4.. We find a better

point during the second search in the zone around point B. The best MV is located at one of the four points in zone D, i.e., top-left, top-right, bottom-left, and bottom-right relative to point A, respectively. If we do not stop search at this point, the consequent third searching result will cover the following three conditions:

- Condition 1: one of the two points closest to the current best MV, i.e., the point in zone F. As described earlier, zone F will change along with the changing of zone D.
- Condition 2: one of the ten points near the two points in zone F.
- Condition 3: no better point.

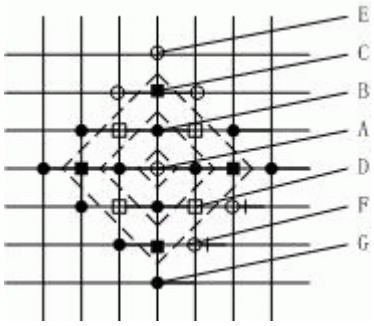


Fig. 4. Pattern III, IV

3.2 Analysis of searching process

The following Tables 1., 2. shows the search complexity of the APDZS algorithm.

The data shown in Table 1. are obtained from encoding according to the pattern classified earlier. Column two is the test video sequence. To obtain the results in different experimental conditions, we choose three sequences with motion activities varying from small to large. They are Akiyo, Carphone, and Football. All sequences are in common intermediate format (CIF) and are encoded at 30 frames per second. All frames except the first one are encoded as P-frames. Column 3 is the count of the total searching points, which represent the search complexity. Column 4 to 6 is the count of total searching points under the condition 1 to 3, respectively. Column 7 is equal to Cond. 1 / (Cond. 1 + Cond. 2).

From the results shown in Table 1., we find that all percentages are large except for the condition at pattern I of sequence football, i.e., the points counted of Cond. 1 occupy a large proportion of the total of Cond. 1 and Cond. 2. It indicates that there is high correlation between the latter best MV and the former best MV.

Table 1. The data of searching complexity of APDZS

Pattern	Object	Total count	Cond. 1	Cond. 2	Cond. 3	Percentage
I	Carphone	41104	1144	192	13827	85.6%
	Akiyo	13593	67	0	2794	100.0%
	Football	34727	1075	538	15983	66.6%
II	Carphone	19947	3459	445	5475	88.6%
	Akiyo	632	33	3	252	91.7%
	Football	46261	23104	1511	11933	93.9%
III	Carphone	2782	978	64	629	93.9%
	Akiyo	6	0	0	4	-
	Football	17435	8654	508	3070	94.5%
IV	Carphone	2458	336	66	1193	83.6%
	Akiyo	97	4	0	44	100.0%
	Football	8793	3158	491	3300	86.5%

For example, in pattern I we find that the probability of the best MV locating at zone C shown in Fig. 2. is up to 80% while the first search is failed and the second search is successful.

Similarly, for pattern II, III, and IV, the probability of the best MV which is located at the zone nearest to the former MV is up to 90%, 95% and 85%, respectively while the current search is successful.

In Table 2., the count of first search is the count of searches around $MV_{current}$. The count of second search is the search count around MV (0, 0). The percentage of the second search is equal to the count of the second search divided by the count of the first search. It is the probability of the first failed search.

Table 2. The data of searching count around MV (0, 0)

Object	Count of early stop	Count of first search	Count of second search	Percentage of second search
Carphone	37553	61051	4163	6.8%
Akiyo	103783	14225	95	0.7%
Football	21180	80988	24584	30.4%

In the APDZS algorithm, it will search around MV (0, 0) again if it is not satisfied with the early termination criteria after searching around $MV_{current}$. From the data shown in Table 2., we find that the probability of searching around MV (0, 0) is very small. The probability is less than 7% for the small motion cases.

3.3 Improved Algorithm

We find that there is more redundancy the APDZS algorithm by analyzing the searching process discussed earlier. As a result, we present the improved algorithm as follows.

- We only search the four points in zone C shown in Fig. 2. when performing the second search if the first search is failed.
- We only search the three points nearest to the best MV and perform the second search if the first search is successful. It is in zone C shown in Fig. 3..
- We only search the three points nearest to the best MV. It is zone E shown in Fig. 4.. We perform the third search if the second search is successful and the best MV is one of the four points of zone C shown in Fig. 4..
- We only search the two points nearest to the best MV. It is in zone F shown in Fig. 4.. We perform the third search if the second search is successful and the best MV is in zone D shown in Fig. 4..
- We add MV (0, 0) as the initial predictor candidate besides $MV_{predicted}$, MV_{left}, MV_{top}, $MV_{top\text{-}right}$, and $MV_{collocated}$. Here we inherit the idea of the PMVFAST algorithm.
- We cancel the second search around MV (0, 0) in the algorithm APDZS.

4. SIMULATION RESULTS OF THE PROPOSED ALGORITHM

The proposed algorithm was embedded in the MPEG-4 VM encoder software and was tested using the video sequences similar to the previous one. The quantification parameter is equal to 2 and the size of the search window is 16. The results are shown in Table 3.

In Table 3. the proposed algorithm is compared in terms of the number of search points, PSNR-Y and average length per frame versus the FS, APDZS, MVFAST and PMVFAST algorithm. From the results we can find that the search number of the proposed algorithm is about 50% less than those of APDZS for the high motion activity such as Football, while the search number of the proposed algorithm is about 80% less than those of APDZE for the low motion activity such as Akiyo. At the same time, the PSNR-Y and average length per frame are almost equal to those derived from the APDZS algorithm.

Table 3. The simulation results

Object	Algorithm	Count of search point	PSNR-Y(dB)	Bitstream length per frame(B)
	FS	25242624	41.31	13492
	APDZS	708221	41.18	13727
Carphone	MVFAST	716344	41.11	13788
	PMVFAST	561704	41.16	13776
	Proposed	482181	41.16	13776
	FS	30210048	43.69	3469
	APDZS	96021	43.65	3501
Akiyo	MVFAST	158512	43.65	3506
	PMVFAST	70324	43.65	3504
	Proposed	52354	43.65	3504
	FS	26155008	42.00	27246
	APDZS	1260910	41.99	27205
Football	MVFAST	1041487	41.88	27289
	PMVFAST	961565	41.86	27300
	Proposed	853757	41.89	27246
	FS	30210048	42.70	7175
	APDZS	265113	41.67	7495
Mother	MVFAST	395408	41.63	7551
	PMVFAST	273893	41.67	7520
	Proposed	227601	41.67	7496
	FS	25242624	40.69	36319
	APDZS	799356	40.67	36230
Funfair	MVFAST	747899	40.67	36228
	PMVFAST	542119	40.67	36257
	Proposed	492630	40.67	36216

We also notice that the average search number is about 18% less than the PMVFAST algorithm while the PSNR-Y and the average length per frame are nearly the same to those of PMVFAST algorithms adopted by the MPEG-4 standard.

5. CONCLUSION

In this paper, we propose a new highly efficient algorithm, direction field base fast block matching motion estimation algorithm. The algorithm combines two techniques, the reduction of searching range based on the direction-field, and adopting MV (0, 0) as one of the predictor motion vector candidates in APDZS. Our simulation results demonstrate the superiority of the proposed algorithm by comparing the other fast motion estimation algorithms such as APDZS, MVFAST and PMVFAST, in terms of PSNR-Y, the number of search points, and the average bit stream length per frame. The proposed algorithm can produce similar visual quality without longer bit stream length compared with the other algorithms including FS.

6. REFERENCES

[1] A.M. Tourapis, O.C. Au, et al, "Status Report of Core Experiment on Fast Block-Matching Motion Estimation using Advanced Diamond Zonal Search with Embedded Radar," ISO/IEC JTC1/SC29/WG11 MPEG99/m4980, Melbourne, Australia, Oct. 1999

[2] J.Y. Tham, S. Ranganath, et al, "A Novel Unrestricted Center-Biased Diamond Search Algorithm for Block Motion Estimation," IEEE Trans. Circuits & Systems for Video Technology, 1998, 8(8)： 369-377.

[3] A.M. Tourapis, O.C. Au, et al, "Optimizing the MPEG-4 Encoder – Advanced Diamond Zonal Search," Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium On Volume 3, 2000, 3(5)：674-677.

[4] A.M. Tourapis, O.C. Au, et al, "New Results on Zonal Based Motion Estimation Algorithms – Advanced Predictive Diamond Zonal Search," Circuits and Systems, 2001. ISCAS 2001. The 2001 IEEE International Symposium On Volume 5, 2001, 5(5)：183-186.

[5] P.I. Hosur, K.K. Ma, "Motion vector field adaptive fast motion estimation," in Proc. 2nd Int. Conf. Information, Communications and Signal Processing (ICICS '99), Dec. 1999

[6] A.M. Tourapis, O.C. Au, et al, "Predictive Motion Vector Field Adaptive Search Technique(PMVFAST)-Enhancing Block Based Motion Estimation," Proceedings of Visual Communications and Image Processing 2001 (VCIP 2001), 2001, 883-892.

Hybrid Concatenation of Single Parity Check Codes and Performance Analysis

Jiaxuan Yang, Chuanying Jia and Guoyou Shi
Institute of Nautical Science and Technology, Dalian Maritime University
Dalian, Liaoning Province, China
Email: ytyjx@163.com, chuanyingjia@163.com, shiguoyoudmu@163.com

ABSTRACT

Concatenated codes not only have the serial and parallel concatenated coding structures, but also have the uncommon hybrid one. Single parity check codes were utilized as constituent codes to be considered as a uniform way of three coding structures. A uniform description of the three coding structures was illustrated. A kind of generalized iterative decoding structure, which was suitable for the three coding structure, was also presented. The results show that different coding structures have difference performance but have regular variation respectively. The performance of various hybrid concatenated single parity check codes is always lying between that of parallel and serial ones. When code rate increases, the performance crossover between parallel and serial concatenated single parity check codes moves to lower bit-error rate (BER) region and the difference between their performances is decreases.

Keywords: Channel Coding, Hybrid, Single Parity Check Codes, Performance Analysis.

1. INTRODUCTION

Single parity check (SPC) codes are a very simple kind of linear block codes. Recently, many scholars have used them as component codes to construct multiple (or multi-dimension) concatenated SPC codes (MC-SPC) [1，2]. MC-SPC have shown good performance both in AWGN and independent Rayleigh fading channels with BPSK or higher modulation schemes such as 16-QAM, but are relatively simple to encode and decode [3-4]. The main advantages of these codes are: high coding efficiency, low decoding complexity, flexible block length and code rate, and improved performance by increasing the dimensionality (or the number of component codes) without changing the block length and overall code rate etc. [4]. On the theoretical research, [5] proves this class of codes can achieve performance close to the Shannon limit asymptotically.

Generally, in order to obtain high performance mentioned above, two commonly used concatenated coding schemes, i.e., parallel and serial concatenated structures are applied in MC-SPC. However, in [4] the author pointed out that with the same dimensionality, multiple parallel concatenated SPC codes (M-PC-SPC) show superior performance to the multiple serial concatenated SPC codes (M-SC-SPC) at low-to-mode- rate signal-to-noise ratios (SNRs), but when at higher SNRs, they began to encounter error floors and M-SC-SPC started to show superiority. Actually, concatenated codes have not only the generally known serial and parallel-concatenated coding structures, but also an uncommon hybrid coding structure, which combines aforementioned two coding structures together. Although an adequate discussion about hybrid concatenation strategy is absent, [6] illustrates this kind of codes can also achieve good performance as well as their parallel and serial counterparts.

With the aim to further investigate the impact of different concatenation strategies on coding performance and also to find a performance versus complexity tradeoff between parallel and serial concatenated codes, utilizing SPC codes as constituent codes, we consider the three coding structures above in a uniform way. A uniform description of parallel, serial and hybrid concatenated coding structures is illustrated in this paper, and we also propose a generalized iterative decoding structure suitable for the three coding structures which is approved by the simulation results. By simulation results on multiple concatenated SPC codes, we find that in the equal decoding system, different coding structures have different but regular performance variation. The performance of various hybrid concatenated SPC codes is always lying in between the performance of parallel and serial concatenated SPC codes and when code rate increases, the crossover between the latter two and their performance difference also shows regular variation.

The paper is organized as follows. The unified encoding structure of MC-SPC and its generalized decoding structure are described in Part 2. Simulation results are presented in Part 3. Finally, conclusions are given in Part 4**.**

2. HYBRID CONCATENATED SPC CODES

2.1 Unified Encoding

Fig. 1 shows the unified encoding structure of MC-SPC including parallel, serial and hybrid concatenation. Let

$$\sum_{i=1}^{m} l_i = n \qquad (1)$$

represent n SPC encoders in Fig.1, following the constraint condition

$$\begin{cases} 1 \le l_i \le n & (1 \le i \le m) \\ 1 \le m \le n \end{cases} \qquad (2)$$

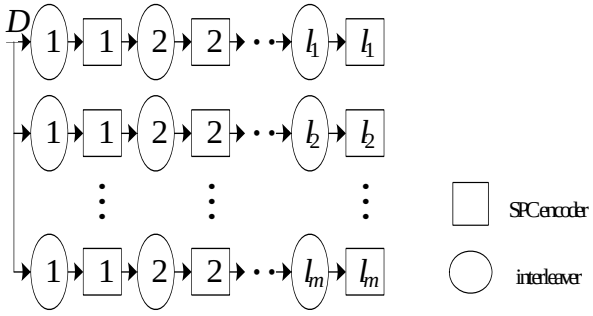


Fig. 1. Unified encoding structure of MC-SPC

We can describe parallel and serial concatenation as $m = n, l_i = 1$, $i = 1, 2, \cdots, m$ and $m = 1, l_1 = n$ respectively [7]. Obviously parallel and serial concatenations are two extreme cases under the unified encoding structure. When variables m and l_i take different values within the constraint condition and comply with Eq. (1), hybrid concatenation is formed. For concrete examples, we give several hybrid concatenated coding structures where $n = 3$ and $n = 4$ shown in Fig. 2.

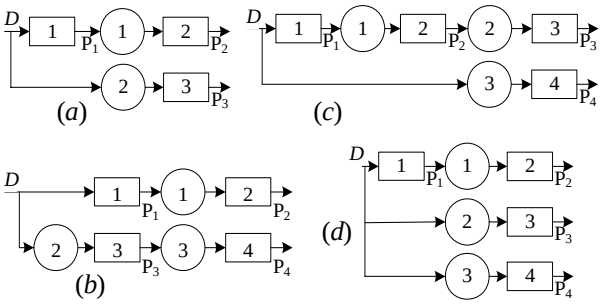


Fig. 2. Hybrid concatenated SPC encoding structures
(a) 3-HC-SPC, (b)4-HC1-SPC, (c) 4-HC2-SPC (d) 4-HC3-SPC

Apparently, after determining the total number of encoders, we can obtain more than one hybrid concatenated coding structures due to the different combination types of encoders.

Furthermore, the number of interleavers in Fig.1 is equal to the number of encoders, however, in general the first interleaver in the first branch can be considered as a special case of no interleaving, but the other interleavers can be of any type. In this paper, random interleavers were investigated. Let D be an input information matrix with length $K = I \times J$ and $\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_n$ be the parity check vectors generated by n SPC encoders. Here, we use the same rules as in [3] to produce these vectors, so the transmitted codeword consists of D and $\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_n$. From Fig.2 it can be seen that with the same number of component codes, all hybrid concatenated structures only contain a middle quantity of parity-on-parity bits between parallel and serial concatenation schemes, therefore the encoding process of hybrid concatenation is more complex than parallel concatenation but less than serial case. The parameter of the first $(J+1, J, 2)$ SPC encoder in the first branch in Fig.1 can be set as [4]

$$J = \frac{K}{N-K} \times M = \frac{R}{1-R} \times M \qquad (3)$$

where K is the length of input information bits, N is the length of output coded bits, M represents the number of concatenations, and the overall code rate $R = K / N$. When having the first parameter J set, we can adjust the other parameters of SPC encoders in terms of the coding structure appropriately. Take 4-HC1-SPC as an example, let SPC1= $(J+1, J, 2)$, then SPC2= $(J+2, J+1, 2)$, SPC3= $(J+1, J, 2)$ and SPC4= $(J+2, J+1, 2)$.

2.2 Generalized iterative decoding structure

Although MC-SPC has various encoding structures, one generalized decoding structure can be used for their iterative decoding. Fig. 3 shows the generalized global iterative decoding structure, where the decoder is a soft-in-soft-out unit performing symbol-by-symbol MAP decoding algorithm [8]. The arrangement of decoders is serial according to the encoding structure, starting with the first encoder in the first branch and ending with the last encoder n. In addition, for the sake of simplicity, the interleaver and de-interleaver parts are omitted in Fig. 3.

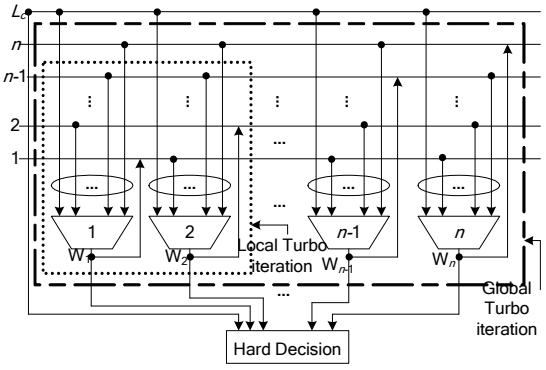


Fig. 3. Generalized global iterative decoding structure for MC-SPC

In Fig. 3, each component decoder combines log-likelihood ratio (LLR) values L_c from the demodulation process and a- priori information L_i to perform iterative decoding [3]. A-priori information for each decoder comes from the summation of the extrinsic information $\mathbf{W}_i$ from all other decoders, which prevents the self-feedback of extrinsic information generated by itself. If let f represent the decoding algorithm for decoders and m denotes the m-th iteration, the iterative decoding steps for n multiple concatenated SPC codes can be concluded as

1) Initiation, let $\mathbf{W}_i^0 = 0$, $i = 1, 2, \cdots, n$

2) Calculate $\mathbf{Li}_i^m = \sum_{k<i} \mathbf{W}_k^m + \sum_{k>i} \mathbf{W}_k^{m-1}$, $i = 1, 2, \cdots, n$, $k = 1, 2, \cdots, n$

3) Summate $\mathbf{X}_i^m = \mathbf{Lc} + \mathbf{Li}_i^m$ and compute $\mathbf{W}_i^m = f\left(\mathbf{X}_i^m\right)$, $i = 1, 2, \cdots, n$

4) Repeat 2 and 3 until achieving the iterative decoding stop criterion or iteration times, then make hard decision on soft values $\mathbf{\Lambda}$. The expression of $\mathbf{\Lambda}$ is

$$\mathbf{\Lambda} = \mathbf{Lc} + \sum_{i=1}^{n} \mathbf{W}_i \qquad (4)$$

The decoding structure in Fig.3 is originally used for the decoding of multiple parallel turbo codes [10], in [5] the author extended it to M-SC-SPC. Now, we explain the reason that it is also suitable for the decoding of hybrid-concatenated codes. Take 3-HC-SPC as an example. In this encoding structure, there are two parallel branches where two encoders are serial concatenated in the first branch and one single encoder in another branch. If we view the two encoders in the first branch as a supercode encoder and since such a supercode encoder with another single SPC encoder constructs the two-parallel concatenated coding structure, it is natural to decode it with the generally known turbo decoder, which is called global turbo decoder here. Within the supercode, since it is serial concatenated, we can apply another iterative decoder to decode it, which is called local turbo decoder here. In order to reduce the decoding

complexity and also be comparable to the decoding of parallel and serial counterparts, we can simplify the whole decoding process to perform only once local turbo iteration and several global turbo iterations, and hence it becomes the generalized iterative decoding structure. The above analysis is also suitable for other hybrid concatenated SPC codes and the assumption is approved by our simulation results.

On the decoding complexity aspect, in general it is simple to decode MC-SPC because of the low complexity of MAP decoding for the constituent SPC code. Using MAP decoding algorithm, the decoding complexity is proportional to the length of both information and parity check bits received by each decoder. Assume the length of information bits is J, the concatenation dimensionality is M, then the length of total received bits for M-PC-SPC in once decoding operation is $M(J+1)$, M-SC-SPC is $\sum_{i=1}^{M}(J+i) = MJ + \sum_{i=1}^{M} i$, and hybrid concatenation is $MJ + p$, where $M < p < \sum_{i=1}^{M} i$, due to the middle quantity of parity-on-parity bits. So it is expected that the decoding complexity of all hybrid concatenated SPC codes is between their parallel and serial counterparts.

Furthermore, in the decoding process, the passing of extrinsic information of parity check bits to the different decoders and interleaving, de-interleaving of extrinsic information of all bits should be carefully considered.

3. SIMULATION RESULTS

In this seFig.4 and Fig.5 are computer simulation results. In our simulation we set the code rate R =0.75 and the input information length K =900 bits constantly. The soft-in-soft-out decoder adopts the modified Max-Log-MAP algorithm in [11]. The global iterative decoder is the proposed generalized global iterative decoding structure with 5 iterations. We use BPSK modulation and AWGN channel in simulation.

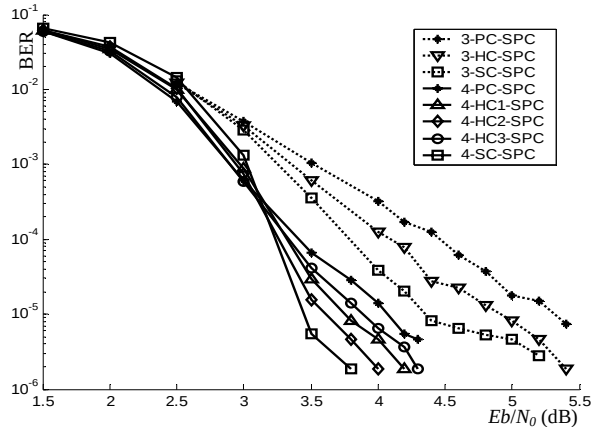


Fig. 4. Simulation results for MC-SPC codes

Fig.4 shows the performance comparison of MC-SPC under different coding structures. As it can be seen from the figure, an obvious performance crossover is between 4-SC-SPC and 4-PC-SPC. Parallel-concatenated SPC codes outperform serial concatenated SPC codes before the crossover but after the crossover serial concatenated SPC codes perform better. All the performance curves of the proposed hybrid concatenated SPC codes in Fig.2 lie in between the performance curves of their serial and parallel counterparts. We consider the reason as follows: since no parity check bits of parallel concatenate codes participate in interleaving but all the parity check bits of serial concatenated codes take part in the process, when in the SNR region before the performance crossover, it is likely for the interleaving parity check bits to offer the negative information (under statistical sense) which deteriorates the iterative decoding performance; but when in the SNR region beyond the crossover, the situation is just the opposite. For hybrid concatenation, since only partial parity check bits take part in interleaving, their performance tends to lie in between serial and parallel concatenation.

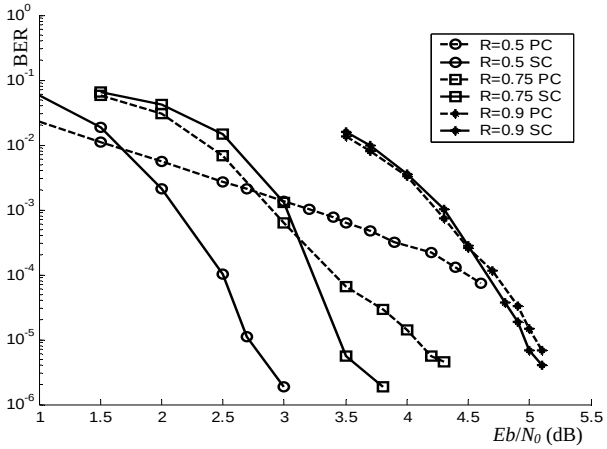


Fig. 5. Performance comparison for serial and parallel-concatenated SPC codes at different code rates

Fig.5 compares the performance of 4 serial and parallel-concatenated SPC codes at different code rates. Code rates are set to 0.5, 0.75 and 0.9. According to Fig. 5, when code rate increases, the performance crossover between serial and parallel concatenation moves to lower BER region gradually. Parallel-concatenated SPC codes encounter severe error floors at low code rates but become better when code rate increases. Serial concatenated SPC codes have significantly better performance at all code rates. In addition, at high code rates, the performance between serial and parallel SPC codes has become very close, so it is better to choose M-PC-SPC when high code rates are needed because of their lowest decoding complexity among multiple concatenated SPC codes.

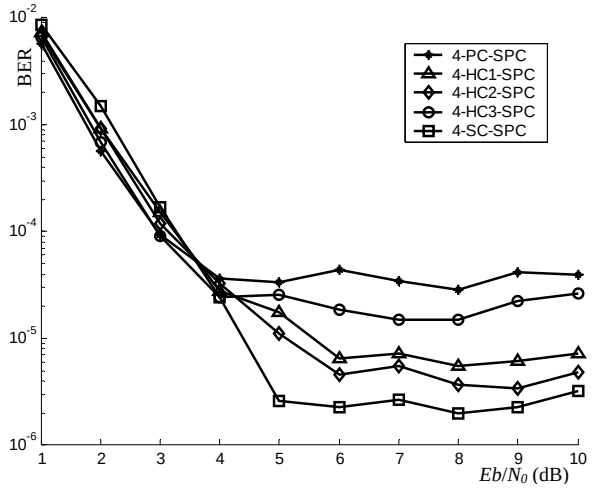


Fig. 6. Convergence of MC-SPC with different encoding

structures.

Fig. 6 illustrates the convergence behavior of MC-SPC with different encoding structures under a channel SNR of $E_b / N_0 = 0.6dB$. From the figure, we can see that after 5 iterations, 4-SC-SPC achieve the lowest error floors among 4 concatenated SPC codes and 4-PC-SPC achieve the highest among them. The error floors of all hybrid concatenated SPC codes lie in between those of M-SC-SPC and M-PC-SPC. It is noticed that for the first 3 iterations, the performance in the figure is not the same as in Fig. 4. We believe this is caused by insufficient iterations. The interleaved parity check bits are likely to provide negative information which deteriorates the iterative decoding performance under insufficient iterations so that 4-PC-SPC show the best performance up to first 3 iterations because they have no interleaved parity bits at all. Since all codes well converge after 5 iterations, limiting the decoding at five iterations is enough for practical applications.

4. CONCLUSIONS

A unified encoding structure of serial, parallel and hybrid MC-SPC and a generalized iterative decoding structure suitable for the unified encoding have been presented. The results show that M-SC-SPC outperform M-PC-SPC under certain conditions; the performance of hybrid concatenated SPC codes is lying in between their serial and parallel counterparts and can be a good performance versus complexity tradeoff between the latter two.

5. REFERENCES

[1] J.S.K. Tee, D.P. Taylor, “Multiple parallel concatenated single parity check codes,” 2001 IEEE International Conference on Communications, 2001,1: 60-64.

[2] D.M. Rankin， T.A. Gulliver., “Single parity check product codes,” IEEE Trans. on Communications, 2001, 49:1354-1362.

[3] P. Li, S. Chan, et al, “Iterative decoding of multi-dimensional concatenated single parity check codes,” 1998 IEEE International Conference on Communications, 1998, 1:131-135.

[4] J.S.K. Tee, D.P. Taylor, et al, “Multiple serial and parallel concatenated single parity check codes,” IEEE Trans. on Commu- nications, 2003, 51:1666-1675.

[5] D.W. Yue, E. H. Yang, “Asymptotically Gaussian weight distribution and performance of multi-component turbo block codes and product codes,” IEEE Trans. on Communications, 2004,52:728-736.

[6] D. Divsalar, F. Pollara, “Hybrid concatenated codes and iterative decoding,” The Telecommunications and Data Acquisition Progress Report, Jet Propulsion Laboratory, Pasadena, California, 1997.

[7] S. Benedetto, G. Montorsi, “Unveiling turbo codes: some results on parallel concatenated coding schemes,” IEEE Trans. on Inform. Theory, 1996,42:409-428.

[8] S. Benedetto, G. Montorsi, et al, “Serial concate- nation of interleaved codes: Performance Analysis, Design, and Iterative Decoding,” IEEE Tran. on Inform. Theory, 1998,44: 909-92.

[9] S. Benedetto, D. Divsalar, et al, “Analysis, design and iterative decoding of double serially concatenated codes with interleavers,” IEEE Journal on Selected Areas in Communications, 1998, 16:231-244.

[10] D. Divsalar, F. Pollara, “Multiple turbo codes,” in Proc. IEEE MILCOM’ 95, San Diego, CA, 1995,1: 279-285.

[11] N. Stojanovic, A. Farbert, et al, “Concatenated three-dimensional SPC turbo codes for optical transmission systems,” 5th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services, Telsiks, 2001, 2:504-507.

Efficient and Secure Quantum All-or-Nothing Disclosure of Secrets Scheme*

Qiyan Wang[1], Liusheng Huang[2], Wei Yang[3], Yonglong Luo[4]
Department of Computer Science and Technology, University of Science and Technology of China,
National High Performance Computing Center (Hefei), Hefei, 230027, P.R.China.
[1]Email: xxwqy@mail.ustc.edu.cn [2]Email: lshuang@ustc.edu.cn
[3]Email: smartyw@mail.ustc.edu.cn [4]Email: ylluo@ustc.edu

ABSTRACT

In an All-or-Nothing Disclosure of Secrets protocol, Alice has *n t*-bit secrets and is willing to disclose one of them to Bob, with guarantee that no information about the other secrets will be gained. Moreover, Bob can freely choose his secret and has a guarantee that Alice will not be able to find out which secret he picked. We stress that previous constructions for this cryptographic task are not desirable in efficiency and security. In this paper, we present a new Quantum All-or-Nothing Disclosure of Secrets scheme, which is of high efficiency and unconditional security for both participants.

Keywords: All-or-Nothing Disclosure of Secrets, Oblivious Transfer, Bit Commitment, Privacy Amplification, Quantum Information.

1. INTRODUCTION

The concept of All-or-Nothing Disclosure of Secrets (*ANDOS*) was first introduced in 1986 by Brassard, Crépeau and Robert in [1]. It involves two parties, a vendor Alice and a buyer Bob, and allows Alice, who holds several secrets, to disclose one of them to Bob, with guarantee that no information about the other secrets will be gained. Furthermore, Bob can freely choose his secret and has a guarantee that Alice will not be able to find out which secret he picked. *ANDOS* has many applications, such as zero-knowledge proof, secrets exchange, mental poker and message authentication. In classical environment, *ANDOS* can be constructed based on the existence of factoring or discrete logarithm difficulty assumptions [2, 3, 4]. However, with Shor's discovery of polynomial-time quantum algorithms for factoring and discrete logarithm problems [5], classical *ANDOS* will be no longer secure if quantum computers become available.

Quantum All-or-Nothing Disclosure of Secrets (*QANDOS*) can realize all functionality of its classical couterpart, and it ensures the security by physical laws independent of difficulty assumptions. In 1997, Adrian Kent proposed the first *QANDOS* scheme in [6], utilizing unconditionally secure Quantum Bit Commitment (*QBC*). However, Mayers [7] subsequently proved that previous *QBC* protocols were not unconditionally secure. As a result, Adrian Kent's protocol is insecure, and furthermore its communication complexity is not ideal either.

*Supported by the National Natural Science Foundation (No.60573171), the National Grand Fundamental Research 973 Program of China (No. 2003CB317000), and the Research Program of Anhui Province Education Department (No.2006KJ024A and No. JYXM2005166)

ANDOS is the natural extension of 1-out-of-2 oblivious transfer (OT^l_2), in which Alice has two secret bits instead of strings. In 2004, Z. Chen put forward a scheme to construct QOT^m_n [8], however which is not perfectly secure due to Alice's non-negligible probability to infer Bob's selections without detection. Even so, this approach gives us some illumination in our work. Afterwards, Z. Chen presented a new *QANDOS* protocol in [9], depending upon the reduction relationship between *ANDOS* and OT^l_2 [1]. Unfortunately, cheating Bob can build $|\varphi_1'\rangle = (|0\rangle|B_{00}\rangle + |1\rangle|B_{11}\rangle)/\sqrt{2}$ instead of $|\varphi_1\rangle$ to get $b_1 \oplus b_2$ (b_1, b_2 are Alice's two secret bits) in its sub-protocol QOT^l_2. Thus, this *QANDOS* is vulnerable because in classical cryptography the exclusive -or of two plaintext English messages allow easy recovery of them both. In addition, the efficiency of the protocol is not desirable either, for repeating QOT^l_2 for many runs.

In 2005, H. P. Yuen pinpointed the gaps in Mayer's proof of the impossibility of unconditionally secure *QBC* and proposed an unconditionally secure *QBC* protocol in [10]. Owing to this result, we construct a new *QANDOS* scheme in this paper, which is more efficient and secure than previous ones [6, 9]. In particular, our scheme is immune to any technologically feasible attacks, regardless of the computing power available to would-be cheaters. In addition, we utilize privacy amplification to ensure the relatively high efficiency of our protocol, which involves QOT^m_n only once. Moreover, we don't make any restrictive assumptions on either party's computing power. Thus, our scheme is of unconditional security.

This paper is organized as follows. First, some preliminaries and definitions are given in section 2. Then, we describe a new *QANDOS* scheme in section 3, and prove its correctness and security in section 4. Finally, section 5 is the conclusion of this paper and our further work.

2. PRELIMINARIES

We suppose that readers have fundamental knowledge about quantum computation and quantum information.

2.1 Privacy Amplification

Privacy amplification is a technology of distilling highly secret shared information, perhaps for use as a cryptographic key, from a larger body of shared information that is only partially secret. Let Alice and Bob be given a random variable W, such as a random n-bit string, while an eavesdropper Eve learns a correlated random variable V, providing at most $t<n$ bits of information about W, i.e. $H(W|V) \geqslant n-t$. The details of the distribution P_{vw} are generally unknown to Alice and Bob, except that it satisfies this constraint as well as possibly some further constraints. They may or may not know P_w. Alice and Bob wish to

publicly choose a compression function g: $\{0,1\}^n \to \{0,1\}^r$ such that Eve's partial information on W and her complete information on g give her arbitrarily little information about $K=g(W)$, except with negligible probability (over possible choices for g). The resulting K is virtually uniformly distributed given all Eve's information, and hence it can be used safely as a cryptographic key.

2.2 Bit Commitment (*BC*)

Assume that Alice has a bit x in mind, to which she would like to be committed toward Bob. That is, Alice wishes, through a procedure commit(x), to provide Bob with a piece of evidence that she has a bit x in mind and that she cannot change it. Meanwhile, Bob should not be able to tell from that evidence what x is. At a later time, Alice can reveal, through a procedure unveil(x), the value of x and prove to Bob that the piece of evidence sent earlier really corresponded to that bit. For a *BC* protocol, there are two security requirements:

- *Binding*: no matter what Alice does, it is impossible to open both 0 and 1 with non-negligible probability of success.
- *Concealing*: Bob cannot obtain more than a negligible amount of information about the committed bit without Alice's help.

3. UNCONDITIONALLY SECURE *QANDOS* SCHEME

In this section, we give a new construction for *QANDOS,* as well as QOT^N_{nN} contained in it. In order to maintain the paper's consecution, we place the detailed introduction of QOT^N_{nN} behind as a sole part. Let $[\ ]_i$ denote the selection function such that $[a_0, a_1,\ldots, a_n]_i=a_i$. Also, let $+ = \{|\rightarrow\rangle, |\uparrow\rangle\}$ and $\times = \{|\nearrow\rangle, |\searrow\rangle\}$ denote the basis of rectilinear and diagonal polarization in the quantum state space of a photon respectively. In addition, we let all random variables used in the protocol be uniformly distributed. Let the compression function at step II be a class of universal hashing $\boldsymbol{G}$ [11].

3.1 *QANDOS* Scheme

Suppose that Alice has n t-bit secrets: $\vec{S}_j = \{s_1^{(j)}, s_2^{(j)},\ldots, s_t^{(j)}\}$ ($1 \leqslant j \leqslant n$, $s_i^{(j)}$ is the i^{th} bit of $\vec{S}_j$), out of which Bob intends to choose the secret $\vec{S}_c = \{s_1^{(c)}, s_2^{(c)},\ldots, s_t^{(c)}\}$.

$QANDOS$ ($\vec{S}_1, \vec{S}_2,\ldots, \vec{S}_n$) (c)

1) Alice and Bob set up a shared vector $\vec{B}_c$.

- Alice chooses a random Boolean matrix.

$$B_{n\times N} = \begin{pmatrix} \vec{B}_1 \\ \vdots \\ \vec{B}_n \end{pmatrix} = \begin{pmatrix} b_1^{(1)} & \cdots & b_N^{(1)} \\ \vdots & \ddots & \vdots \\ b_1^{(n)} & \cdots & b_N^{(n)} \end{pmatrix}_{n\times N}$$

- Using QOT^N_{nN}, Bob obtains N bits out of $B_{n\times N}$, i.e.

$$\vec{B}_c = \{b_1^{(c)}, b_2^{(c)},\ldots, b_N^{(c)}\} .$$

2) Bob gets $\vec{S}_c = \{s_1^{(c)}, s_2^{(c)},\ldots, s_t^{(c)}\}$ from $\vec{B}_c$ with g_i ($1 \leqslant i \leqslant t$) provided by Alice.

$DO^t_{i=1}$

- Alice randomly picks $g_i \in \boldsymbol{G}$, and reveals it to Bob.
- Alice obtains n m-bit vectors, by computing:

$$g_i(\vec{B}_j) = {}^i\vec{C}_j = \{{}^i c_1^{(j)}, {}^i c_2^{(j)},\ldots, {}^i c_m^{(j)}\} \quad (1 \leqslant j \leqslant n)$$

Where m is a safety parameter.

- Alice announces to Bob.

$$\hat{s}_i^{(j)} \leftarrow s_i^{(j)} \oplus (\overset{m}{\underset{l=1}{\oplus}} {}^i c_l^{(j)}) \quad (1 \leqslant j \leqslant n)$$

- Bob obtains $s_i^{(c)}$ by computing

$$g_i(\vec{B}_c) = {}^i\vec{C}_c = \{{}^i c_1^{(c)}, {}^i c_2^{(c)},\ldots, {}^i c_m^{(c)}\}$$

$$s_i^{(c)} \leftarrow \hat{s}_i^{(c)} \oplus \left(\overset{m}{\underset{l=1}{\oplus}} {}^i c_l^{(c)} \right)$$

3.2 QOT^N_{nN} Scheme

Now we present how to construct QOT^N_{nN}. Firstly, we stress that, one strategy cheating Bob can take in quantum oblivious transfer is called photon-storing attack. That is, he does not measure the photons and stores them until Alice has announced all her bases, then he can obtain all Alice's bits by measuring stored photons with these bases. In our *QANDOS* protocol, we use H. P. Yuen's *QBC* protocol to resist photon-storing attack. Suppose that Alice has $n \times N$ bits $\{b_1, b_2,\ldots, b_{nN}\}$ (i.e. $B_{n\times N}$) and is willing to disclose exact N bits of them to Bob at his choice $\{b_{\varsigma_1}, b_{\varsigma_2},\ldots, b_{\varsigma_N}\}$ (i.e. $\vec{B}_c$); on the other hand, Bob doesn't want Alice to know his selections $\{\varsigma_1, \varsigma_2,\ldots, \varsigma_N\}$. Thus, Bob needs to set up $n \times N$ subsets: $I_1, I_2,\ldots, I_{nN} \subset \{1, 2,\ldots, M\}$, where M is a safety parameter and $|I_1| = |I_2| = \ldots = |I_{nN}|$. Thereinto, $\{I_{\varsigma_1}, I_{\varsigma_2},\ldots, I_{\varsigma_N}\}$ can enable Bob to obtain $\{b_{\varsigma_1}, b_{\varsigma_2},\ldots, b_{\varsigma_N}\}$, while the others will spoil the remnant b's.

Before we go into the detailed introduction of QOT^N_{nN}, another point necessary to be explained is that, to ensure that Bob can obtain exact N bits, the ratio of shared bases (i.e. $\beta_i = \beta_i'$) in $I_1 \cup \ldots \cup I_{nN}$ should satisfy (after step IV)

$$\frac{N}{nN} \leq \frac{|\{i \mid \beta_i = \beta_i', i \in I_1 \cup \ldots \cup I_{nN}\}|}{|I_1 \cup \ldots \cup I_{nN}|} < \frac{N+1}{nN} \qquad (1)$$

In this paper, we let the ratio be

$$\frac{1}{2}\left(\frac{N}{nN} + \frac{N+1}{nN}\right) = \frac{2N+1}{2nN} . \qquad (2)$$

Obviously, $\lim_{M\to\infty} \frac{|\{\beta_i = \beta_i'\}|}{M} = \frac{1}{2}$. Thus, there exist redundant shared bases that should be removed (at step IV) in order to meet Eq.(1). Let x be the number of bases to be removed. Then, x can be calculated as follows:

$$\frac{M/2 - x}{M - x} = \frac{2N+1}{2nN}$$

$$x = \frac{nN - (2N+1)}{2nN - (2N+1)} M \qquad (3)$$

Where M should satisfy $2nN-(2N+1) \mid M$.

Let $\beta_{\sigma_1}, \beta_{\sigma_2},\ldots, \beta_{\sigma_x}$, $\sigma_1, \sigma_2,\ldots, \sigma_x \in \{1,2,\ldots, M\}$, denote these x bases.

QOT^N_{nN} ($b_1, b_2,\ldots, b_{nN}$) ($\varsigma_1, \varsigma_2,\ldots, \varsigma_N$)

I. $DO^{2M}_{i=1}$

- Alice chooses a random bit β_i **,** and defines her emission Basis: $\{|\theta_i\rangle, |\theta_i^{\perp}\rangle\} \leftarrow [+,\times]_{\beta_i}$

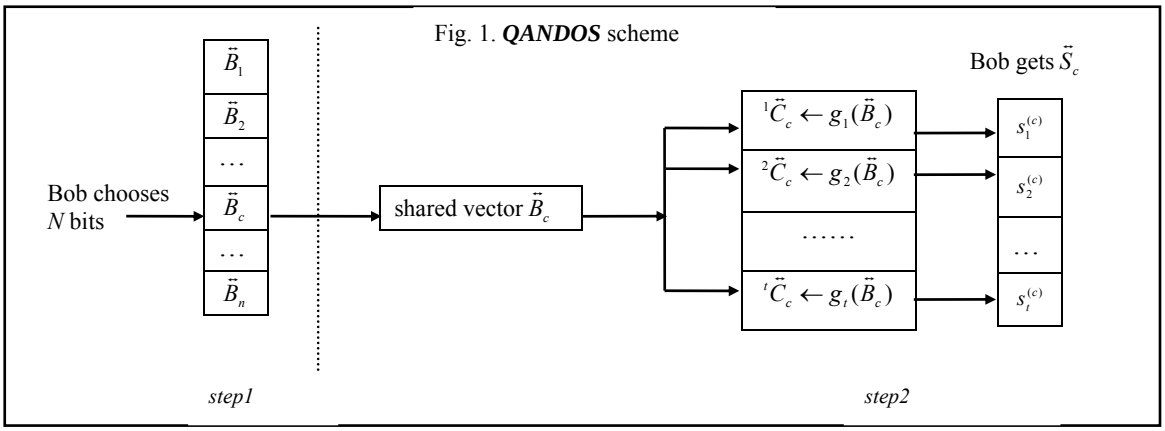


Fig. 1. ***QANDOS*** scheme

- Alice chooses a random bit r_i, and sends to Bob a photon: $|\varphi_i\rangle \leftarrow [|\theta_i\rangle, |\theta_i^{\perp}\rangle]_{r_i}$
- Bob chooses a random bit β_i', and measures $|\varphi_i\rangle$ with the basis: $\{|\theta_i'\rangle, |\theta_i'^{\perp}\rangle\} \leftarrow [+,\times]_{\beta_i'}$
- Bob sets
  - $r_i' \leftarrow 0$, if $|\varphi_i\rangle$ was seen as $|\theta_i'\rangle$.
  - $r_i' \leftarrow 1$, if $|\varphi_i\rangle$ was seen as $|\theta_i'^{\perp}\rangle$.

3) $DO_{i=1}^{M}$

- Bob runs commit (β_i'), commit (r_i'), commit (β_{M+i}') and commit (r_{M+i}') with Alice.
- Alice chooses a random bit d_i, and reveals it to Bob.
- Bob runs unveil (β_{Md_i+i}'), unveil (r_{Md_i+i}').
- Alice checks Bob's commitments.
  - If $\beta_{Md_i+i}' = \beta_{Md_i+i}$, and $r_{Md_i+i}' = r_{Md_i+i}$, protocol proceeds.
  - If $\beta_{Md_i+i}' = \beta_{Md_i+i}$, and $r_{Md_i+i}' \neq r_{Md_i+i}$, protocol returns to the beginning.
- If d_i =0, then
  - Alice sets $r_i' \leftarrow r_{M+i}$, $\beta_i \leftarrow \beta_{M+i}$.
  - Bob sets $r_i' \leftarrow r_{M+i}'$, $\beta_i' \leftarrow \beta_{M+i}'$.

4) Alice announces to Bob $\beta_1, \beta_2, ..., \beta_M$.

5) $DO_{j=1}^{x}$

- Bob reveals to Alice $\beta_{\sigma_j}', r_{\sigma_j}'$.
- Alice checks Bob's commitments.
  - If $\beta_{\sigma_j}' = \beta_{\sigma_j}$, and $r_{\sigma_j}' = r_{\sigma_j}$, protocol proceeds.
  - If $\beta_{\sigma_j}' \neq \beta_{\sigma_j}$, or $r_{\sigma_j}' \neq r_{\sigma_j}$, protocol returns to the beginning.

6) Bob partitions $\{1,2,...,M\}/\{\sigma_1, \sigma_2, ..., \sigma_x\}$ into N disjoint subsets $I_1, I_2, ..., I_{nN}$, satisfying:

- $|I_1| = |I_2| = ... = |I_{nN}| = \frac{M-x}{nN} = \frac{M}{2nN-(2N+1)}$
- $\forall j \neq k, I_j \cap I_k = \phi$
- $\forall j \in I_{\varsigma_1} \cup ... \cup I_{\varsigma_N}, \beta_j' = \beta_j$

and announces to Alice $< I_1, I_2, ..., I_{nN} >$.

7) After receiving $< J_1, J_2, ..., J_{nN} > = < I_1, I_2, ..., I_{nN} >$, Alice reveals to Bob:

$$\hat{b}_i \leftarrow b_i \oplus \left(\oplus_{j \in J_i} r_j\right) \quad (1 \leqslant i \leqslant nN)$$

8) Bob obtains $\{b_{\varsigma_1}, b_{\varsigma_2}, ..., b_{\varsigma_N}\}$ by computing:

$$b_{\varsigma_i} \leftarrow \hat{b}_{\varsigma_i} \oplus \left(\oplus_{j \in I_{\varsigma_i}} r_j'\right) \quad (1 \leqslant i \leqslant N)$$

4. ANALYSIS

Firstly, we show the correctness of our *QANDOS* protocol, and then demonstrate that it can resist any feasible attacks from the aspects of two participants' cheating strategies respectively. To be simple, we suppose that the channel used for quantum transmission is noise-free, so that $r_i = r_i'$ whenever $\beta_i = \beta_i'$. Note that, a more practical scheme can also be proved secure, with some modifications similar to those used in [12]. We omit this proof for brevity.

4.1 Correctness

A correct *QANDOS* should be that if both Alice and Bob are honest, Bob will end up with only one secret, ignorant of the others except with arbitrary small probability. We first consider the case where Bob and Alice have set up a shared vector $\vec{B}_c = \{b_1^{(c)}, b_2^{(c)}, ..., b_N^{(c)}\}$ (after step I in *QANDOS*). Obviously, Bob can get $\vec{S}_c = \{s_1^{(c)}, s_2^{(c)}, ..., s_t^{(c)}\}$, with g_i and $\hat{s}_i^{(j)} \leftarrow s_i^{(j)} \oplus \left(\bigoplus_{l=1}^{m} {}^{i}c_l^{(j)}\right)$ $(1 \leqslant i \leqslant t,\ 1 \leqslant j \leqslant n)$ offered by Alice (at step II), by computing:

$$g_i(\vec{B}_c) = {}^{i}\vec{C}_c = \{{}^{i}c_1^{(c)}, {}^{i}c_2^{(c)}, ..., {}^{i}c_m^{(c)}\}$$

$$s_i^{(c)} \leftarrow \hat{s}_i^{(c)} \oplus \left(\bigoplus_{l=1}^{m} {}^{i}c_l^{(c)}\right)$$

Where g_i, $\hat{s}_i^{(j)}$, $\vec{B}_c$ are shared by both parties, and

$$\left[s_i^{(c)} \oplus \left(\bigoplus_{l=1}^{m} {}^{i}c_l^{(c)}\right)\right] \oplus \left(\bigoplus_{l=1}^{m} {}^{i}c_l^{(c)}\right) = s_i^{(c)}$$

On the other hand, the correctness of the QOT^{N}_{nN} protocol has been proved in [6], which ensures Bob to exactly get $\{b_{\varsigma_1}, b_{\varsigma_2}, ..., b_{\varsigma_N}\}$ (i.e. $\vec{B}_c$), except with negligible probability. Thus, our scheme is of correctness.

4.2 Alice's Strategy

Note that there is very little that Alice can do to cheat Bob, unless she offers Bob junk instead of valuable secrets.

However, this does not count as genuine cheating, and nothing can prevent this type of cheating unless Alice has to commit to her secrets before the start of the protocol. What would count as a genuine success cheating for Alice would be if she would determine (or at least get an indication about) which of her secrets was of interest to Bob. But Bob does not offer any useful information about his choice c except $< J_1, J_2, ..., J_{nN} >$. Moreover, c is purely random and information-theoretically hidden from Alice, because she cannot tell which of Bob's sets is entirely made up of shared bases. Thus, for Alice, the list of $< J_1, J_2, ..., J_{nN} >$ is chosen randomly, helpless to infer Bob's choice. Therefore, Alice is impossible to confirm c, regardless of her computing power and available technology, provided that Bob faithfully follows the protocol.

4.3 Bob's Strategy

Compared with Alice, Bob has more chance to cheat, and his strategies are more complex as well. Differently, Bob's aim of cheating is to obtain more than one secret from Alice. To realize that, Bob must gain information on two shared vectors at least. In this section, we prove our *QANDOS* robust against Bob's any technologically feasible attacks.

1) Cheating by Luck

As analyzed before, only the set I_j entirely made up of shared bases enables Bob to get b_j reliably. As a result, if Bob gets over N bits in QOT^N_{nN}, he is likely to finish *QANDOS* with more than one secret. To be technically exact, we first consider the case that Bob is luckier than average, getting more shared bases (at step I in QOT^N_{nN}).

Lemma 1. (Hoelfding inequality [13]) Let $X_1, X_2, ..., X_n$ be n independent random variables with identical probability distribution, each ranging over the (real) interval $[a, b]$, and let μ denote the expected value of each of these variables. Let $Y = (X_1 + ... + X_n)/n$ and $\eta \geq 0$, then

$$\Pr(|Y - \mu| \geq \eta) \leq 2e^{\frac{-2n\eta^2}{b-a}} \tag{4}$$

Theorem 2. In QOT^N_{nN}, if Alice follows the protocol, there exists a positive constant $\varepsilon<1$, such that Bob ends up with more than N bits with the probability at most ε^M.

***Proof*:** The probability that Bob gets over N bits in QOT^N_{nN} is given by:

$$\Pr\left[\frac{|\{\beta_i = \beta_i'\} - x|}{M - x} \geq \frac{N+1}{nN}\right]$$

$$= \Pr\left[\sum_{i=1}^{M} \beta_i \oplus \beta_i' \leq M - \left(M - \frac{nN-(2N+1)}{2nN-(2N+1)}M\right)\frac{N+1}{nN} - \frac{nN-(2N+1)}{2nN-(2N+1)}M\right]$$

$$\leq \Pr\left[\left|\frac{1}{M}\sum_{i=1}^{M}\beta_i \oplus \beta_i' - \frac{1}{2}\right| \geq \frac{1}{2} - \frac{nN-(N+1)}{2nN-(2N+1)}\right]$$

$$< 2e^{-2M\left(\frac{1}{2} - \frac{nN-(N+1)}{2nN-(2N+1)}\right)^2}$$

$$< e^{-M\left(\frac{1}{2} - \frac{nN-(N+1)}{2nN-(2N+1)}\right)^2}$$

$$= \varepsilon^M$$

Where we use lemma 1 in the second inequality, and $\varepsilon = e^{-\left(\frac{1}{2} - \frac{nN-(N+1)}{2nN-(2N+1)}\right)^2} < 1$. Thus, the probability that Bob gets more than N bits in QOT^N_{nN} is negligible, provided that M is sufficiently large. □

2) The Breidbart Attack: Bob can adopt another strategy to obtain over N bits in QOT^N_{nN}, by performing illegal measurements on Alice's photons with *Breidbart basis* [14]. C. H. Bennett, etc. have proved QOT^1_2 robust against the Breidbart attack in [12], for Bob would have an exponential number of equally likely candidates for Alice's original data except with exponentially small probability. Similarly, we can also demonstrate our *QANDOS* scheme immune to the Breidbart attack. However, due to space limitation we don't give the proof in this paper.

Moreover, in our *QANDOS* protocol, the situation is even worse for Bob if he attempts to cheat in this way. Obviously, no measure can do better than the legitimate measurements with shared bases, which can ensure Bob to safely obtain Alice's bits. Our protocol allows Bob to use the shared bases to perform measurements and get exact N bits eventually. However, before Alice reveals her emission bases (at step III in QOT^N_{nN}), Bob cannot distinguish the shared bases from false bases, and thus he is unable decide which photons should be measured with *Breidbart basis* and which should not. As a result, if Bob measures all photons with *Breidbart basis*, he not only fails to obtain one more bit, but also loses his chance to get his legitimate N bits at the same time. Therefore, the Breidbart attack is trivial for our scheme.

3) Split Attack

We have proved that, the probability that Bob can get over N bits in QOT^N_{nN} is arbitrary small. However, if cheating Bob attempts to gain information on other vectors besides the legitimate one $\vec{B}_c$, he can take another approach by splitting his N legitimate bits into several groups. For example, in order to get both $\vec{S}_c$ and $\vec{S}_j$ ($j \neq c$), he distributes $N/2$ bits for $\vec{B}_c$, and $N/2$ bits for $\vec{B}_j$. Now, we show that this strategy doesn't work in our protocol.

Lemma 3 [15]. Let W be a random n-bit string with uniform distribution over $\{0,1\}^n$, let $V=e(W)$ for an arbitrary eavesdropping function e: $\{0,1\}^n \to \{0,1\}^t$, for some $t<n$, let $s<n-t$ be a positive safety parameter, and let $r=n-t-s$. If Alice and Bob choose $K=g(W)$ as their secret key, where g is chosen at random from a class of universal hashing from $\{0,1\}^n$ to $\{0,1\}^r$, then Eve's expected information about the secret key K, given G and V, satisfies

$$I(K; g, V) \leq \frac{2^{-s}}{\ln 2} \tag{5}$$

Theorem 4. If Bob attempts to cheat Alice by splitting his N legitimate bits into several groups, there exists a positive constant $\varepsilon<1$, such that Bob finishes the *QANDOS* protocol with more than one secret, with the probability at most ε^N.

Proof: Suppose that, Bob finishes QOT^N_{nN} with ηN bits of $\vec{B}_c$, and $(1-\eta)N$ bits of $\vec{B}_j$ (let $\vec{B}_j' = \{b_1^{(j)}, b_2^{(j)}, ..., b_{(1-\eta)N}^{(j)}\} \subset \vec{B}_j$ denote these $(1-\eta)N$ bits), where $0.5 \leq \eta < 1$[1]. Then, according to lemma3, we have

[1]This case $0 < \eta < 0.5$ is equivalent to that case $0.5 \leq \eta < 1$, if we exchange $\vec{B}_c$ with $\vec{B}_j$.

$$I\left({}^{i}\vec{C}_j; g_i, \vec{B}'_j\right) \le \frac{2^{N+(1-\eta)N+m}}{\ln 2} = \lambda\varepsilon^N \quad (6)$$

Where $\lambda = \frac{2^m}{\ln 2}$, $\varepsilon = 2^{-\eta} < 1$. We let $\varepsilon_1 = 0.9\varepsilon < 1$, then

$$I\left({}^{i}\vec{C}_j; g_i, \vec{B}'_j\right) = \varepsilon_1^{\,N} \quad (7)$$

When *N* is sufficiently large. Thus, the probability that Bob obtains over one secret is negligible.

In addition, if Bob does so, the chance of getting his legitimate secret is influenced at the same time. In particular, the information that Bob gains about ${}^{i}\vec{C}_c$ is $I\left({}^{i}\vec{C}_c\right) = \varepsilon_2^{\,N}$, where $\varepsilon_2 = 0.9 \times 2^{-(1-\eta)} < 1$. Clearly, it is difficult for him to get $\vec{S}_c$ when *N* is large.

5. CONCLUSION AND FURTHER WORK

Our *QANDOS* scheme is superior to previous ones in two aspects. Firstly, it is unconditionally secure since both Alice and Bob have no limitation on their computing power, and it can resist any technologically feasible attacks and ensure two parties' privacies effectively. On the other hand, our protocol is of high efficiency, due to its relatively low communication complexity.

A secure $QOT^l{}_2$ protocol always needs to include *QBC* to resist photon-storing attack. Unfortunately, secure *QBC* is always implemented complicatedly, which influences the efficiency of $QOT^l{}_2$. Thus, how to construct a $QOT^l{}_2$ scheme immune to photon-storing attack without using *QBC* is interesting for the further work.

REFERENCES

[1] G. Brassard, C. Crépeau, et al, "Information theoreticreduction among disclosure problems", Proceedings of 27th Annual IEEE Symposium on Foundations of Computer Science, 1986, 168-173.

[2] G. Brassard, C. Crépeau, "All or nothing disclosure of secrets", in Advances in cryptology - CRYPTO'86, Lecture Notes in Computer Science 1403, Springer-Verlag, Berlin, 1987.

[3] A. Salomaa, L. Santean, "Secret selling of secrets with several buyers", In the 42nd EATCS Bulletin, 1990, 178-186.

[4] V. Niemi, A. Renvall, "Cryptographic protocols and voting", In Result and Trends in Theoretical Computer Science, Lecture Notes in Computer Science 812, 1994, pp. 307-316.

[5] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer [J]", SIAM Journal on Computing, 1997, 26 (5):1484-1509.

[6] A. Kent, "Unconditionally Secure All-or-Nothing Disclosure of Secrets", Arxiv preprint quant-ph /9711025, 1997.

[7] D. Mayers, "Unconditionally Secure Quantum Bit Commitment is Impossible", Physical Review Letters, 1997, 78(17):3414 –3417.

[8] Z. D. Chen, H. Zhu, "Quantum m-out-of-n Oblivious Transfer", Arxiv preprint cs.CR/ 0311039, 2003.

[9] Z. D. Chen, T. M. Bu, et al, "Unconditionally Secure Quantum m-out-of-n Oblivious Transfer", arXiv: quant-ph /0401057, 2004, 3(16).

[10] H. P. Yuen, "Unconditionally Secure Quantum Bit Commitment", ArXiv:quant-ph/0505132, 2005, 1(18).

[11] J. L. Carter, M. N. Wegman, "Universal classes of hash functions", Journal of Computer and System Sciences, 1979, 18:143-154.

[12] C. H. Bennett, G. Brassard, et al, "Practical Quantum Oblivious Transfer", In Advances in Cryptology –CRYPTO'91, Proceedings, LNCS, Springer-Verlag, 1992, 576: 362-371.

[13] W. Hoefding, "Probability Inequalities for Sums of Bounded Random Variables", Journal of the American Statistical Association, 1963, 58: 13-30.

[14] C. H. Bennett, G. Brassard, et al, "Quantum cryptography, or unforgeable subway tokens", Advances in Cryptology: Proceedings of Crypto'82.

[15] C. H. Bennett, G. Brassard, et al, "Generalized Privacy Amplification", IEEE Translation on Information Theory, 1995, 41(6).

An Efficient Fast Block Type Selection for Intra prediction in Inter frame of H.264/AVC FRExt *

Qiong Liu, Ruimin Hu, Li Zhu, Zhen Han
Computer School, Wuhan University
Wuhan, Hubei 430079, China
Email: liuq_whu@163.com

ABSTRACT

With advanced intra prediction modes, intra coding of H.264/AVC offers significant coding gains compared with previously video coding standards, but the computational burden is extremely high. Recently, H.264/AVC adopts FRExt profile to improve performance in higher-fidelity video, which added intra8×8 predictions. We proposed an efficient fast block type selection algorithm for intra-prediction of H.264/AVC FRExt. The algorithm utilized smoothness feature of blocks to predict the optimal intra block types. By avoiding unlikely block types, the computation can be much reduced. Experimental result shows the proposed algorithm can reduce the computation complexity up to 46.28%, while little PSNR degradation or bit-rate increase.

Keywords: H.264/AVC, Intra prediction, Block type selection, Inter frame, Fast algorithm.

1. INTRODUCTION

H.264/AVC [1] adopted the Fidelity Range Extensions (FRExt) including a set of new coding tools, to meet the growing demand for coding of higher-fidelity video material.

H.264/AVC has improved the coding gain over a wide range of bit rates significantly by allowing a rich set of coding modes. In inter frames, the full search algorithm will search both inter and intra modes. Respectively, there are seven block sizes of inter prediction, while two or three block types of intra prediction. The main difference between intra prediction of FRExt and non-FRExt H.264/AVC coding is the use of intra8×8 in addition to intra4×4 and intra16×16 of luma samples. Intra8x8 prediction also uses rate distortion optimization (RDO) computation to decide the best prediction mode. And the full decision algorithm of H.264/AVC will search all the possible block types and modes. Therefore, the complexity and computation load of intra-prediction, which is already extremely large, becomes much larger. On the other hand, the statistics of the number of intra mode in inter frame is 3% on average [2]. So the most computation of intra prediction is wasted.

There are three kinds of solution theoretically. The first one is pre-decision of using inter or intra prediction to current marcoblcok (16×16 samples). Research of [3] proposed a fast intra/inter mode selection scheme using the variance and the sum of absolute differences (SAD). To facilitate video transmission over networks, it is reasonable to use the maximum likelihood (ML) criterion[4]. However, these schemes are not suitable for H.264/AVC since the selected features are too simple to provide accurate mode prediction. In [5], it decided the mode using the expected risk of choosing the wrong mode in a multidimensional feature space. It gets little R-D loss, but the computation reduction is limited. Another kind of method is early termination. Perform the inter prediction firstly. If the RDO cost of the best inter mode is smaller than a threshold, there is no need to perform any intra prediction. The accuracy deeply depends on the threshold. And for getting more time saving, it should improved by the fast intra mode selection, which is the third kind of solution.

Some approaches have been proposed about fast intra prediction algorithm. In [6], it proposed a threshold to early terminate the computation of the most probable mode. The proposed algorithm of JVT-G013 [7] is based on the local edge information, and thus adopts the edge direction to predict the possible mode. It is a quite efficient algorithm and is recommended by JVT. However, the most of exist algorithms are only implied in fast mode decision, not in fast block type selection. It is desirable to make a coarse level mode decision about proper block types to try at the first stage.

In this paper, we proposed a fast block type decision algorithm. Prior to select the best mode of each block type, we choose the proper intra coding block type to reduce the computational load. Then, we choose the specific mode within each class at the second stage for a fine-level mode decision. Experimental results indicate that the proposed algorithm, compared by full-search algorithm, can achieve up to 46.28% time saving while maintaining the same rate distortion performance.

The paper is organized as follow: In section 2, we proposed our fast algorithm. In section 3, it showed the experiment result. And a conclusion was made in section 4.

2. PROPOSED ALGORITHM

To get high efficiency of intra coding, H.264/AVC employs spatial intra prediction. A prediction block is formed based on previously coded and reconstructed blocks. As illustrated in Fig. 1, the intra8×8 prediction uses neighboring 25 pixels to predict the 8×8 block and has the same 9 directions as intra4×4 prediction.

* This work was supported by National Natural Science Foundation of China (NO. 60472040).

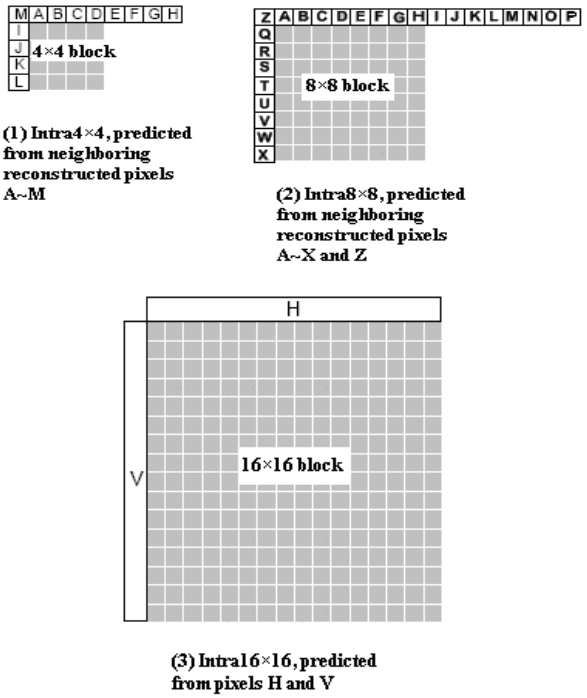


Fig. 1. Three intra block types of H.264/AVC: intra4×4, intra8×8 and intra16×16

Intra coding block type is highly depended on the smoothness of blocks. A small intra coding block size is well suit for a marcroblock with detailed information, while larger size is well suit for smooth one. In many higher-fidelity applications, the resolution of pictures is high. Its spatial correlation is higher than lower resolution pictures. For some marcroblocks, intra16×16 is not suitable to present the detailed information and intra4×4 cost much, so intra8×8 will be chosen.

Our proposed algorithm is based on the high correlation between the smoothness and the coded block type. In this paper, we use edge information to present the smoothness of blocks. In [7], it is also use edge information. So it is easy to add our proposed algorithm to the fast mode decision algorithm in [7].

At first, apply the Sobel operator to each pixel of current image. It is calculated by the Eq.(1), and then the amplitude is derived by Eq.(2).

$$\begin{aligned} dx_{i,j} &= P_{i-1,j+1} + 2\times P_{i,j+1} + P_{i+1,j+1} \\ &- P_{i-1,j-1} - 2\times P_{i,j-1} - P_{i+1,j-1} \\ dy_{i,j} &= P_{i+1,j-1} + 2\times P_{i+1,j} + P_{i+1,j+1} \\ &- P_{i-1,j-1} - 2\times P_{i-1,j} - P_{i-1,j+1} \end{aligned} \quad (1)$$

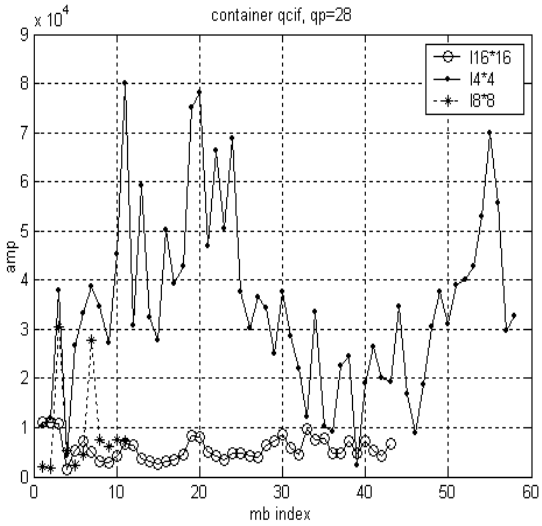
$$Amp(D_{i,j}) = |dx_{i,j}| + |dy_{i,j}| \quad (2)$$

We select the 16×16 block as the basic unit to create edge histogram. Then, we calculate $Amp_{16\times16}$ of 16×16 block using Eq.(3), and use $Amp_{16\times16}$ to present the spatial characteristics in the following algorithm.

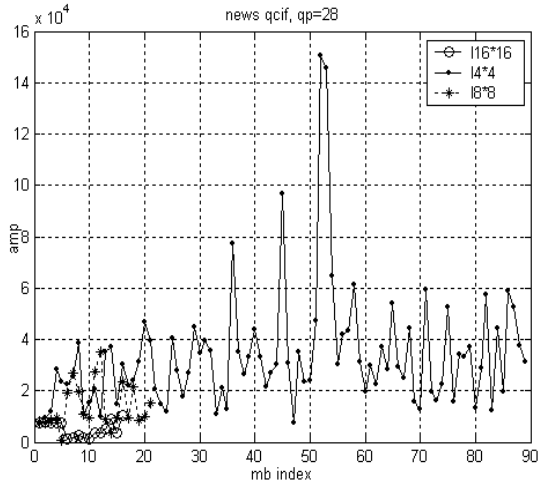
$$Amp_{16\times16} = \sum_{j=0}^{15}\sum_{i=0}^{15} Amp(D_{i,j}) \quad (3)$$

After the smoothness detection, if we can predict the proper block type by $Amp_{16\times16}$ without other computation, the complexity of intra prediction should be greatly reduced. Therefore, how to predict the proper block type according to $Amp_{16\times16}$ is the key of the whole algorithm. It makes a decisive influence for the accuracy of prediction.

To find the correlation between $Amp_{16\times16}$ and the coded block type, we experimented by recoding the $Amp_{16\times16}$ of each block type and comparing them. To find the general situation, we performed the experiments on different sequences. Fig. 2 shows the statistic results of sequence *container* and *news*.



(a) *container*, *QP* =28



(b) *news*, *QP*=28

Fig. 2. The statistic of $Amp_{16\times16}$ of blocks for different intra prediction type

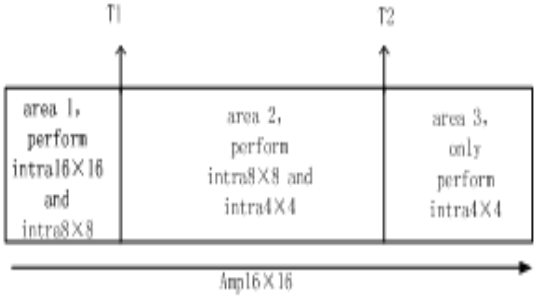


Fig. 3. Classification of the three block types

From the results, we analyzed as below:

(1)The curve of intra4×4 and intra16×16 is obviously separated. On average, the $Amp_{16\times16}$ of intra 4×4 blocks is much lager than $Amp_{16\times16}$ of intra16×16 blocks.

(2)The curve of intra8×8 is mixed with the other two curves. Its higher part is mixed with the lower part of intra4×4, and its lower part is mixed with intra16×16.

(3)Although some intra8×8 samples are mixed with intra4×4, they are only limited in a small zone. The largest value of $Amp_{16\times16}$ of intra 8×8 is much less than intra 4×4's.

(4)In general, the amount of the three type blocks is quit difference. If *QP* is not very large, the majority is intra4×4.

Based on these facts, we developed our fast block type algorithm. According to the value of $Amp_{16\times16}$, we use two thresholds to divide the whole feature space into three areas, which is denoted in Fig. 3. If $Amp_{16\times16}$ of current block is higher lower a threshold T_1, the block intends to be coded as intra16×16 or intra8×8. Otherwise, if $Amp_{16\times16}$ is lower than a threshold T_2, the current block intends to be coded as intra8×8 or intra4×4. Otherwise, only choose the intra4×4.

For each macroblock of frames, the flowchart of proposed algorithm is showed in Fig. 4. However, it should be noticed that Fig.4 only presented the intra predictions. In inter frame, we should perform inter prediction and then perform this fast algorithm. The best inter mode and best intra mode should be compared to decide the optimal mode.

3. EXPERIMENT RESULTS

The proposed algorithm is compared with full-search algorithm of JM9.5 [8] in timesaving, bit-rate rising and PSNR.

In the experiment, we test 4 classical sequences, which are *news*, *container*, *silent* and *paris*, and each sequence has 100 frames. The simulation condition is in Table 1. In this experiment, we set the two thresholds as: T_1=10000 and T_2=25000. Calculation of average PSNR differences and average bit rate differences follows the specification as below:

$$\Delta Time = \frac{Time_{proposed} - Time_{JM}}{Time_{JM}} \quad (4)$$

$$\Delta PSNR = PSNR(y)_{proposed} - PSNR(y)_{JM} \quad (5)$$

$$\Delta bits = \frac{bits_{proposed} - bits_{JM}}{bits_{JM}} \quad (6)$$

Table 1. Test Conditions

Profile	Main
Codec	JM9.5
MV Search Range	-16~+16
Number of Reference frame	1
RDO	1
Entropy coding method	CABAC
Hadamard transform	1
Sequence Structure	IPPP or IIII
Intra period	1 for IIII, and 10 for IPPP
PC	Pentium IV 3.0G, 1G RAM

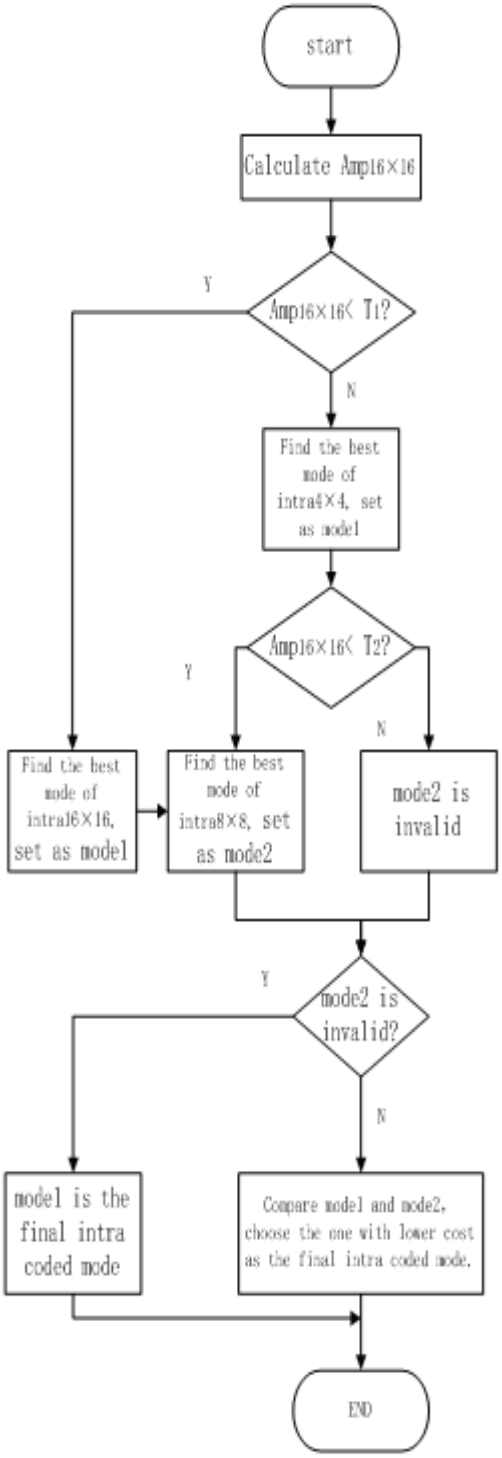


Fig. 4. The complete flow of proposed fast block types selection algorithm

Table 2~3 show the comparison results. In these tables, positive number means increasing, and negative number means decreasing. H.264 FRExt is applied to higher-fidelity

video material, so the *QP* could not be very large. In this experiment, we set *QP*=20, 28.

To find the accuracy impact of proposed algorithm for the intra prediction, we test it on all I frames sequence. And the result is tabulated in Table 2 and Table 3. However, in the general applications, the all-I sequence structure is rare. So we also test it on IPPP sequence structure. And the result is tabulated in Table 4 and Table5.

Table 2~3 showed the results tested on all I frame sequences, the time saving can be up to 46.28%, while neglected RD loss. So this fast algorithm can reduce the intra encoding greatly.

In Table 4~5, we can see that the time saving is smaller than the result of Table 2 and Table 3. That is because of less computation proportion of intra prediction in the IPPP sequence structure. The motion estimation of inter prediction occupied more computation than intra prediction. Even though, the algorithm still can gain 12.2% time saving on average.

Table 2. The comparison results *QP*=20 all I frames

Sequence	*⊿Time (%)*	*⊿PSNR (db)*	*⊿bits (%)*
news	-38.01	-0.03	1.00
container	-46.46	0.03	0.85
silent	-30.41	0.02	1.02
paris	-43.61	0	0.07

Table 3. The comparison results *QP*=28 all I frames

Sequence	*⊿Time (%)*	*⊿PSNR (db)*	*⊿bits (%)*
news	-36.66	-0.04	0.88
container	-46.28	-0.02	1.7
silent	-30.08	-0.04	0.87
paris	-43.21	0	0.16

Table 4. The comparison results *QP*=20 IPPP frames

Sequence	*⊿Time (%)*	*⊿PSNR (db)*	*⊿bits (%)*
news	-12.97	-0.04	0.37
container	-14.67	-0.01	0.21
silent	-13.79	0.01	0.54
paris	-8.32	0	0.14

Table 5. The comparison results *QP*=28 IPPP frames

Sequence	*⊿Time (%)*	*⊿PSNR (db)*	*⊿bits (%)*
news	-12.88	-0.04	0.12
container	-14.13	-0.04	1.41
silent	-13.76	-0.05	0.41
paris	-8.13	0	-0.12

Compared the result of *QP*=20 and *QP*=28, it can be concluded that proposed algorithm's work performance are almost the same under the two different *QP*.

Compared the performances of four difference sequences, we can observe that the timesaving of sequence *paris* under IPPP structure is less than others. That is because motion estimation of *paris* is in larger proportion of the whole computation than other three sequences. So the timesaving of the whole is limited only by fast intra prediction.

Fig. 5 shows the RD curves of the IPPP frames of sequence *News*, by using full-search and our proposed algorithm.

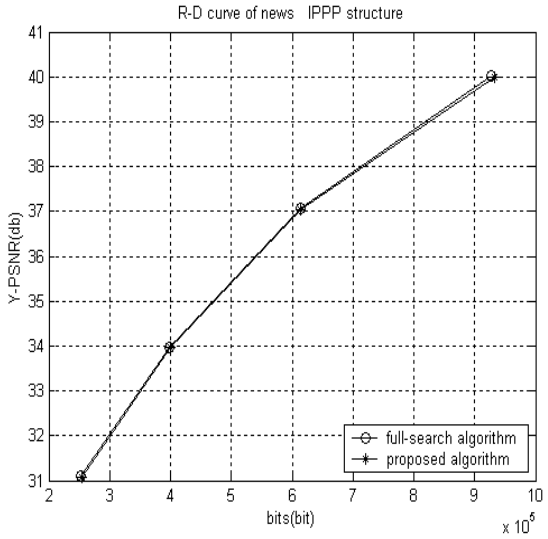


Fig. 5. RD curves of news

4. CONCLUSION

This paper proposed an efficient fast block type selection algorithm for H.264/AVC FREext. The motivation is to predict proper intra block types before conducting RDO computations by using correlation of edge information and block types. Be different from previous work, it distinguished not only intra4×4 and intra16×16, but also intra8×8. Experimental results indicate that the proposed algorithm achieves up to 46.28% time saving on average while maintaining the same rate distortion performance.

In future work, we believe that if we could find an efficient method to distinguish intra/inter mode, the time saving can be further improved by avoiding unnecessary inter prediction procedure.

5. REFERENCES

[1] ITU-T Recommendation H.264 & ISO/IEC 14496-10 (MPEG-4) AVC," Advanced Video Coding for Generic Audiovisual Services", (version 1:2003, versions 2: 2004) version 3: 2005

[2] J. Lee, B. Jeon, "Pruned Mode Decision based on Variable Block Sizes Motion Compensation for H.264," Lecture Notes in Computer Science (LNCS), vol.2899, Nov.2003, pp410-418.

[3] Y.K. Chen, A. Vetro, et al, "Optimizing intra/inter coding mode decisions," ISO/IEC JTC/SC29/WG11 (Coding of Moving Pictures and Associated Audio)

M2884, Oct. 1997.
[4] D.S. Turaga, T. Chen, "Classification based mode decisions for video over networks," IEEE Transection Multimedia, 3(1), 2001, pp41-52.
[5] C.S. Kim, C.C. Kuo, "A Feature-based Approach to Fast H.264 Intra/Inter Mode Decision", Conference Proceeding of Circuits and Systems, Vol. 1, May 2005, pp308-311
[6] B. Meng, O.C. Au, "Fast intra-prediction mode selection for 4x4 blocks in H.264" in Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal, April 2003, pp389-392.
[7] F. Pan, X. Lin, et al, "Fast Mode Decision Algorithm for Intra Prediction in JVT", JVT-G013, 7th JVT meeting, Pattaya, March 2003.
[8] JVT reference software version 9.5, http://bs.hhi.de/~suehring/tml/download/reference_software/

A New Kind of Hybrid Encryption System *

Youan Xiao, Zude Zhou, Layuan Li
School of Information Engineering, Wuhan University of Technology
Wuhan, Hubei 430070, P.R.China
Email: youan@21cn.com

ABSTRACT

In this paper, a new kind of hybrid encryption scheme named as XHES (Xiao's Hybrid Encryption Scheme) is proposed and discussed. This scheme combines the advantage of public-key scheme and symmetry encryption scheme. Unlike the other encryption scheme, XHES scheme has all the advantages of public key encryption scheme and symmetry encryption scheme: not only has quick encryption speed and highly security, but also has the benefit on key distribution and key management. No costly secrecy channel special for key distribution or no additional operations such as Identify, Provable and Authentication is need again. The scheme is simple and high-efficiency, and it can be applied in many domains, such as Distributed Network Computing, Electronic Commerce, etc.

Keywords: Hybrid Encryption System, Key Distribution, Symmetry Encryption Scheme, Xiao's Hybrid Encryption Scheme.

1. INTRODUCTION

Data encryption is one of the basic tasks in the cryptography scheme. Today, there are two types of data encryption schemes: Symmetry Encryption Scheme and Public-Key Encryption.

1) Symmetry Encryption Scheme:

Symmetry Encryption Scheme has a history of thousands of years, is also referred as conventional encryption. It is a form of cryptosystem in which encryption and decryption are performed using the same key, which is called as secret-key. A typical symmetry encryption scheme is as Fig. 1. And it has five ingredients:

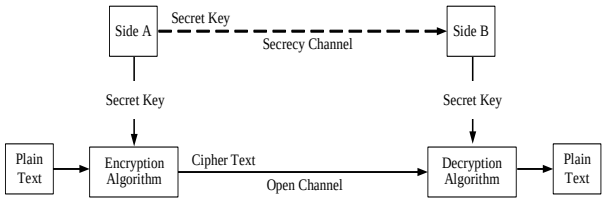


Fig. 1. Symmetry Encryption Scheme

(a) Plaintext: Readable message or data.
(b) Encryption Algorithm: Performs various substitutions and transformations on the plaintext.
(c) Secret Key: The secret key is also input to the algorithm.
(d) Cipher text: The scrambled message produced as output.
(e) Decryption Algorithm: This is essentially the encryption algorithm run in reverse.

There are many different symmetry encryption schemes, such as Data Encryption Standard (DES), which was standardized in 1977 as National Bureau of Standards; Advanced Encryption Standard (AES), a new standard to replace the DES; International Data Encryption Algorithm (IDEA), etc.

All these schemes try to convert a plaintext into something unreadable (ciphertext) using techniques called confusion and diffusion. So the symmetry scheme has many features such as quick encryption speed, high security category, etc. Because the process of encryption and the process of decryption are symmetrical, and the encrypter and decrypter use the same key to encrypt or decrypt a plaintext message, the distribution and management about the secret-key is very different because the secret-key can only be known by the legal two users. To solve these problems, the operations such as Identify, Provable and Authentication or the secrecy channel special for key distribution are needed. But either of them is complicated or costly. So the applicability range of the symmetry scheme is restricted by the problem in the Key Distribution.

2) Public-Key Encryption Scheme [1, 2]

The main problem with symmetric systems is the Key Distribution. In 1976, Whitfield Diffe and Martin Hellman firstly published a solution known as Public-Key Encryption Scheme which successes to solve this problem. The public-key scheme is based on the NP-complete one-way trap function. In this scheme, everyone has a pair of keys: one public key and one private key, which is different from the other one. So the public-key scheme is also called as asymmetric encryption scheme.

A typical public-key encryption scheme has also five key ingredients, which are described in figure 2.

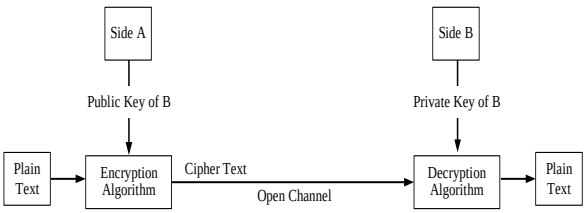


Fig. 2. Public-Key Encryption Scheme

(a) Plaintext: Readable message or data.
(b) Encryption Algorithm: Performs various transformations on the plaintext.
(c) Public and Private Keys Pair: This is a pair of keys that have been selected so that if one is used for encryption, the other is used for decryption.
(d) Ciphertext: The scrambled message produced as

* This work was supported by the Sunshine Young Project in Wuhan City of China under Grant No.20055003059-5, and it has been applied for the Chinese patent of invention (CN. 03128222.9).

output.

(e) Decryption Algorithm: It generates the ciphertext, the matching key and produces the original plaintext.

Since the discovery of public-key cryptography, there are been many realizations of public-key systems. Their security all depends on different computational problems which is NP-complete. The mathematics involved is, however, very complex and very slow. Commonly, the speed of a public-key system is only 1/100~1/1000 of the speed of a symmetry scheme which has the same level security. And it is the reason for its limited use so far.

2. THE XIAO'S HYBIRD ENCRYPTION SCHEME

In order to combine the advantage of public-key scheme and symmetry encryption scheme, a new kind of encryption scheme named as XHES (Xiao's Hybrid Encryption Scheme) is proposed. Unlike the other hybrid encryption scheme such as the PGP method, the XHES scheme unites the Public-Key Distribution Scheme and Symmetry Encryption Scheme. It has not only the advantage of symmetry scheme such as quick encryption speed and highly security, but also has the benefit on Key Distribution and Key Management from the public-key scheme.

The model of the Xiao's Hybrid Encryption Scheme is described as Fig. 3.

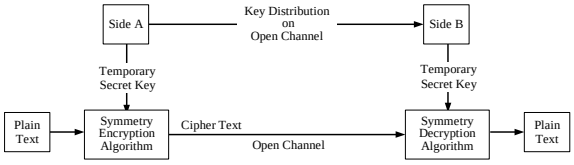


Fig. 3. Xiao's Hybrid Encryption Scheme

As described in Fig. 3. , there are two phases in the XHES Scheme: Key Distribution and Data Encryption.

1) Key Distribution:

This phase is based on the key distribution protocol in the public-key scheme. The sender chooses a temporary secrecy key randomly, and sends it to the legal receiver via the open channel with the key distribution protocol.

The key distribution technique in this part should have the ability of secrecy, identity authentication, integrality proving, etc. Also, it can be integrated with the symmetry encryption scheme seamlessly and safely.

Support the two sides are A and B, their private key are SK_A and SK_B, and the public key are PK_A and PK_B. The public-key of each side is stored in the trusty Center of Authentication (CA).

Firstly, side A chooses a temporary secrecy key "K" randomly. Then using the B's public-key PK_B and the key distribution protocol, side A distributes the temporary secrecy key "K" to B.

For receiver B, he uses his private key SK_B, executes the key distribution protocol, get the temporary secrecy key "K" which is distributed by A.

2) Data Encryption / Decryption:

This part is the conventional symmetry encryption scheme. It uses the temporary secrecy key "K" which is gotten in the last phase to finish the task of data encryption and data decryption.

In this part, on the base of the key distribution phase, A uses the chosen temporary secrecy key "K" as the secret-key of the symmetry encryption system, encrypt the plaintext message M, get the ciphertext C. Then send the ciphertext C to the receiver B via the open channel.

For the receiver B, he can decrypt the received ciphertext C via the symmetry encryption system with the distributed temporary secrecy key "K", and get the plaintext message M.

3. IMPLEMENTATION

In this section, we give a detailed implement example of the XHES scheme [3, 4].

As discussed in the above, the key distribution technique in the XHES scheme should have the ability of secrecy, identity authentication, integrality proving, etc. Also, it can be integrated with the symmetry encryption scheme seamlessly and safely.

After compare all the public-key schemes, we choose the elliptic curve public-key scheme as the public-key scheme in the implementation of XHES, and choose the XECDS scheme, which is introduced by the authors [3, 5] as the key distribution method. For the symmetry encryption scheme, we choose the Advanced Encryption Standard (AES) as the Data Encryption and Decryption Algorithm.

And the detailed implementation scheme can be described as below:

System Parameter:

(a) A large prime p which is chosen randomly.

(b) A safe elliptic curve E defined on the finite field F_p: $E(F_p): y^2 = x^3 + ax + b (\mathrm{mod}\ p)$, the base point of E is G.

(c) The order of E is n and r is a large prime gene of n.

(d) The private key of A and B noted as SK_A and SK_B is both random positive integers small than $r - 1$.

(e) The public key PK_A and PK_B can be gotten with the expressions below, and will be placed on the trusty Center of Authentication (CA).

$$\begin{cases} PK_A = SK_A^{-1} \times G \\ PK_B = SK_B^{-1} \times G \end{cases}$$

1) Data Encrypt Phase:

When sender A wants to encrypt the plaintext message M, and send to the receiver B safely, he follows the steps below:

① Sender A chooses a large integer $k \in [1, r-1]$ randomly, then the temporary secrecy key is $K = k \times G$. It will be used as the secret-key of the AES symmetry encryption system.

② Sender A uses the temporary secrecy key K to encrypt the plaintext message M via the AES symmetry encryption system, get the ciphertext C.

③ Sender A gets the public key PK_B from CA, and computes the $R = k \times PK_B$.

④ Sender A signs the message (R, C) with his private key SK_A, namely $S = \mathrm{Sig}_A(R, C)$ [6].

⑤ At last, A sends the message (R, C, S) to B.

2) Data Decrypt Phase:

When receiver B gets the message (R, C, S) sent by A, he can decrypt it by executes the following steps:

① B gets the public-key PK_A of A from CA, and checks the authenticity and integrality of the message received.

② B computes $K = SK_B \times R$ with his private key SK_B, and K is the same temporary secrecy key distributed by sender A.

③ with the temporary secrecy key K and the AES scheme, B decrypts the received ciphertext C, gets the plaintext message M encrypted by sender A.

4. ANALYSIS OF XHES SCHEME

There is an example of hybrid encryption scheme named as PGP. It uses the IDEA symmetry encryption system to encrypt the message data. And use the RSA encryption algorithm to encrypt the temporary secrecy key which is used by the IDEA system [7~9].

Compared with the PGP scheme, the proposed XHES scheme uses the XKDS Key Distribution Technique instead of the Public Data Encryption Technique because the XKDS scheme has higher efficiency, shorter key length, and lower resource requirement than RSA Data Encryption technique in the same security. On the other hand, due to the using of key distribution technique, the XHES scheme has the extra ability such as identity differentiate, authentication etc.

In the data encrypt phase, the AES symmetry encryption system is used instead of the IDEA system because the AES symmetry encryption is stronger than IDEA system, and it works better with the public-key system based on the elliptic curve discrete logarithm problem.

So the XHES scheme has is a simple and high-efficiency hybrid encrypt scheme.

5. CONCLUSIONS

On the Windows 2000 platform, we implement the XHES scheme with the Delphi development environment. With 1 MB plaintext, on an Intel Pentium III 700 MHz PC, the data encryption speed can reach 27.3MB/s, and the elliptic curve encryption system can only reach 0.7MB/s (EC-ElGamel Encryption Scheme).

The data show that the XHES scheme has all the advantages of public key encryption scheme and symmetry encryption scheme: it not only has quick encryption speed and highly security, but also has the benefit on key distribution and key management. And the additional operations such as Identify, Provable and Authentication are not need again. The costly secrecy channel special for key distribution isn't need too. The XHES scheme is, can be applied in many domain such as computer, distributed network computing, communication network, smart card, mobile phone, electronic commerce, etc.

6. REFERENCES

[1] D.V. Bailey, C. Paar, Efficient Arithmetic in Finite Fields Extensions with Application in Elliptic Curve Cryptography, Journal of Cryptology, 2002.6.

[2] IEEE 1363-2000, Standard Specifications for Public-Key Cryptography. 2000.12.
Available at http://www.ieee.org/grouper/1363.

[3] H.X. Xu, B. Li, "Security proof for distributed key distribution scheme". Journal of Software, 2005, 16(4): 570-576.

[4] Y.A. Xiao, "Researches on Elliptic Curve Public Key Cryptosystems in Network Information Security [PhD Thesis]". Wuhan University of Technology, 2003.

[5] Y.A. Xiao, L.Y. Li, "A New Quick Public Key Crypto System Based on the Difficulty of Factoring Very Large Numbers", 2001 International Symposium on Distributed Computing and Applications to Business, Engineering and Science — DCABES'2001, 230-234.

[6] Y.A. Xiao, L.Y. Li, "Research on the Xiao's key agreement scheme". Journal of Wuhan University of Technology (Transportation Science and Engineering), 2003, 27(6): 758-762.

[7] Y.A. Xiao, L.Y. Li, "Study of digital signature techniques". Journal of Wuhan University of Technology (Transportation Science and Engineering), 2002, 26(6): 737-740.

[8] H.F. Chau, "Unconditionally secure key distribution in higher dimensions by depolarization". IEEE Transactions on Information Theory, 2005, 51(4): 1451-1468.

[9] D.W Hong, J.S. Kang, An efficient key distribution scheme with self-healing property. IEEE Communications Letters, 2005, 9(8): 759-761.

Simulation research of base-band digital signal Super long distance transmission along twisted pair cable

Xinnan Fan, Xuewu Zhang, Hua Liu
College of Computer & Information Engineering, Hohai Univ.,
Changzhou ,213022, China
Email: fanxn@hhuc.edu.cn

Abstract

This paper introduces a super long distance transmission of base-band digital signal with the adoption of band-pass amplifier to the magnifying of signals, the voice transformer to enhance power and driven capacity, the low-pass filtering technology to filter corresponding components and the audio frequency transformer to isolate bus. Based on the data transmission circuit theory and the characteristic of practical using twisted pair cable, We not only build the Simulation model of transmission line, but hardware circuit is designed, and simulating the sending and receiving circuit that we designed by Matlab emphatically. The results of simulating show our research can improve transmission characteristic and distance of base-band digital signal. The research solve the practical problem of base band digital signal transmission, and applied to the signal command transmission of high-speed road emergency telephone system successfully.

Keywords: base-band digital signal, long distance transmission, simulating model, hardware circuit.

1 Introduction

Long distance transmission of base-band digital signal is a key issue of data communication. It is applied widely in many fields.(e.g.: high-speed road emergency telephone system, long distance water lever monitor system, industry control field).In such systems, signal command is transmitted by base-band in order to enhance resist disturbance and reliability. In the most common applications, RS485 bus communication standard is applied widely, but its transmission distance is limited within 1.2 kilometers. So technology of extend transmission distance is a key issue in the fields of data transmission of long distance (super tens of kilogram meters). First, we present construction of research platform. In section 2, we design hardware circuit for base-band digital signal long distance transmission. Simulating research to digital base band signals transmission according to the data transmission circuit theory and the characteristic of practical using twisted pair cable in section 3, Section 4 includes this paper.

2 Constructing system research platform

This system takes the theory and application of research of long distance transmission of base-band signal by bus construction, which follows the OSI/RAM by ISO. All the terminals shared a common physical channel (bus) and it transmits the signals among terminals by the means of host and slave terminals and poll modes by half duplex and base-band signals coupling the shared channels via audio frequency transformer. Link layer guarantee the transmission of datum transparently, and application layer user can ensure the transmission reliably by self-defining protocol [1,2].

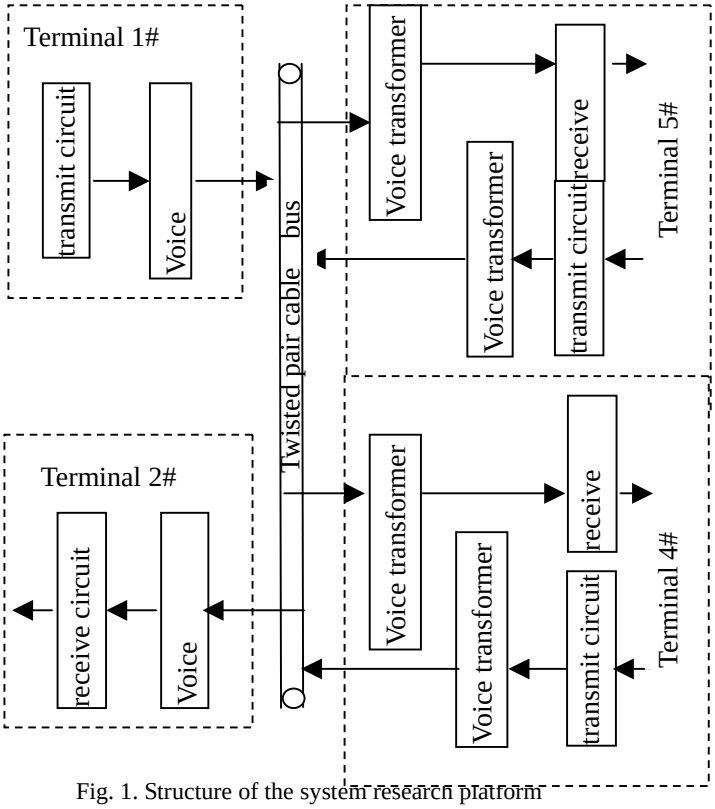


Fig. 1. Structure of the system research platform

3 The hardware design of digital base band signal long distance

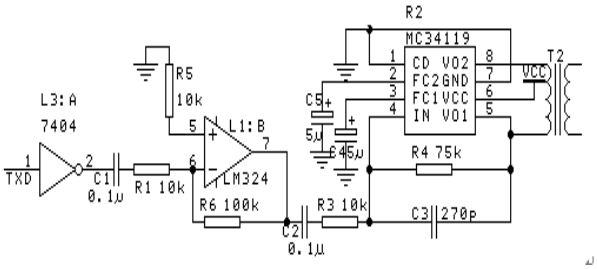


Fig. 2. Sending circuit

Receiving circuit of digital base band signals is showed as figure 2 and figure 3[3].

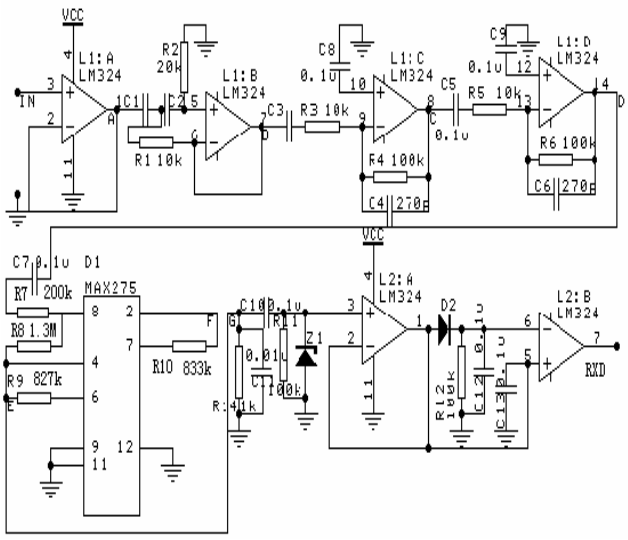


Fig. 3. Receiving circuit

4 Emulation research of digital base band signals transmission.

4.1 Emulation modal analysis of data circuit

Circuit data emulation is based on the circuit transmission theory. According to the data transmission theory and combining practical characteristic of utilizing cable, we can build the emulation model of transmission circuit to get the primary and secondary parameter and accordingly the amplitude value and phase delay of sine wave of different kind frequency, which pass through the circuit. Firstly we build the circuit model based on the relative theory. This system adopts the 0.9mm communication cable and the emulation model as it is shown in Fig. 4.

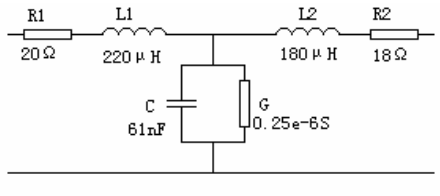


Fig. 4. Circuit model (km)

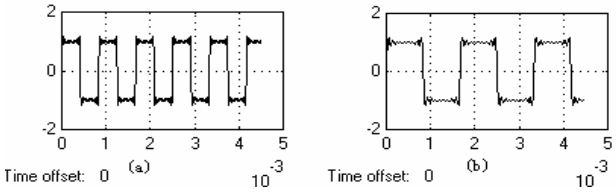
The primary parameter of this circuit is listed as the following : $R = 38\Omega$, $L = 400\mu H$, $C = 61nF$, $G = 0.25e-6S$ .we can calculate the attenuation α and phases-shift constant β [4,5].

1) Proceeding procedure of circuit data emulation

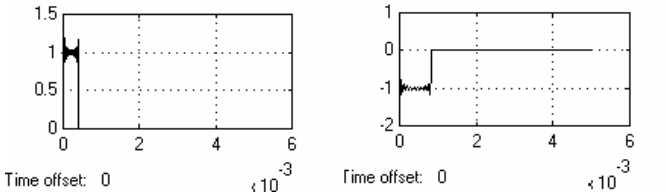
1. Decompose square wave according to Fourier formula to get the fundamental frequency and the coefficient of every harmonic wave;
2. Analyze the circuit model, programming and compile the M file and then run this procedure to get the amplitude value attenuation and phase delay of different kind harmonic wave.
3. Build emulation model, regard the variable in M file as model parameter and get the approximately waveform of square wave of this frequency after passing the circuit.
4. Deal with the wave and get emulation data.

The processing course to the circuit data is shown in Fig.5: the wave shown in Fig.5-(1) are 1200HZ and 600HZ ones ,which are compounded with Fourier series(take the signal 32 for example here, because the binary code in the transmission circuit of 32 is 0010011001,and only two bit are the same in this sequence, so it can be regarded as the combination of 12000HZ and 600HZ and just compound the square waves of the two frequency. The other datum can be concluded like this and omitted here).What is shown in Fig.5-(2) is the sub-waveband extracted from the compounded square wave, which need intercepting and time lapse and nestification to get the signals in the Fig.5-(3).

The carrying out of this process will be introduced in the following.

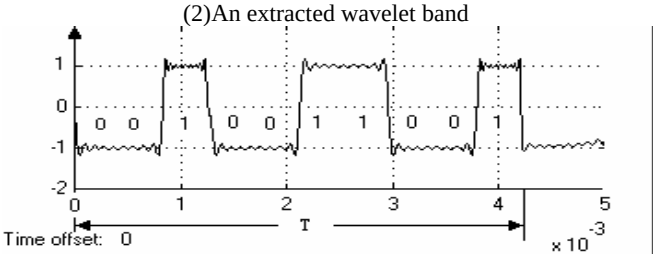


（1） Approximate square wave of 1200HZ and 600HZ after using Fourier series



(a) A positive level in 1200HZ square wave
(b) A negative level in 600HZ square wave

(2)An extracted wavelet band



(3)The approximate wave of the data 32 after processing (0 kilometre)

Fig. 5. The procession done with the inputing signals

2) Emulation model and M files

Fig. 6 shows the wavelet band model of 600HZ that we need. The left part of the dashed line is the compound square model, and the right part is created by setting the parameters of the STEP and taking out the wave band that we need. 9 modules of SINE WAVE make up the fundamental frequency signals,3、5、7、9、11、13、15、17 orders harmonic wave . and their parameters are transmitted through M files [5] .

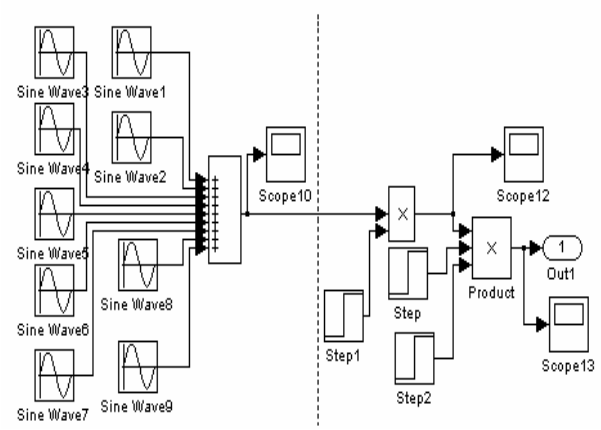


Fig. 6. One of the simulation models that create the wavelet band

4.2 The Receiving circuit simulation model

The Receiving circuit simulation model is designed as figure 7.

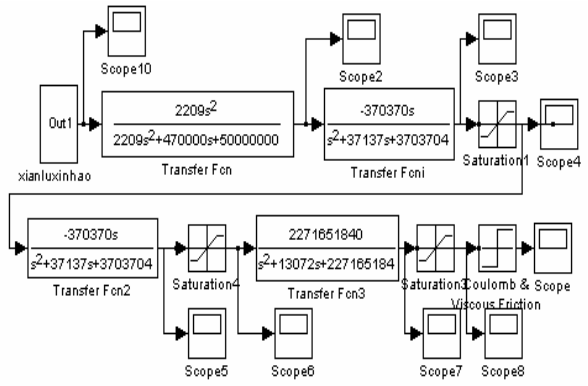


Fig.7 The Receiving circuit simulation model

4.3 The results of simulation and analysis

After building the simulating model and setting the proper parameters, we can simulate the models. When requiring observing signals from the circuit of the different length, what we have to do is nothing but double clicking the line signal module, setting the length of circuit and regulating the parameters. And then we can observe the results. Fig 8 shows the output waveform of the circuit of 1、7、15 and 20 kilometers.

We can see clearly from the waveform, when the length of circuits increase, the aberrance of the waveform prick up. The results of the experiment and the results of the simulation show no differences. And when the length of the circuit increases, the width of the high levels decreases gradually. And when this kind of change gets to a certain degree, the single chip can't receive the data rightly. It can be proved by the experiment that the distance of the data transmission can reach the 15 kilometers, and when it is beyond the 15 kilometers the error rate of the data after transmission increase rapidly, and this system can't work well. And when the distance of the circuit is short then 15 kilometers, the error rare is below 5%. And after software error correction, the system can work normally.

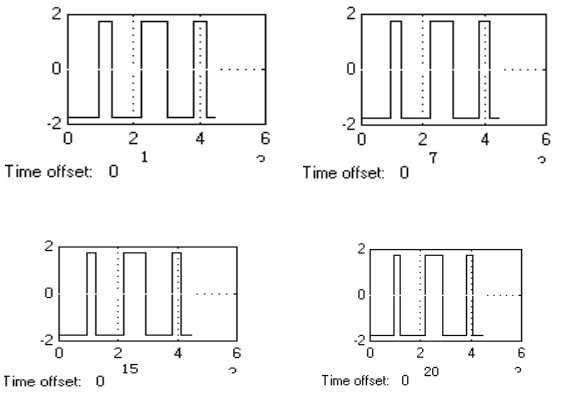


Fig. 8. The results after simulation

5 Conclusions

Low-pass implement the base-band digital signal transmission through the practical circuit and simulation technique. This technique can be applied to some the signal transmission of some definite purpose communicating system (such as signaling transmission of high-speed road emerge system). And the research meets the requirement of the remote transmission systems, which transmit only a few data.

References

[1] L.M. He, MCS-51serias application system designing and interface technology. Beijing Aeronautics and Astronautics publication .1991.3.
[2] Steve Lebson: The Handbook of Microcomputer Interfaceing, 1987.
[3] J.B. Robert, Microprocessors in Instruments and control, 1986.
[4] K.L. Vinay, G. John, Signal processing and implement by Matlab.
[5] M.D. Beer, Microcomputer Interfaceing, 1985.
[6] W. Leon, Digital and Analog Communication System. Tinghua University publication, Prentice Hall. 1998.6.

Non-primitive Turbo codes in MIMO System over Flat Fading Channel

Mingxin Tan
the Department of Information Technology, Central Chinese Normal University
Wuhan City, Hubei Province, 430079, China
Email: tmingxin@yahoo.com.cn

Abstract

In this paper non-primitive turbo code (NOPTC) is investigated based on the joint Turbo code and MIMO technology. Turbo code secures excellent bit error rate (BER) performance at moderate and higher signal to interference (SIR), however simulation results indicate that this code yields comparatively poor BER perform-ance when SIR is lower. At a lower and moderate SIR we here show that in multiple input multiple output (MIMO) system NOPTC provides better BER performance over Flat Fading channel than primitive turbo code (PTC) that is used in 3GPP standard as well as other applications. Thus NOPTC promotes both capacity and reliability of the communications system.

Keywords: Non-primitive Polynomial, Iterative Decoding, MIMO, Flat Fading, Turbo Code.

1. INTRODUCTION

In ICC'93 Berrou's paper [1] about Turbo code marks the channel coding enters new times in view of its extremely powerful coding scheme. In recent years, Turbo codes are gradually developed to combine with other technologies. In reference [2], simulation results demonstrate that the turbo-equalized system using convolution turbo codes is the most robust system for all code rates investigated. Reference [3] focuses on receiver structures for iterative channel coded systems and achieves both robustness and overall performance improvement. The integration of power control and turbo coding is adopted [4]. In SISO (single input single output) system BER performance of NOPTC is given over Added White Gausian Noise (AWGN) channel [5]. However, in spite of its excellent performance with respect

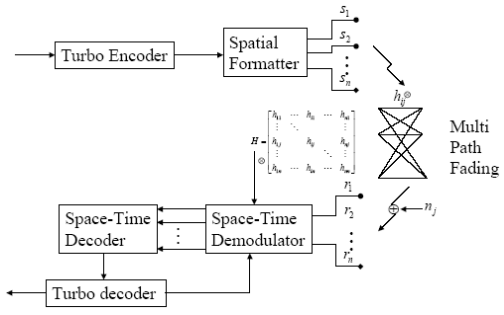


Fig. 1. System Block Diagram

to BER, the Berrou code has some disadvantage in its BER curve. Therefore, turbo code will be more widely applied in wireless communications once its Low SIR Performance is improved. So far, nearly all-previous research work concerning turbo codes only concentrated on PTCs whose encoders/decoders consist of both primitive feedback polynomials and primitive generator polynomials. This paper explores in MIMO system NOPTC's performance over Flat Fading channel, and discusses their merits when operating at lower and moderate SIR. It is arranged as follows: Section 2 describes the performance of NOPTC at a lower and moderate SIR, Section 3 gives simulation results and analysis. Section 4 concludes the paper.

2. THE SYSTEM MODEL OF JOINT MIMO AND TURBO CODE

In radio communications theory and wireless research, Turbo and space diversity [6] are two of the most popular technologies. Space diversity is used to combat various types of channel impairments [7]. Unlike time and frequency diversity, space diversity doesn't induce any loss in bandwidth efficiency. This property is very attractive for future high data rate wireless communications. Space diversity includes both receive and transmit diversity [8]. In this paper both transmit and receive diversity are employed.

In our design, Turbo code is considered as a channel code, and multiple antennas at both transmitter and receiver are used to overcome the interference over the Rayleigh Flat Fading channel. Figure 1 portrays the system block diagram.

The aim of the 3GPP (3rd generation partnership project)

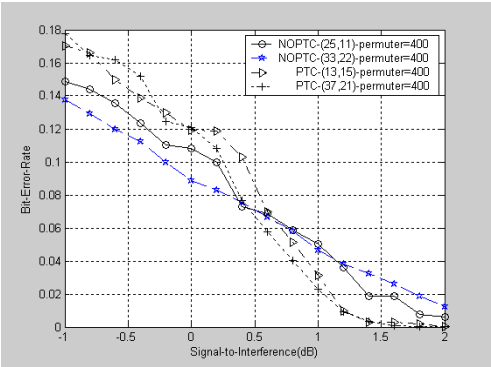


Fig. 2. BER comparison between NOPTC and PTC Over AWGN channel

Systems are to provide an "any where, any time" service, which has higher BER specification, and turbo code's remarkable characteristics make it possible to use in 3G. Reference [5] only investigates the performance of BER over AWGN channel (portrayed in Fig. 2.). But realistic scenario includes all kinds of interferences. This paper extends our research range to Rayleigh Flat Fading channel.

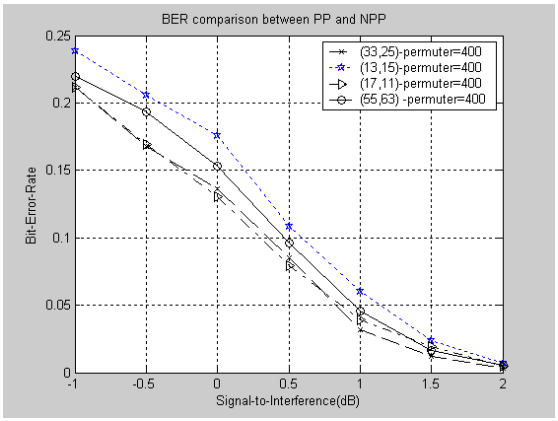
Primitive polynomial is a mathematic concept from 'Modern Algebra', and usually in coding theories encoders/decoders consisting of primitive polynomials have

better BER/FER (frame error rate) performance than non-primitive polynomials. Here several kinds of FEC (forward error correction) examples are provided to show PTC's wide application. For example, Cyclic Hamming codes [7, 4, 3], of which 7, 4, 3 stand for code length, information length and minimum Hamming distance, respectively, has the generator as follows: $g(x) = x^3 + x^2 + 1$. It is seen $g(x)$ is a primitive polynomial. So is the BCH code, such as BCH code [31, 11, 5], of which 31, 11, 5 denote code length, information length and maximum number of rectifiable error bits, respectively. Here the generator [31, 11, 5] is a primitive polynomial as well [9]. Likewise, convolution codes consisting of primitive polynomials are more considered, and usually convolution codes have fewer states (such as convolution codes [2, 1, 2] and so on) in order to reduce complexity. As to turbo code, since it was invented in ICC'93, a lot of research work has been developed and it has been applied in 3GPP standard and Digital Video Broadcast (DVB) standard for Return Channel Satellite (DVB-RCS). Almost all applications and previous research about turbo code are emphasized on PTC. Turbo code $(7,5)_8$ and $(13,15)_8$ are familiar PTC examples in authorized publications such as IEEE Journals and so on. In deed, in SISO system PTC has better BER/FER (frame error rate) performance than NOPTC when at moderate or higher SIR. Some probable reasons and analysis of poorer BER performance have been given at reference [10]. However, it does not provide practical measures to promote BER performance at lower SIR.

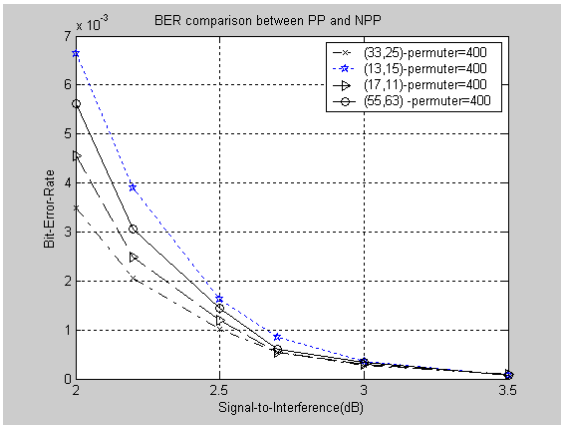
It is known the BER performance is poorer when at a lower SIR [11], and in everyday life we often face wireless communications environment of lower and moderate SIR. In a real channel, the received signal consists of a combination of attenuated, reflected, refracted, and diffracted replicas of the transmitted signal. In addition to all these factors, the channel adds noise to the signal and can cause Doppler effects. Because the fading of different antenna pairs is independent, there are more transmission paths, and there is smaller probability that deep fading occurs on all paths. The block fading channel model is used in this paper. In this model, the channel state remains fixed over a block (or frame) of a given size; the channel states may be independent or correlated between successive blocks. This model is a reasonable approximation to a wide variety of practical communication systems, such as OFDM, TDMA, FHSS systems and slow fading channels.

3. SIMULATION WITH ANALYSIS

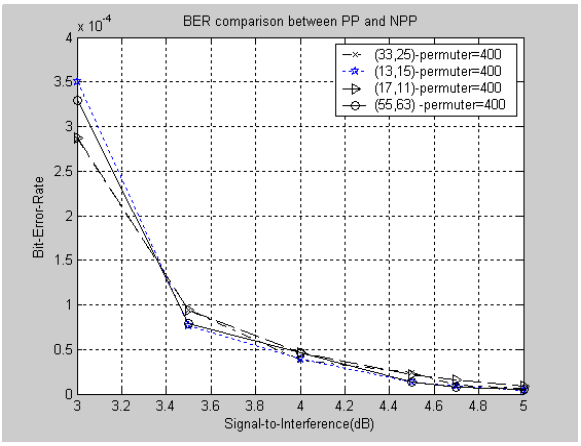
Figure 3 compares the BER curves, and they are shown here on a linear scale better to allow comparison resulting from the simulations of PTC and NOPTC. As stated above, the simulation circumstance is block fading. In our simulation, NOPTC $(17,11)_8$, $(33,25)_8$ and $(55,63)_8$ are selected that are 8-state, 16-state and 32-state non-primitive turbo code, respectively. PTC $(13,15)_8$ is selected as well, which is what we often read in publications that is chosen as 3GPP standard [12,13] for its better performance. The simulation conditions also include the following: interleaved depth 400 bits; code rate $1/3$; the number of iteration 5; the number of transmit antenna 8; the number of receive antenna 8.



(a) BER vs SIR at lower SIR



(b) BER vs SIR at moderate SIR



(c) BER vs sir higher SIR

Fig. 3. BER comparison between NOPTC And PTC over Flat Fading channel

From (a), (b) and (c) of figure 3, we can see that PTC $(13,15)_8$ has the worst BER performance before SIR is more than $3.4dB$. The BER value of the crossover point between NOPTC and PTC is between 1.3×10^{-4} and 1.5×10^{-4}. When $\mathrm{SIR} > 3.4dB$, PTC $(13,15)_8$ has best BER performance. Evidently, the cross-over point of curves in figure 3 occurs at a BER lower than the required specification for some practical services [14,15], especially

digital speech, which is usually the most tolerant to errors of widely used communication services.

Among the selected NOPTCs, $(55,63)_8$ has most states, and the transfer function for this 32-state constituent code is

$$G(D)=\left[1,\frac{g_1(D)}{g_0(D)}\right] \quad (1)$$

Where

$$g_0(D)=1+D^2+D^3+D^5 \quad (2)$$

$$g_1(D)=1+D+D^4+D^5 \quad (3)$$

From the simulation results of figure 3, it is seen that NOPTC $(55,63)_8$ doesn't have the best BER performance at lower or moderate SIR, or at higher SIR. NOPTC $(17,11)_8$ has smallest BER at lower SIR, and NOPTC $(33,25)_8$ at moderate SIR. The assertion 'the more state numbers turbo code has, the better performance turbo code achieves' doesn't meet in SISO system, neither in MIMO system. In a word it is not the state numbers that decide NOPTC's BER performance when the state number is not more than 32 at least.

It is seen those feedback polynomials that contain more '1' (not '0') in the feedback polynomials have lower BER value when working at the same SIR. The number of '1' doesn't include both Least Significant bit (LSB) and Most Significant bit (MSB). In convolution code and turbo code theory, LSB and MSB corresponding to a polynomial item stand for output and input, respectively. In feedback polynomial structure, apart from LSB and MSB, other '1's mean the information feeds back to the input, '0's mean the information doesn't feed back to the input. Therefore, in feedback polynomial structure, more '1' will contribute to recovering the original transmission information reliably.

In practical applications, for the 3G speeches code and the MPEG-4 video code [14,15], specification of $\mathrm{BER}=10^{-4}$ can implement some visible communications traffic, and speech communications can be finished at $\mathrm{BER}<10^{-3}$. In fact, sometimes $\mathrm{BER}=10^{-2}$ is enough to some practical application, such as slow fading channel. According to this standard, in speech communications, NOPTCs outperforms PTC when at lower and moderate SIR.

4. CONCLUSION

This paper compares the BER performance between NOPTC and PTC over Flat Fading channel. As a result, NOPTCs outperform PTC at lower and moderate SIR. According to 3GPP technical specifications, for speech services some visible artifacts communications such BERs are enough to establish and maintain communications in wireless communications circumstances.

5. REFERENCES

[1] C. Berrou, A. Glavieux, et al, 'Near Shannon limit error correcting coding and decoding: turbo codes', ICC, Proceedings, May 1993, 1064-1070.

[2] B.L. Yeap, J.Hamorsky, et al, Comparative Study of Turbo Equalization Schemes Using Convolutional, Convolutional Turbo, and Block-Turbo Codes, IEEE Transactions on Wireless Communications, April, 2002, Vol. 1(2): 266-273.

[3] Nenad, Veselinovic, Iterative Receivers for Interference cancellation and Suppression in Wireless Communications, Oulu University Press, 2004.

[4] Wenzhen, V. K. Dubey et al, The Performance of Turbo coding Over Power-controlled Fading Channel in Ka-Band LEO Satellite Systems, IEEE Transac- tions on Vehicular Technology, July, 2003, vol. 52(4): 1032-1043.

[5] M. Tan, A.G. Burr, Non-primitive Turbo codes, Proceedings of First International Symposium on Broadband Communications, Hollgate, England, 12-15, Dec. 2004,page61.

[6] B. Vucetic, J. Yuan, Space-Time Coding, Wiley Editorial Office 2003.

[7] D. Shiu, Wireless Communication Using Dual Antenna Arrays, Kluwer Academic Publishers, 2000

[8] S.L. Ariyavisitakul, Turbo Space Time Processing to Improve Wireless Channel Capacity, IEEE Transac-tions on communications, August， 2000vol. 48(8).

[9] S. Lin. An Introduction to Error Correcting Codes. Prentice-Hall, Englewood Cliffs, NJ, 1970

[10] L. C. Perez, J. Seghers, et al. A distance spectrum interpretation of turbo codes. IEEE Trans. Inform. Theory, vol.42, Nov, 1996: 1698-1709.

[11] O.Y. Takeshita, O. M. Collins, et al, On the Frame Error Rate of Turbo Codes, IEEE Transctions on Communications, Vol. 49, No. 4, April 2001.

[12] Third Generation Partnership Project "Multiplexing and channel coding", 3G TS 25.222 V3.2.0.

[13] Third Generation Partnership Project "Multiplexing and channel coding", 3G TS 25.212 (DRAFT).

[14] Third Generation Partnership Project "Quality of Service concept and architect- ture" (Release 6), 3GPP TS 23.107 V6.2.0 (2004-12).

[15] Third Generation Partnership Project "QoS Concept and Architecture" (Release 1999), 3GPP TS 23.107 V6.2.0 (2004-12).

Access Point Placement for Context-aware System

Chundong Wang
College of Information Technical Science, NanKai University
Tianjin 300071, China
Department of Computer Science & Engineering, Tianjin University of Technology
Tianjin 300191,China
Email: Michael3769@mail.nankai.edu.cn
And
Xiu-Feng Wang
College of Information Technical Science, NanKai University
Tianjin 300071, China
Email: wangxf@nankai.edu.cn
And
Ming Gao
Department of Computer Science & Engineering, Tianjin University of Technology
Tianjin 300191,China
Email: ten_minutes@163.com

ABSTRACT

The location information of the mobile user is an essential contextual element in a context-aware system. We investigated the optimal access point (AP) placement to enhance positioning accuracy, and presented a Being Away From Obstacles (BAFO) algorithm designed for context-aware systems using Received Signal Strength (RSS) to position objects. With BAFO algorithm, we can place APs by considering the circumstances such as the distribution of obstacles. The main objective of BAFO algorithm is to place APs away from obstacles, therefore minimizing the influence of obstacles on the signal strength. We made experiments to compare BAFO algorithm with hexagonal algorithm. Our conclusion is that BAFO algorithm is simpler and more effective.

Keywords: Access Point Placement, Context-aware, BAFO, Wireless Network.

1. INTRODUCTION

In context-aware system, we need to know where a user is in a given time, so a good positioning system is absolutely essential. The wireless devices and the positioning measures we choose determine the positioning accuracy, but especially in context-aware system having a RSS (received signal strength)-based positioning system, an optimal AP placement method is also significant to positioning accuracy.

In a common 802.11 network, the principle of the AP placement is minimizing the number of APs while maintaining a prescribed Quality-of-Service in the entire coverage area. In context-aware system, under the constraint of the cost, we must use more APs to calculate the position of the users, and we also need an effective AP placement algorithm to enhance positioning accuracy.

We mainly discuss how to use Being Away From Obstacles (BAFO) algorithm to calculate where the APs should be placed. In Section 2, we introduce the related works about AP placement. In section 3, we propose BAFO algorithm and discuss the results of the performance evaluation. In section 4, the conclusions are drawn and the directions of further research are pointed out.

2. RELATED WORKS

The current researches about AP placement are for different purposes, and only few of them are for the purpose of positioning. In these few researches, most are based on ideal network environment, which assumes that there are no obstacles in the entire coverage area.

Mari Kobayashi, Shinichiro Haruyama, Ryuji Kohno and Masao Nakagawa investigated the optimal access point placement in simultaneous broadcast system [1]. In the research, the disadvantage is that they only considered some particular terminal, and they couldn't get information of all terminals.

Y. Lee, K. Kim and Y. Choi presented a mathematical model in 2002 [2]. They considered both the optimal AP placement and assignment of channels. The advantage of their research is that they calculate the AP placement by considering the channel interference problem and load balancing.

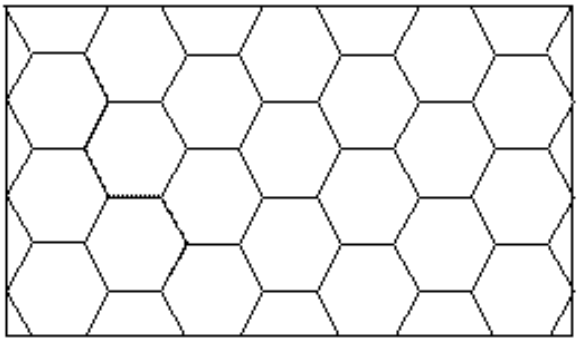


Fig. 1. Using hexagonal lattice packing method covering area

A direct and simple method is hexagonal lattice packing method shown in Finger1, and the BAFO algorithm is based

on this method. Now it is known that the best packing of circles in two dimensions is the 'hexagonal lattice' packing of circles, each touching six others [3]. In Figure.1, we use many hexagons covering the entire area, and place a AP in the center of each hexagon.

We can see that the above methods are not for the purpose of positioning or they don't take into account the actual geographic circumstances in the entire area. For this reason, we project Being Away From Obstacles (BAFO) algorithm.

3. BEING AWAY FROM OBSTACLES (BAFO) ALGORITHM

3.1. Assumptions

Now, we assume that we place the APs for the purpose of precise wireless positioning, and we also assume that the positioning method we will use is as follows. We first maintain the power signature of a set of positions. The position whose power signature is maintained in the database is called a reference point. Finger2 shows the reference frame of one hexagonal. Then we compare the power signature measured by the mobile terminal with the data recorded in the database, and then estimate the mobile terminal's position. According to the above positioning method, we know that our principle of AP placement is minimizing the number of access points in the area whose power signatures are equal.

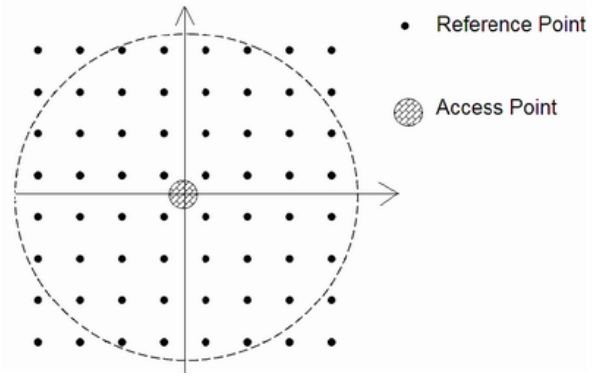


Fig. 2. The reference frame of hexagonal lattice packing method

We assume that there are *m* possible positions that AP can be placed in, and there are *n* reference points. Obviously, the simplest idea is to place AP in all the m possible positions, and for each possible position, we measure power signatures of the *n* reference points, and find the best one in the *m* possible positions. Thus we totally need to take O (*mn*) measurements. In general, the cost is unacceptable, especially when *m* is large or *m* is infinite, and that means AP can be placed in many positions or AP can be placed in any positions in the area. So our objective is to calculate the APs' optimal placement by taking just one measurement of the *n* reference points and to reduce the cost from O (*mn*) to O (*n*).

3.2. Mathematical model

Though our principle of AP placement is minimizing the number of access points in the area whose power signatures are equal, power signatures of each two positions cannot be absolutely equal in theory. So we need a mathematical formula to estimate whether a placement is optimal. Obviously, we can use the Standard Deviation, and choose the position where we get the largest Standard Deviation as the optimal position, but this method is not effective. We can assume that the power signatures can be separated into groups, power signatures in one group is low, but power signatures in the other group is high. Within the group, the power signatures are almost equal. We can get a large Standard Deviation now, but in this situation it doesn't fit for positioning. So a better method is to find out whether the power signatures are well distributed. It is known that the power signatures are well distributed when there are no obstacles, and any obstacles can block the signals and weaken the signal strength behind them, and then make the distribution of power signatures disperse. What we should do is to place the APs away from the obstacles, and to maximize the size of regions where the power signatures are well distributed. In addition, the total signal strength is maximized, and the signal becomes immune to the influence of other disadvantageous factors.

In Figure3, the region painted with thin color denotes the area that are influenced by the obstacle when we place AP in position j, and the sum of the area painted with thin color and the two areas painted with thick color denote the area that are influenced by the obstacle when we place AP in position i. Obviously, since i is nearer to the obstacle, the size of area that is influenced is larger. According to the figure3, we can tell that it's better to place AP in position j than in position i.

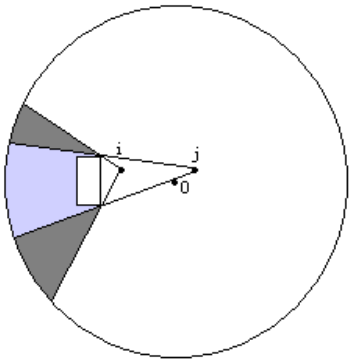


Fig. 3. The principal of BAFO algorithm

To get the AP's optimal position, we use a weight value to average the position of the reference points, and the weight value is $\sqrt{\frac{S'}{S}}$. *S* denotes the ideal power signature (no obstacles) at each reference point, *S'* denotes the actual power signature measured at each reference point. Because the signal strength is inversely proportional to the square of the distance between the AP and the position, we use the radicals of the ratio of *S'* to *S* as the weight value. If there is an obstacle near the reference point, *S'* will be smaller, and the weight value will be smaller correspondingly, and the calculated result will be away from the reference point.

3.3. Process of the BAFO algorithm

In the BAFO algorithm, we should divide the entire area into several independent child areas (these child areas can partly overlap), and in each child area we calculate one AP's position. Thus we can avoid the APs being too centralized in some positions. In this algorithm, we use hexagonal lattice packing method to divide the entire area. Initially we place the APs in the center of each hexagon (position calculated with hexagonal lattice packing method in ideal network

environment), and then we use BAFO algorithm to revise this ideal position.

The process of calculation is as follows:

Step 1: To divide the entire area using hexagonal lattice packing method. Assuming we need to deploy k APs, we place the APs in the center of each hexagon. We define the outer tangent circle of each hexagon as C_1, C_2, …, C_k, and the coordinate of each AP is (0,0) at initialization.

Step 2: To define n reference points in each C, according to each C_i, i=1, 2, …, k, the coordinate of each reference points is denotes as p_h= (x_h, y_h), h=1, 2, …, n, and then calculate the ideal power signature (no obstacles) S_h at each reference point.

Step 3: To measure the actual power signature S'_h at each reference point.

Step 4: To define R_i= (X'_i, Y'_i), and

$$X'_i = \frac{1}{n}\sum_{h=1}^{n} x_h \sqrt{\frac{S'_h}{S_h}}\ ,\ Y'_i = \frac{1}{n}\sum_{h=1}^{n} y_h \sqrt{\frac{S'_h}{S_h}}\quad i=1, 2,\ldots, k.$$

Step 5: To return the coordinate of R_i, (i=1, 2, …, k) as the result where the AP_i should be placed to.

3.4. Experiment Design

The experiments were carried in the sixth floor of Dept of Computer Science, Tianjin University of Technology. The client was implemented on an Intel-processor laptop with a wireless network card. The OS of the laptop is Linux Red Flag 4.0.

Table 1. The experimental results using hexagonal method

NNSS	Correct	Failed	Total	% Correct
Hexagonal method	35	15	50	70
BAFO	38	12	50	76

Table 2. The experimental results using MCCPS algorithm

MCCPS	Correct	Failed	Total	% Correct
Hexagonal method	40	10	50	80
BAFO	47	3	50	94

We conducted experiments based on two different AP placement methods. At first, we used hexagonal method to place the AP at the centre of the each hexagon, and then, we placed AP to the new location obtained from BAFO algorithm. In the two cases, we ran two different position algorithms, one was Nearest Neighbor(s) in Signal Space (NNSS) Algorithm which simply compared the power signature measured by the mobile terminal with the power signatures recorded in the database, and choose the best match. The other was MCCPS algorithm that fully consider of the signal strength of the reference points and signal strength near the reference points [4].

The experimental results show that compared with the hexagonal lattice packing method, BAFO method can enhance the positioning accuracy of NNSS algorithm by 6 percent, and the positioning accuracy of MCCPS by 14 percent. We can undoubtedly see that BAFO algorithm is more effective when we use a positioning algorithm that takes into account power signatures of both reference point itself and its neighbors.

4. CONCLUSIONS AND FUTURE WORK

This paper presented a Being Away From Obstacles (BAFO) algorithm. BAFO can calculate the optimal placement of APs according to the actual geographic circumstances in the entire area. In further research, we can focus on following aspects.

If we have more information of the geographic circumstances of the area, we will have more accurate position results, but the costs of measurement will increase. In the algorithm presented in this paper, we made measurements for all reference points just once when we place the APs in the center of the Hexagons. Can we increase the time of measurements by placing APs in different positions to enhance the positioning accuracy? If we can, how to balance the positioning accuracy and the cost of measurement is the direction of further research.

In BAFO, we use hexagonal lattice packing method to divide the entire area, and the size of overlapped area is small. We can get more information of the entire area by making the child areas largely overlap, but this will increase the amount of computing and may make the APs centralized in particular position. For different purposes, we can use other methods to divide the entire area.

5. REFERENCES

[1] M. Kobayashi, S. Haruyama, et al, “Optimal Access Point Placement in Simultaneous Broadcast System using OFDM for Indoor Wireless LAN”, Proceedings of the 11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2000,1:200-204

[2] Y. Lee, K. Kim, et al, “Optimization of AP Placement and Channel Assignment in Wireless LANs”, Proceedings of the 27th Annual IEEE Conference on Local Computer Networks, 2002,831-836.

[3] N. SLOANE, “The sphere-packing problem”, Overview for ICM-July 1998, July 1998.

[4] M.H. Jin, E.H.K. Wu, et al, “802.11-based Positioning System for Context Aware Applications” Proceedings of Communication Systems and Applications, 2004

QPSO-based Algorithm for Multi-objective Optimization Problem

Fangjing Guan, Wenbo Xu, Jun Sun
Computer Engineering School of Information Technology, Southern Yangtze University, Wuxi Jiangsu, 214122, China
Email: guan_fj@163.com

ABSTRACT

Quantum-behaved Particle Swarm Optimization Algorithm is a global convergent optimization algorithm. In this paper, we propose a Vector Evaluated Quantum-behaved Particle Swarm Optimization algorithm (VEQPSO) to solve multi-objective optimization problem. After described the concepts of VEQPSO, we test it on multi-objective benchmark problems. The experiment results of VEQPSO on four classic multi-objective minimization functions show the efficiency of the algorithm for solving multi-objective optimization problem.

Keywords: Particle swarm optimization, multi-objective optimization, Quantum-behaved particle swarm optimization.

1. INTRODUCTION

Multi-objective optimization problems are very common, especially in scientific research and applications, optimization aimed at these problems are called multi-objective optimization problems. Algorithm for multi-objective optimization problems has developed rapidly. In Multi-objective optimization problems, there are collisions between objects, usually, it is impossible to achieve best solution for every object simultaneously, and solutions for multi-objective problems are a set, not a single one. In traditional multi-objective optimization solutions, the objectives have to be aggregated in a single objective function, and then get the solution by single objective optimization method.

In this paper, based on VEGA (Vector Evaluated Genetic Algorithm), and QPSO (Quantum-behaved Particle Swarm Optimization) [1, 2], A Vector Evaluated QPSO algorithm, VEQPSO for short, is presented to solve multi-objective problems. The experiment results of VEQPSO on some classic multi-objective minimization problems show the efficiency of the algorithm. The conclusion is given supported by some classic functions.

2. PSO ALGORITHM

PSO algorithms are especially useful for parameter optimization in continuous, multi-dimensional search spaces. PSO is mainly inspired by social behavior patterns of organisms that live and interact within large groups. In particular, PSO incorporates swarming behaviors observed in flocks of birds, schools of fish, or swarms of bees. PSO assign direction vectors and velocities to each point in a multi-dimensional search space. Each point then 'moves' or 'flies' through the search space following its velocity vector, which is influenced by the directions and velocities of other points in its neighborhood. The outline of PSO algorithm is given as follow:

Assuming that the search space is n-dimensional, the i-th particle of the swarm is represented by the n-dimensional vector $X_i = (x_{i1}, x_{i2}, \cdots, x_{in})$, the current velocities of the i-th particle is represented by $V_i = (v_{i1}, v_{i2}, \cdots, v_{in})$. The best previous position (*pbest*) of the i-th particle is represented by $P_i = (p_{i1}, p_{i2}, \cdots, p_{in})$, called individual highest fitness; *gbest* is the global best fitness of the swarm.

pbest is formulated as follow:

$$p_i(t+1) = \begin{cases} P_i(t) & if\ f(x_i(t+1)) \geq f(P_i(t)) \\ x_i(t+1) & if\ f(x_i(t+1)) < f(P_i(t)) \end{cases} \quad (1)$$

Velocities and positions of the particles are updated with the following equations:

$$\begin{aligned} v_{i,d}(t+1) &= wv_{i,d}(t) + c_1 r_{1,d}(t)(pbest_{i,d}(t) \\ &- x_{i,d}(t)) + c_2 r_{2,d}(t)(gbest(t) - x_{i,d}(t)) \\ x_i(t+1) &= x_i(t) + v_i(t+1) \end{aligned} \quad (2)$$

where: $d = 1, ..., n$; W is the inertia weight; c_1 and c_2 are two positive constants; $r_{1,d}(t), r_{2,d}(t)$ are two random numbers within the range [0, 1], $pbest_{i,d}(t)$ (the local best position) is the cognitive component because of only thinking of the particle's experience, $gbest(t)$ (the global best position) is the social component.

3. QPSO

QPSO [1, 2] presented previous, not only have few parameters, but also have search ability advantages over the existing PSO algorithms. In QPSO, particles change their positions follow the three formulations:

$$mbest = \frac{1}{M}\sum_{i=1}^{M} P_i = \left(\frac{1}{M}\sum_{i=1}^{M} P_{i1}, ..., \frac{1}{M} P_{ij}\right) \quad (3)$$

*This work is supported by National natural science foundation of china (60474030)

$$PP_{ij} = f \times P_{ij} + (1 - f) \times P_{gj}$$

$$f = rand \tag{4}$$

$$x_{ij} = PP_{ij} \pm a \times | mbest_j - x_{ij} | \times \ln(1/u) \tag{5}$$

$$u = rand$$

where $mbest$ is the middle position of $pbest$, PP_{ij} is a random point between P_{ij} and P_{gj}. a called creativity coefficient, is the vital parameter in QPSO convergence, at T -th iteration,

$$\alpha = 0.5 + 0.5 * (MAXTIME - T) / MAXTIME \tag{6}$$

(the concrete value of a depends on the concrete conditions), where MAXITER is the maximum number of iterations. QPSO is the following:

1) Initialization

For T=1: MAXTIME

2) Compute the values of particles' objective functions according to formulations.

3) Update the new local best position of the particle: P_{ij}.

4) Update the global best position: P_{gj}.

5) Compute the value of ***mbest*** according to Eq. (3).

6) Compute PP_{ij} (the random point of each particle) using Eq. (4).

7) Update new positions of particles according to Eq. (5).

end

Repeat step 2～7 until meet the maximum number of iterations.

The difference between PSO and QPSO is the way of evolution (the way of updating particle's position). In PSO algorithm, particles must be in a limit search space in order to assure the swarm to meet the global convergent optimization solution. But in QPSO, particles can be at arbitrary position in search space according to probability, perhaps the position is better than *pbest*.

4. SOLUTIONS OF MULTI-OBJECTIVE OPTIMI_ ZATION WITH QPSO

4.1 Basic Concepts of Multi-Objective (MO) [7]

Assuming, in a decision making process, k objectives are to be inspected, and the objective functions required to be the smaller when satisfy the constraint conditions. Optimization problem is described as follows:

$$\min_{X \in R} F(X)$$

$$F(X) = (f_1(x), f_2(x), \cdots, f_k(x))^T$$

$$R = \{X \mid g(X) \le 0\}, \tag{7}$$

$$g(X) = (g_1(X), g_2(X), \cdots, g_m(X))^T$$

$$X = (x_1, x_2, \cdots, x_n)^T, X \subset R \subset E^n$$

where: $F(X)$ are the objective functions, $g(X)$ is constrained vector. In most cases it is impossible to obtain for all objectives the global minimum at the same point. The goal of MO is to provide a set of Pareto Optimal solutions to the aforementioned problem.

For multi-objective problems, it is difficult to determine which particle is the best, so Pareto Dominance is used to define the Pareto Optimal points.

Definition 1: Let $u = (u_1, u_2, \cdots, u_k)^T$ and $v = (v_1, v_2, \cdots, v_k)^T$ be two vectors. Then u dominates v, if and only if

$$u_i \le v_i, i = 1, 2, \cdots, k \tag{8}$$

and

$$u_i < v_i$$

for at least one component.

Definition 2: To minimization MO problem, there is a objective vector $u = (f_1(X_u), f_2(X_u), \cdots, f_k(X_u))^T$, where $X_u \in R$ is decision vector, X_u is Pareto optimal solutions, then: only if there does not exist decision vector X_v, such that objective vector $v = (f_1(X_v), , f_2(X_v), \cdots f_k(X_v))^T$, dominates objective vector $u = (f_1(X_u), f_2(X_u), \cdots, f_k(X_u))^T$.

A set of solutions, which meet the Definiaiton2, is called Pareto Optimal Set of an MO problem or Non-inferior set; each solution in the set is a Pareto Optimal solution, called Non-inferior Solution or Effective solution.

4.2 A Vector Evaluated Quantum-behaved Particle Swarm Optimization (VEQPSO) [4, 7]

VEQPS0 (Vector Evaluated QPSO) is inspired on a Vector Evaluated Genetic Algorithm (VEGA) which proposed by Schaffer to solve MO problem.we use the evolution method of multi-objective and adjust the selected method of simple evaluated algorithm in VEGA algorithm. In each iteration, all objectives are looked as single objective optimal problem, we evaluate each individual respectively, calculate fitness of the individual. Then, to each objective function, we select individuals to generate a sub-population with fitness rate. If there are *k* objective functions in optimal problem, *k* sub-populations are generated, the individual in sub-popula_ tion stands for a set of optimal solution of certain objective function. All sub-populations together, crossover and mutation are applied to generate the new swarm.

The main ideas of VEGA were adopted and modified to fit the QPSO, developing the VEQPSO algorithm. The framework of the VEQPSO algorithm is shown in Fig. 1. .

Following, we describe the realization of VEQPSO algorithm for two objects in detail. Assuming, there is an objective vector $F(X) = (f_1(X), f_2(X))^T$, then the algorithm of VEQPSO is the following:

Setup 1: Initialize the population:

Generate N particles randomly in search space, every particle is represented by vector, weight of a vector is uniformly distributed randomly.

Setup 2: Generate sub-populations

Assuming, *Nsg* is the sub-swarm size, then $Nsg = N / A$, where *A* is the objective number. If there are two objectives, then there are two sub-populations, every sub-population size is $N/2$ particles. Using the evaluation method of the single objective, all particles are evaluated according to $f_1(X)$ and $f_2(X)$. Computing the value of all articles is based on $f_1(X)$ and $f_2(X)$, then better *Nsg* particles selected are applied to generate sub-population1 and sub-population2.

Setup 3: Obtain shared information

The best particle of each swarm is assured, then becomes the population's social shared information.

Setup 4: Update particles' position according to Eq. (4) and Eq. (5).

Particles of different sub-populations update their positions according to social shared information from the other sub-population. In terms of QPSO model, the best particle of the second swarm ($gbest2$) was used for the determination of the new velocities of the first swarm's particles, but the best particle of the first swarm ($gbest1$) was used for the determination of the new velocities of the second swarm's particles. After shuffling all particles are applied to generate the new population.

Setup 5: Repeat above-mentioned **Setup 2**～**Setup 4** until terminate condition is reached. Finally, a set of Pareto Optimal solution is selected from two sub-populations to compose Pareto Optimal solution set, and output as optimal result.

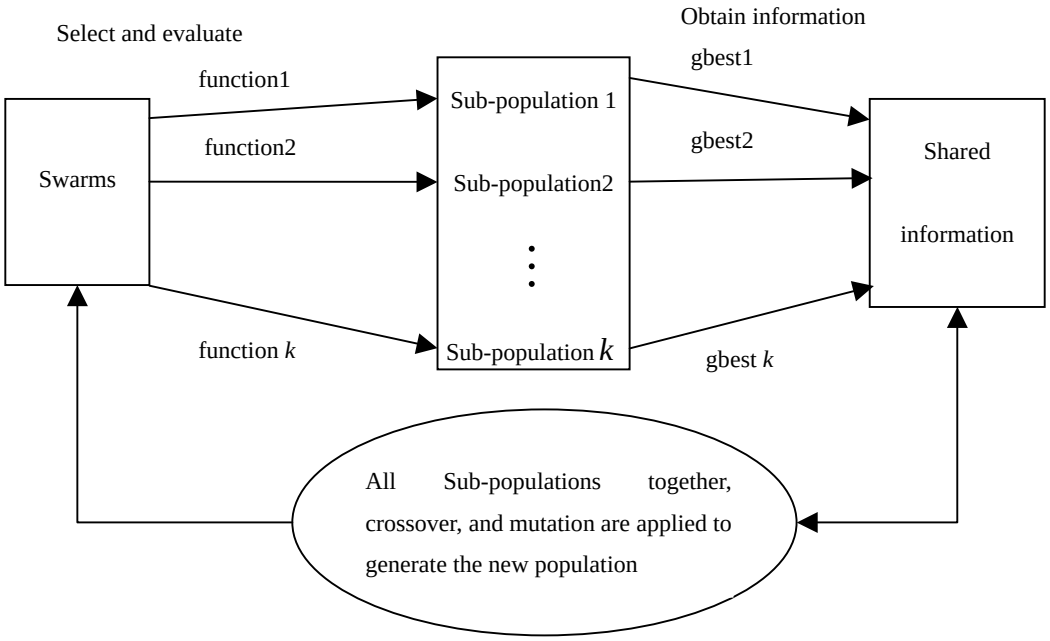


Fig. 1. The framework of VEQPSO

5. NUMERICAL EXPERIMENTS

Four classic multi-objective test functions are used to test the efficiency of VEQPSO.

Test function1 [6, 8]

$$\min f_1 = x^2$$

$$\min f_2 = (x-2)^2$$

where: $0 \le x \le 1$. We apply VEQPSO to solve this problem. The experiment sets as follow:

The population size of QPSO is 20 particles, the maximum number of iteration is 150, the value of $\boldsymbol{a}$ is decreased form 1.0 towards 0.6 following the number of iterations increased

$\alpha = 0.6 + 0.4 \times (MAXTIME - T)/MAXTIME$

where: *MAXITIME* is the maximum number of iterations, *T* is the current number of iterations.

After 150 iterations, 20 non-dominated solutions out of 20 solutions were found, the results are shown in Fig. 2. :

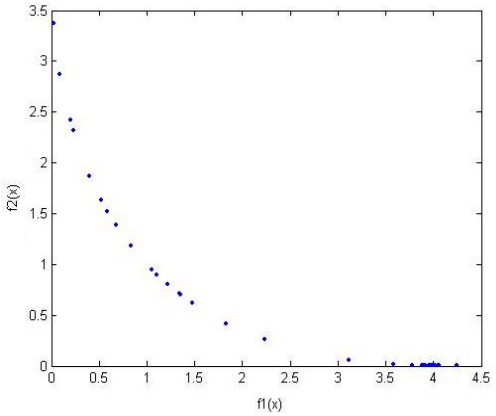


Fig. 2. The results of Test function 1

Test function 2 [5, 8]:

$$\min f_1(x, y) = (x^2 + y^2)^{\frac{1}{8}}$$

$$\min f_2(x, y) = ((x-0.5)^2 + (y-0.5)^2)^{\frac{1}{4}}$$

where: $-5 \le x \le 10$, α=0.7+0.3×(150-T)/150 .

After 150 iterations, 90 non-dominated solutions out of 100 solutions were found; the results are shown in Fig. 3. :

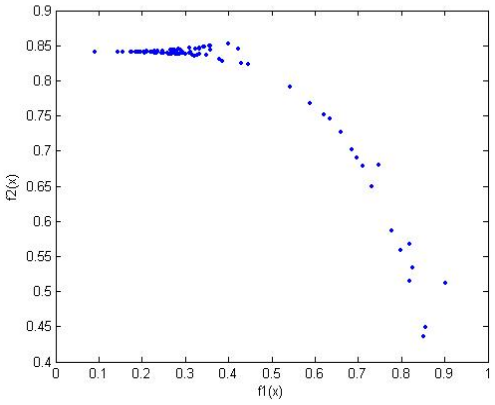


Fig. 3. The results of Test function 2

Test function 3 [3]:

$$\min f_1 = x_1, g = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i$$

$$\min f_2 = 1 - \sqrt{f_1 / g}$$

where: $0 \le x \le 1, n = 30$, α=0.5+0.5×(150-T)/150 .

After 150 iterations, 20 non-dominated solutions out of 20 solutions were found; the results are shown in Fig. 4. :

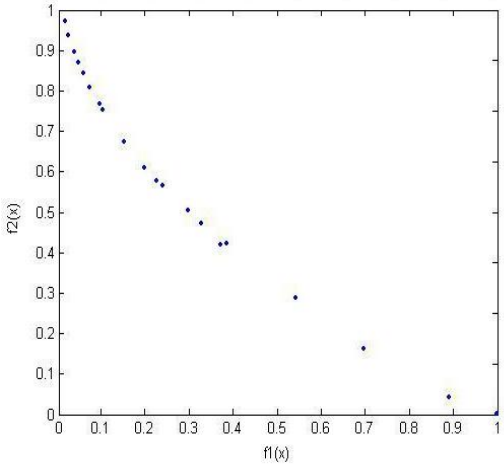


Fig. 4. The results of Test function 3

Test function 4 [3]:

$$\min f_1 = x_1;$$

$$g = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i;$$

$$\min f_2 = 1 - \sqrt[4]{f_1 / g} - (f_1 / g)^4$$

where: $0 \le x \le 1, n = 30$, α=0.5+0.5×(150-T)/150 .

After 150 iterations, 20 non-dominated solutions out of 20 solutions were found; the results are shown in Fig. 5. :

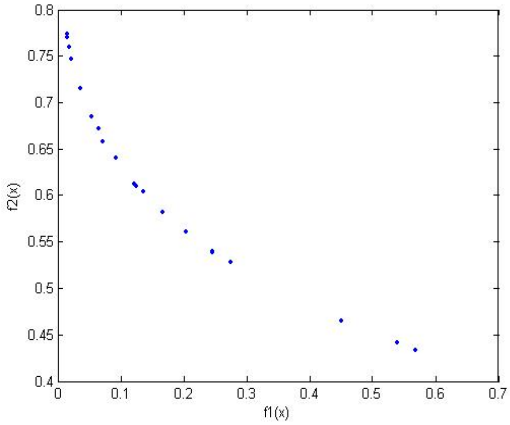


Fig. 5. The results of Test function 4

We also apply VEQPSO algorithm to other multi-objective problems, and the results is validated. When PSO is applied to solve the same problems, for simple problems, results of PSO is similar to QPSO, but for complex problems, QPSO is better than PSO, QPSO have a faster convergence velocity, because QPSO is an optimal algorithm designed for global convergence.

6. CONCLUSION

A Vector Evaluated Quantum-behaved Particle Swarm Optimization (VEQPSO) based on QPSO for solving multi-objective optimization problem is proposed. The algorithm, which is inspired by VEGA, divides the population into several sub-populations to search out the optima of the problem. The experiment results of VEQPSO on four classic multi-objective minimization problems show the efficiency of the algorithm to solve multi-objective problems.

7. REFERENCES

[1]. J. Sun, W.B. Xu, “A Global Search Strategy of Quantum-behaved Particle Swarm Optimization”, Proceedings of IEEE conference on Cybernetics and Intelligent Systems, 2004, pp.111−116.

[2]. J. Sun, B. Feng, et al, “Particle Swarm Optimization with Particles Having Quantum Behavior”, Proceedings of 2004 Congress on Evolutionary Computation, 2004, pp. 325−331.

[3]. K.E. Parsopoulos, M.N. Vrabatis, “Particle Swarm Optimization method in Multiobjective Problems” In: Proceeding ACM Symposium on Applied Computing (SAC 2002), 2002, pp.603−607.

[4]. Y. Jin, M. Olhofer, et al, “Dynamic Weighted

Aggrega_
tion for Evolutionary Multi-Objective Optimization: Why Does It Work and How?" In: Proceedings GECCO 2001 Conference, 2001.

[5]. X.B. Hu, Russell Eberhart, "Multiobjective Optimization Using Dynamic Neighbourhood Particle Swarm Optimization", Proceedings of IEEE conference on Cybernetics and Intelligent Systems, 2002, pp.1677−1681.

[6]. J.D. Schaffer, "Multiple objective optimization with vector evaluated genetic algorithms", The First Int'l Conf on Genetic Algorithms, Lawrence Erlbaum, 1985.

[7]. J.C. Zeng, Q.Jie, et al, "Particle Swarm Optimization" [M]. the Science Publishing House of Peking, first version, 2004,pp.122−127.

[8]. L.B. Zhang, C.G. Zhou, et al, "Solutions of Multi-Objective Optimization Problems Based on Particle Swarm Optimization", Computer research and development, Vol.41, No.7, July 2004, pp.1286−1291.

A Novel Bus-based Interconnection Topology For Small-scale Multiprocessor System And SOC Multiprocessor System *

Yunni Xia, Hanpin Wang, Yu Huang
School of electronics engineering and computer science of Peking University, Beijing, China
Beijing, 100871, China
Email: xiayunni@pku.edu.cn

ABSTRACT

This paper presents a novel topology, partial-bus, as an alternative to traditional single/multiple shared bus. Through a discrete-time homogeneous Markov model which analytically calculates throughput, the research compares the performance of partial-bus with that of single/multiple bus. The performance comparison indicates that partial-bus may well be a good alternative to single/multiple bus for small scale and SOC multiprocessor systems. To ensure bus allocations of partial-bus is deadlock-free, a Petri-net based model is also introduced to formally describe bus occupation behaviors in partial-bus and verify the deadlock-free property.

Keywords: interconnection topology, homogeneous DTMC model, throughput, performance evaluation, Petri net.

1. Introduction

The performance of multiprocessor system depends heavily on its interconnection network. Researches of recent years mainly focus on interconnections of large-scale (distributed shared-memory) multiprocessor systems. However, interconnections for small-scale multiprocessor systems draw less attention. System manufacturers have long been using bus based interconnections, especially single/multiple buses [1], to build small-scale machines. On the other hand, bus-based interconnections are also widely used in on-chip multi-core systems [2,3] because they are cheap, easy to manufacture and energy efficient. AMBA [4], Core Connect [5], Core Frame [6], HIBI, Lottery-bus, Marble are some commonly used bus-based SOC interconnections. Although these SOC interconnections are different in physical details, they all incorporate single-bus-like (such as AHB/ASB in AMBA and MBus in Core-frame) or multi-bus-like (such as PLB/OPB in Core-Connect) devices to connect their processing modules (DSP or CPU) and on-chip memories. Consequently, SOC interconnections mentioned above take a similar form to single/multiple bus.

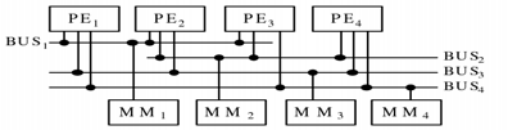


Fig. 1. A partial-bus sample *PB(4)*

For simplicity, notations *SB(n)* and *MB(n,m,b)* are used to denote a single-bus with n processors and a multiple-bus with n processors, m memory modules and b buses, respectively. Note that in *MB(n,m,b)*, m, n, b are subject to $b \le min(m,n)$.

This paper proposes a novel bus-based interconnection topology, partial-bus, for small-scale multi-processor systems (also as the building block of large-scale systems and substructure for hierarchical/multi-stage systems) and on-chip multiprocessor systems. A partial-bus topology consists of n ($n \le 8$) processor elements (PE), n buses each connected to only one memory module and $n-1$ PEs. The n buses cover all combination of $n-1$PEs. For simplicity, a notation *PB(n)* is used to denote such a topology, as illustrated by Fig. 1.

In a *PB(n)*, every bus has n ports (connecting 1 memory module and $n-1$ PEs), in contrast, every bus in *SB(n)* and *MB(n,m,b)* has $n+1$ and $n+m$ ports, respectively. The reduction of ports may endow system based on partial-bus with: lower hardware complexity and simpler circuit, lower manufacturing cost, lower energy consumption (especially beneficial to SOC applications), higher hardware reliability and shorter bus-cycle time.

Also, the Petri-net based formal model in section 3 indicates that deadlock never occurs in the bus occupation behaviors of Partial-bus and its detailed proof is given in appendix. Many simulation models [7,8,9] were proposed to evaluate the performance of multiprocessor systems and interconnection networks. On the other hand, various analytical models have also been developed [10,11,12], most of which are based on queuing theory or stochastic process. Section 4 employs a discrete-time homogeneous Markovian model to analyze the performance of *PB(n)* and to compare it with *SB(n)* and *MB(n,m,b)*.

2. MEMORY OPERATIONS ON PARTIAL-BUS

To characterize memory operations in partial-bus system, definitions listed as below are used

Requesting PE a PE wishing to access memory is defined as a requesting PE

Active/Inactive a PE is defined as active if it is performing a memory access, otherwise inactive

* Supported by supported by the National Grand Fundamental Research 973 Program of China under Grant No. 2002CB312004. The corresponding author is Hanpin Wang

Occupied/Idle a bus is defined as occupied if some PE is occupying this bus, otherwise idle

Local bus for a specific PE, the *n-1* buses it directly connects are defined as its local buses. For instance, in the example shown in Fig.1, Bus_1, Bus_3 and Bus_4 are local buses of PE_1

Remote bus for a specific PE, the bus it indirectly connects is defined as its remote bus. For instance, in Fig.1, Bus_2 is the remote bus of PE_1

Local memory module for a specific PE, the memory modules attached to local buses are identified as its local memory modules

Remote memory module for a specific PE, the memory modules attached to remote buses are identified as its remote memory modules

It is obvious that, to access a local memory module, the requesting PE occupies the local bus attached to the requested memory module. However, to access a remote memory module, the requesting PE has to turn to another PE (which can locally access the requested memory) for assistance. Therefore, four more definitions follow

Local access an access to a local memory module by a PE is defined as a local access. In a local access, the requesting PE occupies the local bus attached to the requested memory module

Medium PE since the requesting PE can not directly access its remote memory module, another PE which can locally access the requested remote memory has to assist the requesting PE. This co-working PE is defined as the medium PE. Note that for any specific requesting PE wishing to access a specific remote memory module, there exist *n-1* candidate medium PEs. It is assumed that the *n-1* candidate medium PEs are equally likely to be selected

Table 1. Definition of Petri-net element

Element	t or p	Description
$Inactive_i$	p	PE_i is idle without memory request
$Wait_i$	p	PE_i waits to occupy bus
$Read\_L_{ij}$	p	PE_i ready for a local access with the local bus Bus_j occupied
$Read\_R_{ijk}$	p	PE_i ready for a remote access with the local medium bus Bus_j and the remote bus Bus_k occupied
B_j	p	A token in this place indicates Bus_j is idle and available
Rq_i	t	PE_i generates a memory request
Occ_{ij}	t	PE_i occupies the local bus Bus_j for a local access
$A\_o\_L_{ij}$	t	PE_i performs a local access with the local bus Bus_j occupied
$A\_op\_R_{ijk}$	t	PE_i performs a remote access with the local medium bus Bus_j and the remote bus Bus_k occupied
Occ_B_{ijk}	t	PE_i occupies both the local medium bus Bus_j and the remote bus Bus_k synchronously, for a remote access

Medium bus in a remote access, the local bus connecting the requesting PE and the medium PE is defined as the medium bus. Note that, for a specific PE wishing to access a specific remote memory module, there exist *n-1*candidate medium buses. It is assumed that the *n-1*candidate medium buses are equally likely to be selected

Remote access an access to a remote memory module by a PE is defined as a remote access. In a remote access, the requesting PE occupies the remote bus attached to the requested memory module, chooses a medium PE and occupies a medium bus

For instance, as shown in Fig. 1. , PE_1 can choose any one of PE_2, PE_3 and PE_4 as the medium PE, to perform a remote access to MM_2. Bus_1, Bus_3 and Bus_4 are the three candidate medium buses which are equally likely to be chosen by PE_1

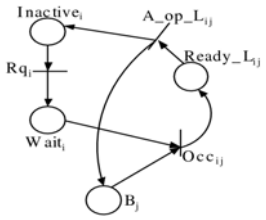


Fig. 2. PT for local access

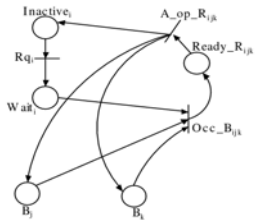


Fig.3 PT for remote access

3. BUS OCCUPATION WITH LIVENESS PROPERTY

This section models partial-bus memory operations through P/T nets (given by Fig. 2. and Fig. 3.) and verifies the liveness property. For brevity, definitions, notations and methods of P/T net formalism are directly used with out introductions. Readers may refer to [13] for details. Since every bus connects only one memory module, a memory module is accessible to a PE only if the PE already occupies the bus attached to the memory module. Therefore, the model based on P/T net only focuses on state changes of buses and the status of PEs are neglected. Definitions of net elements are given by Table 1. In this table, place $Inactive_i$, $Wait_i$, $ready\_L_{ij}$, $ready\_R_{ijk}$ serve as state indicators of PE_i, while B_j indicates the state of Bus_j (a token in this place indicates Bus_j is idle, otherwise occupied). Fig.2 models the local access to MM_j by PE_i, with Bus_j to be occupied. Transition $A\_op\_L_{ij}$ releases an occupied bus by generating a token back into B_j and making Bus_j idle.

On the other hand, remote access (shown by Fig. 3.) is a bit more complicated since every PE has to occupy two buses. To prevent deadlock, the requesting PE occupies both the medium bus and the remote bus synchronously, if both buses are available. Otherwise, the requesting PE occupies neither of them and waits.

It is also assumed that the control policy randomly and non-prioritizedly chooses from competing memory requests and allocates buses. Proof of the liveness property (no deadlock occurs and every memory request is eventually satisfied) is based on **net synthesis** discussed in the appendix.

4. PERFORMANCE EVALUATION MODEL

This section employs a discrete-time homogenous Markovian model to evaluate the performance of *PB(n)*. Throughput (the average number of memory accesses that system serves in a cycle) is used as the performance metric. Note that this model ignores implementation details such as bit-widths, memory/cache capacity, PE/bus frequency, control

signal format, pipelining mechanism, etc. Before the model is presented, some assumptions are given as: (1) Time is divided into discrete time units, bus-cycles [10] (cycle for short), (2) All PEs are statistically identical and independent, (3) Each inactive PE can issue only one memory access request at a cycle, (4) If memory access request is denied by the control unit, the requesting PE remains inactive, (5) Memory requests are distributed equally among all memory modules. That is to say, the probability that a memory request addresses to a specific memory module is *1/n*, (6) Every bus can be occupied by only one PE in a cycle (7) A non-prioritized strategy is used by the control policy. Therefore, the control policy non-prioritizedly selects one of conflicting memory requests and permits its bus occupation, (8) When an active PE accomplishes its memory access, it releases all buses it occupies and becomes inactive, (9) The probability that an inactive PE generates a memory request in the next cycle is *g*, (10) The probability that an ongoing local access continues in the next cycle is a, and the number of cycles local access lasts is geometrically distributed with parameter *1-a*. Therefore, the average cycle length of local access is $(1\text{-a})^{-1}$, (11) the probability that an ongoing remote access continues in the next cycle is a_2. Remote access takes u more cycles than local access on average, therefore $1/(1-a_2)=1/(1-a)+u$, (12) For a remote access, the probability that the requested data exist in the cache of medium PE is *h*. This kind of cache hit is termed as **R_hit**, (13) Every PE can perform its own memory operations while serving as a medium PE, (14) In a remote access, occupations of the medium bus and the remote bus are accomplished in the same cycle, (15) As mentioned in section 2, in a remote access, the *n-1* candidate medium buses are equally likely to be occupied Assumption (7) deals with bus conflicts, by requiring remote and local access requests are equally treated. Assumption (8) means that there is at least one inactive cycle between two active periods of every PE. Assumption (12) indicates that, when R hit occurs, the remote access need not occupy the remote bus to access the remote memory module. Instead, the requesting PE directly fetches data from the medium PE and the remote access becomes a local access. The remote bus speared under this circumstance can be allocated to another memory access (termed as Beneficiary accesses). For simplicity, it is assumed that R hit does not further occur to beneficiary accesses. According to assumption (13), since medium PE can perform its own memory operations as if it were not serving as a medium, it is unnecessary to record whether a PE is a medium or not in the states of the Markov model mentioned later.

Based on assumptions above, a discrete-time homogenous Markov model to calculate the throughput of *PB(n)* is designed. For *PB(n)*, a state of Markov chain is depicted by an n-tuple

$$S_i = [b_1, b_2 ..., b_n], 0 \le b_j \le n, 1 \le j \le n, 1 \le i \le T$$

Where Bus_j is identified as occupied by PE_{b_j} if $b_j>0$, or as idle if $b_j=0$. The Markov chain has T states, each of which represents a possible bus occupation configuration. An iterative procedure is developed to generate state space.

Let *P* denote the transition probability matrix and P_{ij} denote the transition probability from S_i to S_j. Let *Tr(i,j)* be the set of all possible transition configurations that can lead S_i to S_j. For a transition configuration $tr_k \in TR(i,j)$, let $\lambda_{1\text{-}9}$ denote probabilities that, l_1 new local accesses (not including beneficiary accesses) occur, l_2 new remote accesses (not including beneficiary accesses)occur, l_3 of the g_2 remote accesses get **R_hits**, l_4 beneficiary accesses occur, l_5 inactive PEs remain inactive, l_6 local accesses continue, l_7 remote accesses continue, l_8 local accesses terminate, l_9 remote accesses terminate, respectively. Then, $\lambda_{1\text{-}9}$ statistically equal g^{l_1}, g^{l_2}, $h^{l_3}(1-h)^{l_1+l_2}$, g^{l_4}, $(1-g)^{l_5}$, a^{l_6}, $a_2^{l_7}$, $(1-a)^{l_8}$ and $(1-a_2)^{l_9}$, respectively. On the other hand, since *1/n* denotes the probability that one access addresses to one of the *n* memory modules, $(1/n)^{l_1+l_2+l_4}$ denotes the probability that every one of the $l_1+l_2+l_4$ new accesses addresses to one of the *n* memory modules.

Therefore, the transition probability from S_i to S_j (i≠j) is expressed as:

$$p_{ij} = \begin{cases} 0 & if\ Tr(i,j) = \phi \\ \sum_{tr_k \in Tr(i,j)} (1/n)^{l_1+l_2+l_4} \times \prod_{1 \le z \le 9} l_z & else \end{cases} \quad (1)$$

P also satisfies

$$P_{ii} = 1 - \sum_{1 \le j \le T, i \ne j} p_{ij} \quad (2)$$

The Markov chain is aperiodic because every state can return to itself in one cycle ($p_{ii}>0$ for all 1⩽i⩽T); the chain is irreducible since any state is reachable from any other state with in a finite number of cycles. Consequently, the chain is ergodic and has a steady state probability distribution. Let π denote the steady state probability vector and π_i denote the steady state probability of S_i. Then π can be obtained by solving the system of equations

$$\pi = \pi \times P \quad (3)$$

Let op_i denote the number of accesses in state S_i. Then, the throughput of *PB(n)* can be expressed as

$$tp\_P = \sum_{1 \le i \le T} op_i \times \pi_i \quad (4)$$

On the other hand, this paper borrows the model designed in [10] to study the performance of single-bus *SB(n)* and multi-bus *MB(n,m,b)*, where all accesses are local. Since assumptions (1-10) in the previous model for partial-bus also apply to [10], results of the two models are comparable. In this model, the number of occupied buses is taken as the Markov chain state. The transition probability from a state with i occupied buses to another state with j occupied buses is Z_{ij}.

$$Z_{ij} = \sum_{k=\max(0,j-b+1)}^{\min(i,j)} p_{i,k} q_{b-i,j-k} \quad (5)$$

Where $p_{i,k}$ denotes the probability that k out of i occupied buses remain occupied

$$p_{i,k} = C_i^k a^k (1-a)^{i-k} \quad (6)$$

Let *f(i, r)* denote the probability that exactly *r* out of the *m-i* idle memory modules are requested, then

$$f(i,r) = C_{m-i}^r [1-(1-g/m)^{n-i}]^r \times (1-g/m)^{(n-i)\times(m-i-r)} \quad (7)$$

Consequently, the probability that *j-k* out of the *b-i* idle buses become occupied, $q_{b-i,j-k}$, can be obtained as

$$q_{b-i,j-k} = \begin{cases} f(i,j-k) & for\ j-k < b-i \\ \sum_{l=b-i}^{m-i} f(i,l) & for\ j-k = b-i \end{cases} \quad (8)$$

Through solving a function similar to Eq.(3), throughput of single-bus (where *b=m=1*) and multi-bus can be obtained.

5. COMPARISONS AND RESULTS

Based on models designed in previous sections, this section presents performance comparison between *PB(n)* and *SB(n)/MB(n,m,b)*.

Note that in every group of comparison, the numbers of memory-modules/buses of *MB(n,m,b)* are equal to the numbers of memory-modules/buses of $PB(n)$, to guarantee the comparison between partial-bus and multi-bus is fair.

Fig.4 to Fig.11 illustrate throughput as functions of a and g achieved by single-bus, multi-bus and our partial-bus, where *u=5*, *0.3≤a,g≤0.8*, *0≤h≤40%*. In conclusion, the higher *h* can be, the better performance partial-bus achieves. Partial-bus definitely achieves higher throughput than single-bus even if *h=0*. Partial-bus is also superior to multi-bus in that it achieves higher throughput in most circumstances even if *h=0*. Another interesting observation is that, the more PEs partial-bus has the higher throughput advantage it achieves over multi-bus.

6. SUMMARY AND FURTHER STUDY

This paper proposes a bus-based interconnection topology, partial-bus, for small-scale multiprocessor systems and SOC multiprocessor systems. Partial-bus topology features reduced port number on buses than traditional bus-based interconnections. A formal verification of liveness property of bus occupation behaviors is also given.

Although the we do not hastily conclude that partial-bus is absolutely superior to single/multiple bus, results in section 5 at least indicate that partial-bus may well be a promising alternative to traditional single/multiple bus from the performance perspective.

(1) Based on present work, some further topics are to be studied: Sensitivity analysis models to identify performance bottlenecks
(2) Quantitative analysis of energy consumption of partial-bus
(3) Designing models to analyze other performance aspects, such as scalability, dependability/availability, fault tolerance, resource utilization, etc
(4) Applying other optimization techniques into partial-bus, such as data pre-fetch, cache miss-rate reduction, pipelining, etc
(5) Quantitative analysis of cache-coherence-overhead of partial-bus

7. APPENDIX: PROOF OF THE LIVENESS PROPERTY

This section presents the formal verification of the liveness property mentioned section 3. The verification is based on a synthesized P/T net depicting all bus occupation behaviors in partial-bus. The net is synthesized from

(1) *n×(n-1)* subnets identical to Fig. 2. , each of which depicts a possible local access configuration (*n* PEs, each with *n-1* local memory modules)

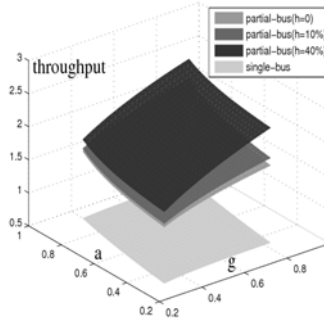


Fig. 4. PB (4) vs. SB (4)

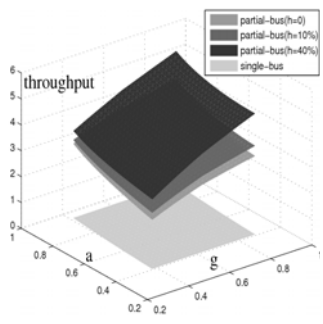


Fig. 5. PB (8) vs. SB (8)

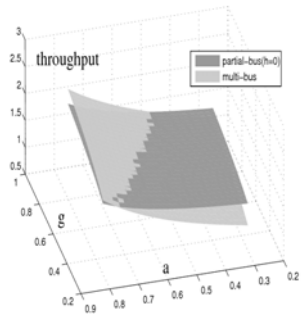


Fig. 6. PB (4) vs. MB (4, 4, 4) where h=0

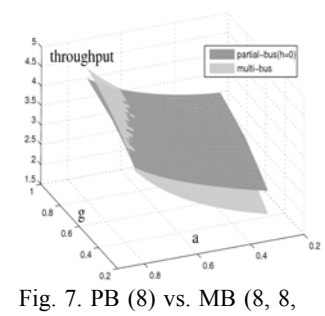


Fig. 7. PB (8) vs. MB (8, 8, 8) where h=0

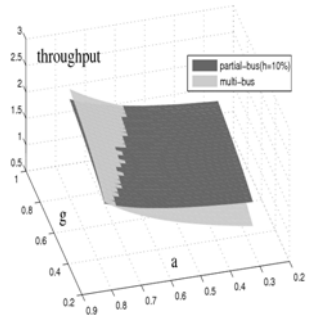


Fig. 8. PB (4) vs. MB (4, 4,4) where h=10%

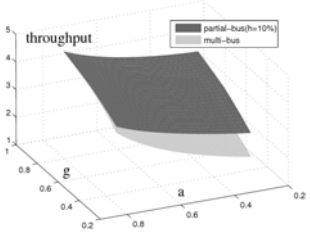


Fig. 9. PB (8) vs. MB (8, 8, 8) where h=10%

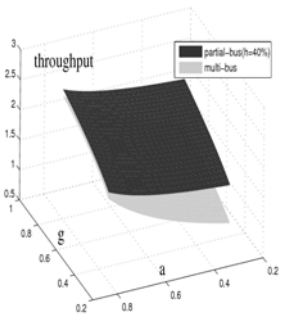


Fig. 10. PB (4) vs. MB (4, 4, 4) where h=40%

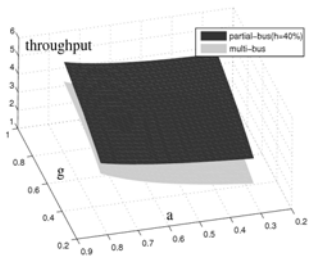


Fig. 11. PB (8) vs. MB (8, 8, 8) where h=40%

(2) $n\times(n-1)$ subnets identical to Fig. 3. , each of which depicts a possible remote access configuration (*n* PEs, each with 1 remote bus and 1 remote memory module). For a given PE wishing to access its remote memory module, there exist *n-1* candidate medium buses, as mentioned in section 2)

The synthesis rule is given as: (1) For every *i* where $1\le i\le n$, places $Inactive_i$ in all subnets are merged into a single place $Inactive_i$ in the synthesized net (the same is done for places $Wait_i$ and transitions Req_i), (2) For every *j* where $1\le i\le n$, places B_j in all subnets are merged into a single place B_j in the synthesized net. For instance, Fig. 12. illustrates how three subnets depicting three local accesses in a partial-bus system *PB(3)* (PE_1 occupying Bus_1 to access MM_1, PE_1 occupying Bus_3 to access MM_3 and PE_3 occupying Bus_3 to access MM_3) are synthesized.

In the synthesized net, conflicting transitions enabled by place B_j indicate memory requests competing for Bus_j. As mentioned earlier, the control policy randomly and non-prioritizedly allocates Bus_j to one of the competing requests.

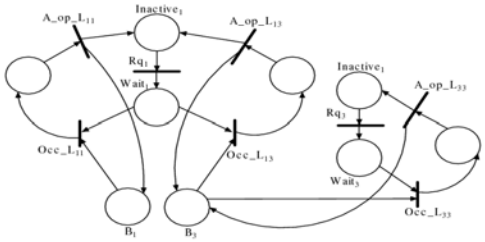


Fig. 12. Synthesis of three subnets

The liveness property is manually verified, because existing automatic verification tools only deal with net with fixed structure while the property to be verified is based on synthesized net with number *n* varied. The liveness property is formalized as

$$\forall M\in[M_0>,\forall t\in T,\exists M'\in[M>: M'[t>$$

Or as

$$\forall M\in[M_0>,\forall t\in T,\exists\sigma: M[\sigma>M'\wedge M'[t>$$

where σ is a firing sequence with a finite length, *T* is the set of all transitions in the synthesized net, and M_0 is the initial marking. The liveness property ensures that no deadlock occurs and every transition is eventually fired, i.e. all memory requests are eventually satisfied. Note that if the length of firing sequence is 0, transition *t* is enabled at marking *M* and *M=M'*. A firing sequence with length 0 is recorded as *Null*. If there exists a firing sequence σ satisfying property above, the liveness property is proved. It is assumed that initially all PEs are inactive and all buses are idle. Therefore, the following property follows

$$\forall i,j: 1\le i,j\le n, M_0(Inactive_i)=M_0(B_j)=1$$

The proof is done in three steps

Property1:

$$\forall M\in[M_0>,\forall i:$$
$$1\le i\le n, M(Inactive_i)+M(Wait_i)$$
$$+\sum_{All\ Ready\_L_{ij}} M(Ready\_L_{ij})+\sum_{All\ Ready\_R_{ijk}} M(Ready\_R_{ijk})=1$$

Proof Since initially there exists only 1 token in each $Inactive_i$ and the firing of every transition maintains the sum of tokens in all places of $Inactive_i$, $Wait_i$, $Ready\_L_{ij}$ and $Ready\_R_{ijk}$, the property holds for all reachable markings.

Property2:

$$\forall M\in[M_0>,$$
$$\forall t\in T\wedge t\notin\{A\_op\_L_{ij}, A\_op\_R_{ijk}, Rq_i\},$$
$$\exists\sigma: M[\sigma>M'\wedge M'[t>$$

Proof To prove this property, the firing sequence σ for each transition involved must be found:

For transition Occ_{ij}

If $M(B_j)=1$

If $M(Wait_i)=1$, then $\sigma=Null$

If $M(Inactive_i)=1$, then $\sigma=<Rq_i>$

If $M(B_j)=0$

If $M(Ready\_L_{ij})=1$, then $\sigma=<A\_op\_L_{ij}, Rq_i>$

If there is a *k* such that $M(Ready\_R_{ijk})=1$, then $\sigma=<A\_op\_R_{ijk}, Rq_i>$

If there is a *k* such that $M(Ready\_R_{ikj})=1$, then $\sigma=<A\_op\_R_{ikj}, Rq_i>$

If there is an *x* such that $x\neq i$ and $M(Ready\_L_{xj})=1$

If $M(Wait_i)=1$, then $\sigma=<A\_op\_L_{xj}>$

If $M(Inactive_i)=1$, then $\sigma=<Rq_i, A\_op\_L_{xj}>$

If there is an *x* and a *k* such that $x\neq k$ and $M(Ready\_R_{xjk})=1$

If $M(Wait_i)=1$, then $\sigma=<A\_op\_R_{xjk}>$

If $M(Inactive_i)=1$, then $\sigma=<Rq_i, A\_op\_R_{xjk}>$

If there is an *x* and a *k* such that $x\neq k$ and $M(Ready\_R_{xkj})=1$

If $M(Wait_i)=1$, then $\sigma=<A\_op\_R_{xkj}>$

If $M(Inactive_i)=1$, then $\sigma=<Rq_i, A\_op\_R_{xkj}>$

For transition Occ_B_{ijk}, the firing sequence σ is given by $\sigma=\sigma_1+\sigma_2+\sigma_3$, where '+' denotes the operation of sequence concatenation

If $M(B_j)=0$
  If $M(Ready\_L_{ij})=1$, then $\sigma_1=<A\_op\_L_{ij}>$
  If there is an x such that $x\neq i$ and $M(Ready\_L_{xj})=1$, then $\sigma_1=<A\_op\_L_{xj}>$
  If there are an x and a t such that $x\neq i$, $t\neq k$ and $M(Ready\_R_{xjt})=1$, then $\sigma_1=<A\_op\_R_{xjt}>$
  If there an x and a t such that $x\neq i$, $t\neq k$ and $M(Ready\_R_{xtj})=1$, then $\sigma_1=<A\_op\_R_{xtj}>$
  If there is an x such that $x\neq k$ and $M(Ready\_R_{ijx})=1$, then $\sigma_1=<A\_op\_R_{ijx}>$
If $M(B_k)=0$
  If $M(Ready\_L_{ik})=1$, then $\sigma_2=<A\_op\_L_{ik}>$
  If there is an x such that $x\neq i$ and $M(Ready\_L_{xk})=1$, then $\sigma_2=<A\_op\_L_{xk}>$
  If there are an x and a t such that $x\neq i$, $t\neq j$ and $M(Ready\_R_{xkt})=1$, then $\sigma_2=<A\_op\_R_{xkt}>$
  If there are an x and a t such that $x\neq i$, $t\neq j$ and $M(Ready\_R_{xtk})=1$, then $\sigma_2=<A\_op\_R_{xtk}>$
  If there is an x such that $x\neq j$ and $M(Ready\_R_{ixk})=1$, then $\sigma_2=<A\_op\_R_{ixk}>$
If $M(B_j)=M(B_k)=0$
  If $M(Ready\_R_{ijk})=1$, then $\sigma_1=<A\_op\_R_{ijk}>$, $\sigma_2=Null$
  If there is an x such that $x\neq i$ and $M(Ready\_R_{xjk})=1$, then $\sigma_1=<A\_op\_R_{xjk}>$, $\sigma_2=Null$
  If there is an x such that $x\neq i$ and $M(Ready\_R_{xkj})=1$, then $\sigma_1=<A\_op\_R_{xkj}>$, $\sigma_2=Null$
If $M(B_j)=1$ then $\sigma_1=Null$
If $M(B_k)=1$ then $\sigma_2=Null$
If $M(Inactive_i)=1$ then $\sigma_3=<Rq_i>$
If $M(Wait_i)=1$ then $\sigma_3=Null$

Property3

$$\forall M \in [M_0>, \forall t \in \{A\_op\_L_{ij}, A\_op\_R_{ijk}, Rq_i\},$$
$$\exists \sigma : M[\sigma > M' \wedge M'[t >$$

Proof The firing sequence σ for each transition involved can be found as

For transition $A\_op\_L_{ij}$

As proven before, there exists a sequence σ' such that $M[\sigma'>M'' \wedge M''[Occ_{ij}>$. Therefore, $\sigma=\sigma'+<Occ_{ij}>$

For transition $A\_op\_R_{ijk}$

As proven before, there exists a sequence σ' such that $M[\sigma'>M'' \wedge M''[Occ\_B_{ijk}>$. Therefore, $\sigma=\sigma'+<Occ\_B_{ijk}>$

For transition Rq_i

As proven before, there exists a sequence σ' such that $M[\sigma'>M'' \wedge M''[A\_op\_L_{ij}>$. Therefore, $\sigma=\sigma'+<A\_op\_L_{ij}>$

Combining **Property2** and **Property3**, it follows that there always exists a firing sequence σ for every transition in the synthesized net such that for any marking $M \in [M_0>$ the transition is enabled at marking M' where $M[\sigma>M'$. Therefore, the liveness property gets proven.

8. REFERENCES

[1] Information technology microprocessor systems-high-performance synchronous 32-bit bus: MULTIBUS II, IEEE standard, ANSI/IEEE Std 1296-1994, ISO/IEC 10861, 1994, 0-1.

[2] S. Erno, L. Vesa, et al, "Overview of bus-based system-on-chip interconnections", In Proc. IEEE International Symposium on Circuits and Systems, 2002, 372-375.

[3] A.S. Lee, N.W. Bergmann, "On-chip communication architectures for reconfigurable System-on-Chip", In Proc. IEEE International Conference on Field-Programmable Technology, 2003, 332-335.

[4] AMBA Specification Rev 2.0, ARM Ltd. www.arm.com

[5] IBM Corporation, The Coreconnect Bus Architecture, 2004, www.chips.ibm.com

[6] B. Cordan, "An effective bus architecture for system-on-chip design", In Proc. IEEE Integrated Circuit Conference, 1999, 623-626.

[7] T. Thin-Fong, W. Yamamoto, "Queuing simulation model for multiprocessor systems",In IEEE transaction on Computers, 2003,36(2) :58-64.

[8] S. Morrison, I.M. Obeng, et al, "Simulation Study of the Performance of a Cluster-based hypercube multicomputer", In Simulation-transactions of the society for Modeling and simulation international, 1997, 68(4).

[9] N.F. Tzeng, H. Kumar, "Traffic analysis and simulation performance of incomplete hypercubes", In IEEE Transactions on Parallel and Distributed Systems, 1996,7(7): 740-754.

[10] N. Hamee, "A Markov model for multibus multiprocessor system under asynchronous operation", In Information Processing Letters, 1995,54(1): 11-16.

[11] G. Rubino, B. Sericola, "Interval Availability Analysis Using Denumerable Markov Processes: Application to Multiprocessor Subject to Breakdowns and Repair", In IEEE Transaction on Computers, 1995,44(2):286-291.

[12] G.Y. Min, M. Ould-Khaoua, "Performance modeling and evaluation of k-array n-cube networks in the presence of bursty traffic", In Proc. 10th IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunications Systems MASCOTS ,2002, 383- 390.

[13] C. Girault, R. Valk, "Petri Nets for Systems Engineering: A Guide to Modeling, Verification, and Applications", Springer-Verlag, 2002.

Integrated Modeling and Simulation of Activated Sludge Process Based on SASM

Jianbo Chen, Dongwen Shi, Yangjun Gao, Danli Xi
College of Environmental Science and Engineering, Donghua University
Shanghai, 200051, China
Email: pop_qq@126.com
And
Zhanjun Zhang
Zhengzhou Water Supply Corp
Zhengzhou, Henan Province, 450013, China
Email: zhj001_zhang@126.com

ABSTRACT

Dynamic models and process simulation can be very useful to control strategy of wastewater treatment process. The principle of simplified activated sludge model (SASM) based on ASM1 model is presented here. This model is mainly applied to predict the performance of wastewater treatment. In addition the secondary settler model is also presented based on the solid flux theory and Takács settling velocity model. When coupling the bioreactor to the secondary settler, the activated sludge process integrated model was developed in MATLAB/Simulink and then the simulation system was achieved using VC.NET as development language. At last, SASM model was verified with the obtained information and directly compares the simulate results with the actual value collect from plant. The test results show that SASM model is a good alternative for on-line process monitoring, prediction, control and optimization and the proposed procedure appears to be a practical tool for efficient and reliable model simulation.

Keywords: simulation，activated sludge process, ASAM, secondary settler, MATLAB/Simulink.

1. INTRODUCTION

Dynamic models and process simulators can be very useful in creation of effective control system for wastewater treatment processes. Several simulators and models of activated sludge process have been developed during the last decade. Typical application areas for these kind simulators include process design, controller design, optimization, forecasting, and research. However, most simulators have not been used very much by plant operator. There are several reasons for this, but one major cause is that the simulators so far have been much complex and not been very user friendly. The objective of our research is to develop an integrated and user friendly simulator for activated sludge process, which is suited for operation and control. This paper first presents the basic modeling principle of simplified activated sludge. Then the bioreactor model and secondary settler model are put forward. At last the results of using the integrated model for simulating are given.

2. MATHEMATIC MODEL

2.1 Simplified Activated Sludge Model-SASM

Today the behavior of the biological reactors in a municipal wastewater treatment plant applying an activated sludge process can be described quite well by computer simulations based on the IAWQ model No.1 (ASM1) [1]. The model describes the processes of carbonaceous oxidation, nitrification and denitrification in a mechanistic way, and the dynamics of the organic and nitrogen components can be observed. The model contains 13 different components. The behavior of each component is described by a nonlinear differential equation. ASM1 is probably the most used advanced model for the activated sludge process. However, it also includes following disadvantages. First of all, the calibration of all the (kinetic) parameters is a hard task. Secondly, the model is highly non-linear due to the appearance of Monod like kinetics in the mass balance equations, which lead the model calibration and application to frequently change. So, the complexity of the model should be reduced or simplified with appropriate methods [2].

To facilitate model development and application, a simplified model is developed based on the understanding of ASM1 model and the activated sludge process knowledge. The following simplifications are made. First of all, the goal of the simplified model is to adequately represent the carbonization and nitrification processes, thus the denitrification process is neglected, and the simplified model is so named as SASM model. Secondly, because of the ammonification process being a fast reaction and the ammonification rate being likely proportional to the degradation rate of soluble substrates [3], the ammonification process is neglected and ammonia nitrogen is assumed to form by microorganisms decay in the SASM model. Thirdly, the hydrolysis of particulate organic nitrogen occurring along with the hydrolysis of slowly biodegradable substrate, so the hydrolysis process of particulate organic nitrogen is not considered and two organic nitrogen fractions are cancelled. Fourthly, the hydrolysis rate of slowly biodegradable substrate is controlled by their concentrations [4], so the hydrolysis process is modeled as first-order kinetics. Finally, inert particulate substrates that are produced by metabolism and come from the influent are incorporated one fraction in SASM model [2].

In eight basic processes represented in ASM1 model, only five remain in SASM model: growth as well as decay of heterotrophs and autotrophs and hydrolysis. In the thirteen components only ten remain: S_I (soluble inert substrate), S_S (soluble readily biodegradable substrate), X_I (particulate inert substrate), X_S (particulate slowly biodegradable

substrate), X_{BH} (heterotrophs), X_{BA} (autotrophs), S_O (soluble oxygen), S_{NH} (ammonia), S_{NO} (nitrate and nitrite) and S_{ALK} (alkalinity).

Table 1. The SASM model

Process	Component(Stoichiometric coefficient) $v_{i,j}$										Reaction Rate r_j
	1 S_I	2 S_S	3 X_I	4 X_S	5 X_{BH}	6 X_{BA}	7 S_O	8 S_{NO}	9 S_{NH}	10 S_{ALK}	
Heterotrophs growth		$-\frac{1}{Y_H}$			1		$-\frac{1-Y_H}{Y_H}$		$-i_{XB}$	$-\frac{i_{XB}}{14}$	$\mu_H(\frac{S_S}{K_S+S_S})X_{BH}$
Autotrophs growth						1	$-\frac{4.57-Y_A}{Y_A}$	$\frac{1}{Y_A}$	$-i_{XB}-\frac{1}{Y_A}$	$-\frac{i_{XB}}{14}-\frac{1}{7Y_A}$	$\mu_A(\frac{S_{NH}}{K_{NH}+S_{NH}})$ $(\frac{S_O}{K_{O.A}+S_O})X_{BA}$
Heterotrophs decay			f_p	$1-f_p$	-1				i_{XB}		$b_H X_{BH}$
Autotrophs decay			f_p	$1-f_p$		-1			i_{XB}		$b_A X_{BA}$
Hydrolysis		1		-1							$k_h X_S$

The detailed SASM model is summarized in Table 1. It can be seen from Table 1. that the number of components, processes, stoichiometric coefficients and kinetic parameters of SASM model are less than that of ASM1 model. As a result, the difficulties existing in parameters identification and model application are reduced, and the model practicability is also enhanced.

2.2 Aeration Tank

The system is the one represented in Fig.1. For the aeration tank we consider completely stirred tank reactor (CSTR) in steady state. The implemented process consists of a basin divided into several completely mixed compartments. The number of compartments is not fixed. Meanwhile, the parameters are possible to change in order to simulate.

The mass balances done inside the aeration tank resort to the SASM model. The generic equation for a mass balance around a certain system considering a CSTR is

$$QX_{Ai} + Vr_A = QX_{AO} + V\frac{dX_A}{dt} \tag{1}$$

In what follows we describe the equations resulting from these balances. Where Q is the flow that enters the tank, V is the aeration tank volume, r_A is the reaction (growth) rate, X_{AO} and X_{Ai} are the concentrations of the component inside the compartment and on entry, respectively. It is convenient to refer that in a CSTR the concentration of a compound is the same at any point inside the reactor and at the effluent of that reactor, thus X_{AO} equal to X_A.

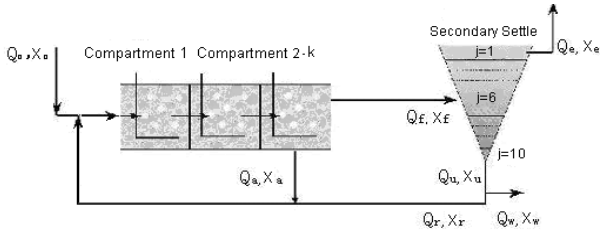


Fig. 1. Schematic representation of the activated sludge system

Thus for compartment 1

$$\frac{dX_1}{dt} = \frac{1}{V_1}(Q_a X_a + Q_r X_r + Q_o X_o + r_1 X_1 - Q_1 X_1) \tag{2}$$

$$Q_1 = Q_a + Q_r + Q_o \tag{3}$$

For compartment *2-k*

$$\frac{dX_k}{dt} = \frac{1}{V_k}(Q_{k-1}X_{k-1} + r_k V_k - Q_k X_k) \tag{4}$$

$$Q_k = Q_{k-1} \tag{5}$$

The reaction rate r_A, is obtained by the sum of the product of the stoichiometric coefficients, v_{ij}, with the expression of the process reaction rate, r_j, of the ASM1 Peterson matrix [1].

$$r_A = \sum_j V_{ij} r_j \tag{6}$$

The SASM model involves 5 processes incorporating different components. For example, the mass balance equation related to the soluble substrate (SS) is the following:

$$-\frac{\mu_H}{Y_H}\frac{S_S}{K_S+S_S}X_{BH} + k_h X_S + \frac{Q}{V}(S_{Sin} - S_S) = 0 \tag{7}$$

We denote all the soluble components by $S_?$ and the particulates by $X_?$. All these equations depend on stoichiometric and kinetic parameters.

2.3 Secondary Settler

When simulating the activated sludge process in a wastewater treatment plant it is also necessary to model a settler since biomass must be settled and recirculated back into the process. In a settler, particulate matter in the water come from the bioreactor sink to the bottom of the settler (thickening), and clear water is produced (clarification) and removed in the top of the settler. The sludge is recirculated to the bioreactor to maintain a desired solids level.

The settler unit is modeled as a traditional one-dimensional exponential model, in which the tank is divided into ten horizontal layers. The solids concentration in each layer of constant height can be calculated from a simple mass balance around each layer. No biological reactions take place in this tank. Fig.2 describes mass balance of each layer in detail.

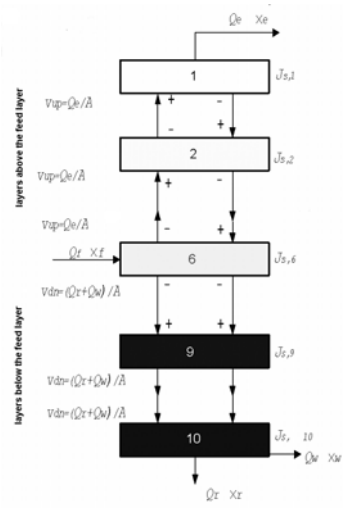


Fig. 2. Mass balance of each layer in secondary settler

Only vertical flux is considered and the solids are uniformly distributed across the entire cross-sectional area of the feed layer (j=6, in our case). This model is based on a traditional solids flux analysis but the flux in a particular layer is limited by the adjacent layer. The settling function, described by Takács et al. in [5], is given by

$$v_{s,j} = \max[0, \min\{v_0', (v_0 e^{-r_h(X_j - f_{ns}X_f)} - v_0 e^{-r_p(X_j - f_{ns}X_f)})\}] \quad (8)$$

Where $v_{s,j}$ is the settling velocity in layer j (m/day), X_j is the total suspended solids concentration in each of the ten considered layers of the settler and v_0, v_0', r_h, r_p and f_{ns} are settling parameters. Note that $X_6=X_f$. The area and height of each layer are A, Z_j respectively.

The solids flux due to the bulk movement of liquid may be up or down, v_{up} and v_{dn} respectively, depending on its position relative to the feed layer, thus

$$v_{up} = \frac{Q_e}{A},$$

$$v_{dn} = \frac{Q_u}{A} = \frac{Q_r + Q_w}{A} \quad (9)$$

As to the subscripts, r concerns the recycled sludge, w the wasted sludge and e the treated effluent.

The sedimentation flux, J_s, for the layers under the feed layer (j=6,…,10) is given by

$$J_{s,j} = v_{s,j} X_j \quad (10)$$

And above the feed layer (j=1,…,5) the sedimentation flux, J_{clar}, is given by

$$J_{clar,j} = \begin{cases} v_{s,j} X_j & if \ \ X_{j+1} \le X_t \\ \min(v_{s,j} X_j, v_{s,j+1} X_{j+1}) & otherwise \end{cases} \quad (11)$$

Where X_j is the threshold concentration of the sludge. The resulting solids balances around each layer, considering steady state, are the following:

For the top layer (j=1)

$$\frac{dX_1}{dt} = \frac{v_{up}(X_2 - X_1) - J_{clar,1}}{Z_1} \quad (12)$$

For the intermediate layers above the feed layer (j=2,…,5)

$$\frac{dX_j}{dt} = \frac{v_{up}(X_{j+1} - X_j) + J_{clar,j-1} - J_{clar,m}}{Z_j} \quad (13)$$

For the feed layer (j=6)

$$\frac{dX_6}{dt} = \frac{\frac{Q_f X_f}{A} + J_{clar,5} - (v_{up} + v_{dn}) X_6 - \min(J_{s,6}, J_{s,7})}{Z_6} \quad (14)$$

For the intermediate layers under the feed layer (j=7, 9)

$$\frac{dX_j}{dt} = \frac{v_{dn}(X_{j-1} - X_j) + \min(J_{s,j}, J_{s,j-1}) - \min(J_{s,j}, J_{s,j+1})}{Z_j} \quad (15)$$

And, for the bottom layer (j=10)

$$\frac{dX_{10}}{dt} = \frac{v_{dn}(X_9 - X_{10}) + \min(J_{s,9}, J_{s,10})}{Z_{10}} \quad (16)$$

For the soluble components (including dissolved oxygen) in the feed layer (j = 6)

$$\frac{dS_6}{dt} = \frac{\frac{Q_f S_f}{A} - (v_{dn} - v_{up}) S_6}{Z_6} \quad (17)$$

For the soluble components S_j (including dissolved oxygen) in the layers j = 1 to 5

$$\frac{dS_j}{dt} = \frac{v_{up}(S_{j+1} - S_j)}{Z_j} \quad (18)$$

For the soluble components S_j (including dissolved oxygen) in the layers j = 6 to 10

$$\frac{dS_j}{dt} = \frac{v_{dn}(S_{j-1} - S_j)}{Z_j} \quad (19)$$

2.4 Model of the Secondary Settler Coupled to the Biological Reactor

When coupling the secondary settler to the bioreactor, the conversion of components in two models must be taken into account. In SASM model, the unit of all organic material is kg (COD)/m^3. While in secondary settler model, the unit for describing the material is kg (SS)/m^3. The particulate material is affected by gravity settling and the true mass for all components must therefore be known. When entering the secondary settler, the transformation coefficient ($f_{rCOD\text{-}SS}$) and the Eq. (20) are used to describe the components conversions in two models.

$$X_f = \frac{1}{f_{rCOD-SS}}(X_S + X_I + X_{BH} + X_{BA}) \quad (20)$$

Where $f_{rCOD\text{-}SS}$ is 4/3, which is based on averages from measurement of many treatment plants [6].

In the effluent from the secondary settler, the components conversions of two models can be described in Eq.(21)

$$\left\{ \begin{array}{l} TSS_e = 0.75 * (X_{S,e} + X_{BH,e} + X_{BA,e} + X_{I,e}) \\ COD_e = S_{S,e} + S_{I,e} + X_{S,e} + X_{I,e} + X_{BH,e} + X_{BA,e} \\ NH_4\text{-}N,_e = S_{NH,e} \\ NO_X\text{-}N,_e = S_{NO,e} \end{array} \right\} \quad (21)$$

3. APPLICATION of ACTIVATED SLUDGE PROCESS INTEGRATED MODEL

3.1 Custom Programming Software—WWTPSS

Based on the model developed above, simulation software

of integrated wastewater system based on SASM was developed using MATLAB language combined with VC.NET. The main simulation program was accomplished in MATLAB/Simulink environment and then compiled to dynamic link library (DLL) subprograms (Fig. 3.). To call the DLL developed with C++ in MATLAB platform can save a great deal of time in simulation. We called this custom GUIs Waste Water Treatment Plant Simulation System(WWTPSS). The main interface is shown in Fig. 4.

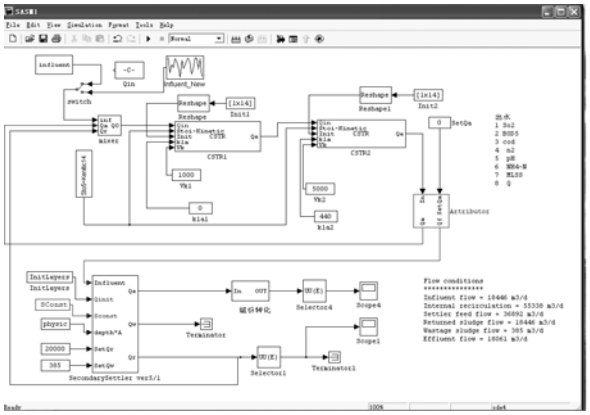


Fig. 3. Realization of mathematical model in Matlab/Simulink environment

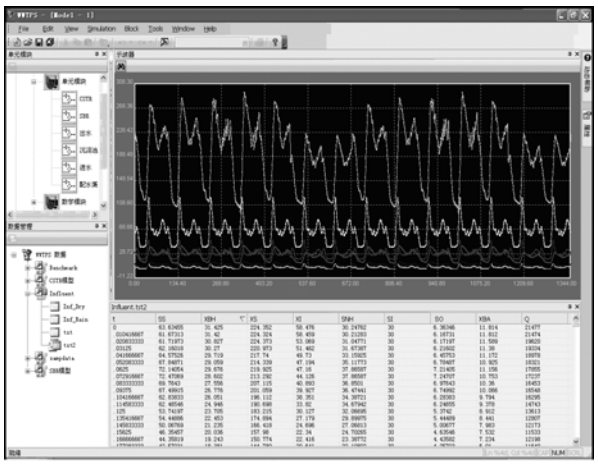


Fig. 4. The main interface of WWTPS

3.2 Simulation Results and Discussion

In order to verify the accuracy of SASM model, a practical wastewater treatment plan is applied to simulate to investigate whether the model explains the dynamical process in the actual phenomena.

Jingdong WWTP, constructed in 1980, is designed as a conventional activated sludge plant for a population equivalent of 200,000 inhabitants, 55000m^3/d total average flow. The influent of the WWTP is about 95 percent domestic wastewater. There are two aerobic reactors, which volume is 6480m^3. There are four secondary settlers, each diameter is 30 meter and the surface load is 1m^3/ (m^2.h).
According to SASM model, water components are divided into the following fractions: S_I, S_S, X_I, X_S, X_{BH}, X_{BA}, S_O, S_{NO}, S_{NH} and S_{ALK}. First of all, it is assumed that the oxygen concentration (S_O) in the influent is zero and the conversion of alkalinity (S_{ALK}) is not considered to be limited in this study. In this way, SASM wastewater components to consider are related to the organic carbon (*COD*) and nitrogen components.

Using respirometric tests [7], physico-chemical tests and the mass balances, the organic carbon components were determined. The design of these experiments was described in more details elsewhere [2]. For the nitrogen fractions the routine approach is used for influent characterization. It is assumed that the influent contained negligible concentrations of nitrate (S_{NO}). The measured *NH_4-N* concentration is considered to be equal to SNH. Using the methods mentioned above, the influent characterization of Jindong WWTP is determined, and the percentage of model components' concentrations to total *COD* in the influent is also gained. So, the available data from the routine measurements are translated into a data set that can be used as input to SASM model.

The dynamic simulation is made using the operation data. The changing trend of *COD*, *SS*, *NH4-N* and *NO_X-N* concentrations in the final effluent is gained. Fig. 5. illustrates the simulation results of SASM in June 2005.

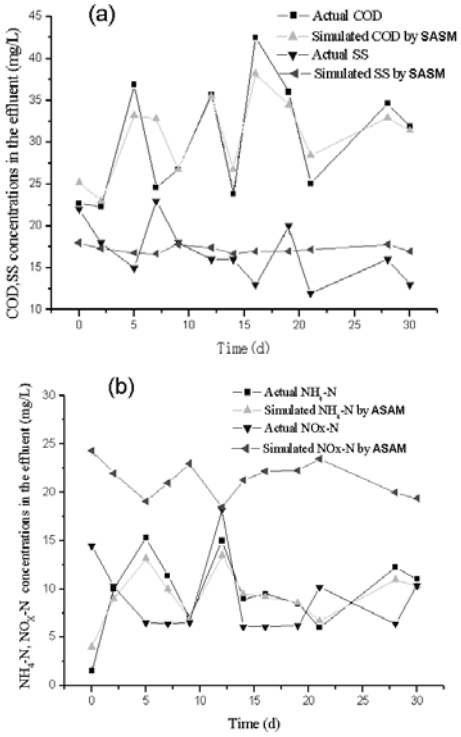


Fig. 5. The concentrations of each index in the effluent in June 2005 (a) Concentrations of *COD*, *SS*; (b) Concentrations of *NH_4-N*, *NO_X-N*

The simulation results are only correlative with the influent conditions when the model parameters are fixed. The changing trend of the predicted *COD*, *SS* and *NH_4-N* concentrations in the effluent is similar to the changing trend in the influent; however, the trend of actual effluent concentrations is somewhat different from that of the influent because of environmental influences. The results show that very good correlations between measured data and predicted data are achieved. In general, the predicted results of *COD*, *SS* and *NH_4-N* are well and the relative errors don't exceed 10 percent. On the contrary, two models predict too high *NOx-N* concentrations in the final effluent and the predicted relative errors all exceed 100 percent. This may

probably attribute to the simultaneous denitrification happening. Although the activated sludge system is fully aerated it is likely that some simultaneous denitrification can take place in the system, however this phenomena has not been described in model, so the actual *NOx-N* concentrations are higher than the predicted concentrations in the effluent. Anyway, judging from the predicted results, SASM model offer satisfied predictions.

4. CONCLUSION

Activated sludge process is the main technology for modern wastewater treatment. The principle and development of integrated dynamic simulation system for activated sludge process was finished here. This simulation system is mainly applied to predict the performance of wastewater treatment based on the theory of simplified sludge model and one-dimension settling model of circular secondary settler. The application example shows that this simulation system is a valuable tool for the plant operator or designer in forecasting or explaining the performance of wastewater treatment.

5. REFERENCES

[1] M. Henze, C.P.L.J. Grady et al, Activated sludge model no.1, IAWPRC Scientific and Technical Report No.1, London: International Association on Water Pollution Research and Contro, 1987.

[2] F. Liu, “Parameters Measurement and Simulation Research on Activated Sludge Model of Wastewater Treatment Plants”, PhD thesis, Department of Environmental Science and Technology, University of Tongji, China, 2004.

[3] P.L. Dold, G.V.R Marais, “Evaluation of the general activated sludge model proposed by the IAWPRC task group”, Water Science and Technology, Vol.18(6), 1986, pp.63-89.

[4] U. Sollfrank, W. Gujer, “Characterization of domestic wastewater for mathematical modelling of the activated sludge process”, Water Science and Technology, Vol. 23(4/6), 1991, pp.1057-1066.

[5] I. Takács, G. Patry et al, “A dynamic model of the clarification thickening process”, Water Science and Technology, Vol.25(1), 1991, pp.1263.

[6] M.Henze, W. Gujer et al, Activated Sludge Models ASM1, ASM2, ASM2D and ASM3. Scientific and Technical Report No. 9, London: IWA Publishing, 2000.

[7] H. Spanjers, P.A. Vanrolleghem, “Respriometry as a tool for rapid characterization of wastewater and activated sludge”, Water Science and Technology, Vol.31, 1995, pp.105-144.

The Research of Sea-floor Transient Electromagnetic Data Sample Process And the Field Experiments

HUI LI [1,2], JUN LIN[1], YANJU JI [1], FENGDAO ZHOU[1], YAN WANG [1], CHANGSHENG LIU[1], GUOHUA ZHOU,

(1.Key Lab of Geo-Information Exploration & Instrumentation，Jilin University , Changchun ,Jilin 130026, China
2. School of Electrical &Electronic Engineering, Changchun University of Technology,Changchun, Jilin 130012,China)
Email: lucylihui@126.com

ABSTRACT

The characteristics of the transient electromagnetic detection signal received by receiver instrument are analyzed in order to design the sampling methods. The amplitude of the induced electromotive force is high and the attenuation speed is fast in the early time. The signal is weak and the attenuation speed is slow in the late time. So three-point exponential approximation nonlinear smooth filter and approximative equal logarithm interval trace are adopted to display the dynamic profile map in time. In this way, the wrong data of the signal caused by ocean wave and factors else can be eliminated. The coplanar magnetic dipole-dipole configuration is adopted to accomplish the sea-floor detection. The canal of oil in the sea floor is detected by rule and line in this way.

Keywords: Sea-floor transient electromagnetic method, Induced electromotive force (EMF), Profile map, Dynamic trace, Signal-to Noise.

1. INTRODUCTION

Over three-fifths of the Earth's surface is covered by the oceans and shallow seas. There is a largely unexplored and unexploited resource base [1]. But little research was shown in the sea-floor environment in the method. The method of sound wave and seism is effective in sea-floor detection in principle. The corrosive layer is deposited in the sea floor. The air bladder is formed between the sea floor and seawater. The waves of the sound and seism are embarrassed. But the electromagnetic wave can penetrate. The electrical conductivity of the sea floor is usually much less than that of the seawater about it. The spread speed of electromagnetic wave in low electrical conductivity is faster than that in high electrical conductivity. Though the electromagnetic wave through sea floor arrives the receive point first. This is the theory base of electromagnetic detection in sea floor. A variety of time-domain and frequency-domain electromagnetic (EM) methods have come into use in minerals exploration for detection of conductive ore bodies in land. [2] To explore on the shallow seafloor, the time-domain electromagnetic methods (transient electromagnetic method or TEM) is preferred in this paper.

2. THE CHARACTERISTICS OF THE TRANSIENT ELECTROMAGNETIC SIGNAL [3,4]

The working process of the transient electromagnetic method is: send pulse earthwards to make a whirlpool in the underground medium, meanwhile send loop produce a electromagnetic field (called the first-time field). With the first-time field eliminating, electromagnetic field (the second-time field) produced by whirlpool doesn't eliminate at once, and it can be reacted by the receive loop to make sure the underground medium conductivity by calculation in order to understand the underground structure.[3]

The typical maps to explain the underground structure include the induced EMF curve, the profile map (the trace curve), and the section map.

The characteristics of transient electromagnetic signal are analyzed first in order to design the profile map.

(1) The response of limited dimensions conductor in the late time attenuates according to exponential rule. The induce electromotive force is

$$\varepsilon(t) = \frac{K}{\tau} e^{-t/\tau} \tag{1}$$

K is the constant, τ is the time constant related to conductor. The experimental attenuation curve under the logarithmic coordinate is shown in Fig. 1..

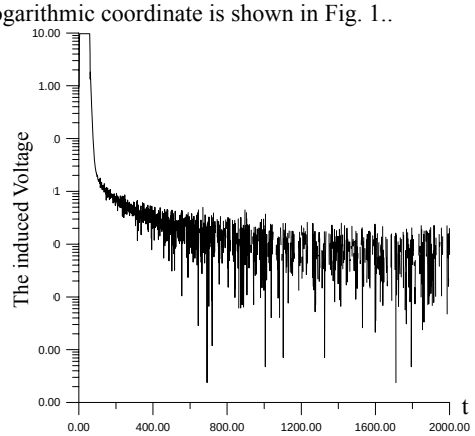


Fig. 1. Induced electromotive force curve

(2) The dynamic range of the induce electromotive force is wide. The amplitude of the signal is from 0.nμV to n×10^{5}μV to the one signal from early time to late time.

(3) Frequency band of the signal is wide. The frequency is from nHz to n×10^{4}Hz.

(4) The attenuation speed is different in different time (early time, middle time and late time). The amplitude is high and the attenuation speed is fast in the early time. The signal is weak and the attenuation speed is slow in the late time.

(5) The signal attenuates to microvolt in the late time. The signal-to noise is very low. The useful signal is difficult to distinguish.

According as the analysis above, the trace principle is confirmed. In the early time, the sampling interval and width of the windows must be narrow in order to distinguish the attenuation characters of the signal. In the late time, the sampling interval and width of the windows

can be wide because the attenuation speed and signal-to noise is very low.

3. THE TRACE METHOD OF THE INDUCED ELECTROMOTIVE FORCE CURVE

3.1 The three-point exponential approximation nonlinear smooth filter

When the system is dragged in the seawater, the ocean wave will cause the great noise. Thought the receiver system adopts the eliminating strange data and the medium value filter to remove the noise, the wrong data may be received and displayed in the attenuation curve. So the signal need be filtered in the software of data process. There are many filter methods in TEM, for example, three-point filter, four-point filter, six-point filter, Kalman filter and so on. And the response of limited dimensions conductor in the late time attenuates according to exponential rule. So the three-point exponential approximation nonlinear smooth filter is adopted in experiments. Its filter effect is better than three-point filter. The strange data is eliminated.

The transient signal received has the characteristic of unilateralism attenuation. If $\varepsilon_{t_{i+1}} > \varepsilon_{t_i}$, then $\varepsilon_{t_{i+1}}$ is the wrong value caused by the noise. The point Q ($\varepsilon_{t_{i+1}}$) is shown in Fig. 2. is the strange point. The filter method is drawing the tangent AB of the two preceding points (ε_{t_i} and $\varepsilon_{t_{i-1}}$) first, than drawing the horizontal line AC, the point P ($\varepsilon'_{t_{i+1}}$) of the angle midline of the angle CAB is the filter value.

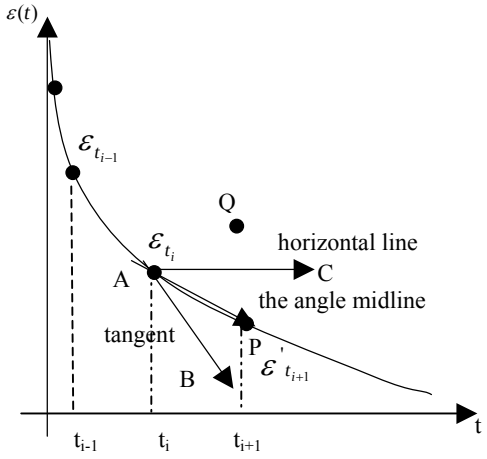


Fig. 2. The three-point exponential approximation nonlinear smooth filter

If $\varepsilon_{t_{i+1}} < \varepsilon_{t_i}$, then

$$\varepsilon'_{i+1}(t) = (\ln \varepsilon_i(t) + 2\ln \varepsilon_{i+1}(t) + \ln \varepsilon_{i+2}(t))/4 \quad (2)$$

3.2 Sample the signal at the approximative equal logarithm interval

The approximative equal logarithm interval is adopted to sample the transient signal in sea-floor detection. The width of window is 1.2 times that approaches the logarithm interval $10^{0.1}$(1.26 times) and the calculational error will be diminished. The center time and the induced electromotive force corresponding to this window are calculated in the method of geometrical average. The geometrical average is the area integral under the double logarithm coordinate. The calculational expression is shown in Eq. (3).

$$LogX = \frac{1}{N}\sum_{i=1}^{N}\log X_i \quad (3)$$

If the trace starting point is adopted the point 8, the first trace adopts the point 8 and 9 (because $8\times1.2=9.6$) to calculate the center time. If the sample ratio is 10μs of the receiving system , the center time of this windows (the first trace) is

$$\log t = \frac{1}{2}(\log 8\times 10 + \log 9\times 10) \quad (4)$$

Then

$$t = 0.0849\text{ms} \quad (5)$$

The induced electromotive force corresponding to this center time is calculated in this way too.

So the starting point of the second trace is the point 9. Then time values of the point 9, 10, 11 (because 9*1.2=10.8) are adopted to calculate the second trace and the rest may be deduced by analogy.

If the transmitting frequency is 25Hz, one cycle is 40ms; the induced electromotive force curve is 20ms. If the sample ratio is 10 μs of the receiving system, 2000 points is received by receiving system to describe a detection point. In the approximative equal logarithm interval to sample, 26 traces will be displayed. In this way, it can be distinguished the place of low resistance body clearly.

4. THE DESIGN OF DYNAMIC PROFILE MAP AND CONTROLLING SOFTWARE

In order to detect the sea floor, the instrument system must be dragged in the seawater. Under the environment of abominable marine, a large number of manpower, material resources and financial resources are cost when field experiments of the sea floor are carried. When the goal area measured very little, memory of data handling capacity is not very heavy. When the area measured is very large, or the position of the sea-floor goal body is unknown in the large scope, the exceptional position need be scan point by point, the quantities of the data are very large in storing amount. The result of detection is hoped to be displayed in real time in order to find the goal body in time, save the data memory space, and improve working efficiency. So the dynamic display software is needed, it can display the profile map real time [5].

Ocean offers the good environment to drag the system. The receiver system adopts PC/104-HXL/P300 with network card. The control computer on the boat finishes to setup parameter and display the profile map in real time. The network communication is adopted between the controlling computer on the boat and apparatus system in the sea floor. The speed of transmission achieves 1M/s.

The dynamic software can send the work parameters and control the transmitting and receiver system in the sea floor, for example, the points of collection and zoom out, periodicities, sampling rate, the name of the line of detected, and so on. In the same time, the controlling computer can receive the data detected and display the profile map in real. The interface of the soft is shown in Fig. 3..

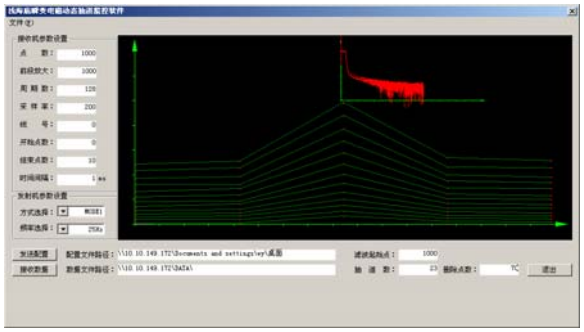


Fig. 3. The software of the dynamic control and profile map

5. THE EXPERIMENT ON THE SEA

The possible transmitter and receiver configurations for a sea-floor EM system are documented (S.J.Cheesman, R.N.Edwards and A.D.Chave, 1987). There are the coaxial electric dipole-dipole system (ERER), the coplanar magnetic dipole-dipole system (HZHZ), the coaxial magnetic dipole-dipole system (HRHR), the vertical magnetic dipole transmitter and the horizontal magnetic dipole receiver system (HZHR), the horizontal electric dipole transmitter and the vertical magnetic dipole receiver system (EPHIHZ). The configurations are shown in Fig. 4.[6]

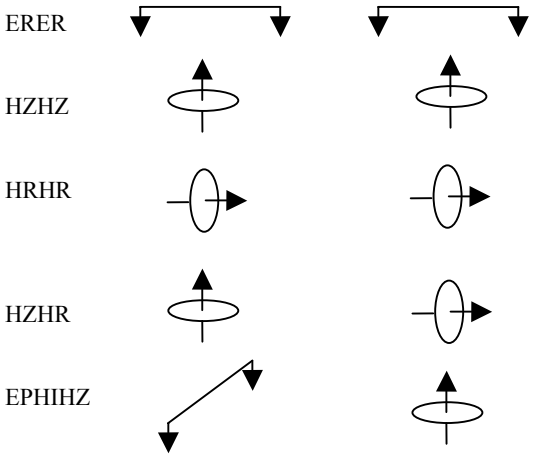


Fig. 4. The possible configurations for a sea-floor EM system

The coplanar magnetic dipole-dipole (HZHZ) is adopted in the sea-floor explorations in HuLu island of ShenYang in China. The transmitting frequency is 50 Hz. The magnetic sensor is adopted as the receiver loop; its availability acreage is 1750m^2. The profile map is shown in fig. 5.. Form the map; the point of 120-130 is the body of low resistance in the sea floor. The result accords with the fact. The method analyzed above is effective.

6. CONCLUSIONS

The method that the controlled source time-domain transient electromagnetic applies to sea-floor exploration is innovative. In order to distinguish the place of the low

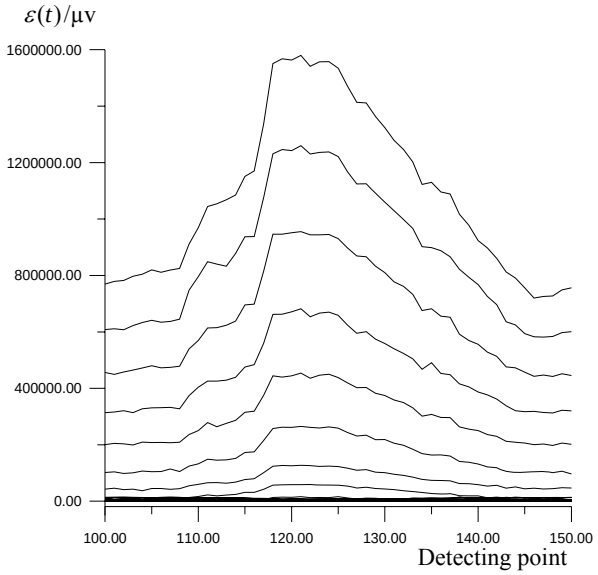


Fig. 5. The profile of the experiment in the sea floor

Resistance, the profile map should be display based on the data received. The coplanar magnetic dipole-dipole system (HZHZ) is used in the field experiments. Detailed conclusions are shown below.

(1) Analyzing the characteristic of the received signal. The amplitude is high and the attenuation speed is fast in the early time. The signal is weak and the attenuation speed is slow in the late time. So in the early time, the sampling interval and width of the windows must be narrow in order to distinguish the attenuation characters of the signal. In the late time, the sampling interval and width of the windows can be wide because the attenuation speed and Signal-to Noise is very low.

(2) The three-point exponential nonlinear smooth filter and the sampling of the approximative logarithm interval is adopted in the design of dynamic control and profile displaying soft.

(3) The result of experiment in the sea floor is shown in the paper. The canal of oil is detected by rule and line .The method analyzed above is effective.

7. ACKNOWLEDGMENTS

This paper is supported by Hi-Tech Research and Development Program of China (863 Program of China) (ID: 2002AA615030) and Science & Technology Plan of Changchun City (ID: 05GG38).

8. REFERENCES

[1] Alexander A. Kaufman, “The Influence of Currents Induced in the Host Rock on Electromagnetic Response of a Spheroid Directly Beneath a Loop”, Geophysics, August 1981, 46(8). 1121-1136.

[2] S.J. Cheesman, R.N. Edwards, et al, On the Theory of Sea-Floor Conductivity Mapping Using Transient Electromagnetic Systems, Geophysics, February 1987,52(2). 204-217.

[3] Y.J. Ji, “Sampling and Processing Techniques of Digitized Signals in Transient Electromagnetic Method”, Geophysical & Geochemical Exploration, April 1987, 27(2). 142-145.

[4] Y.J.Ji, "All-time Secondary Electromagnetic Field Extraction in High Resolution Transient Electromagnetic System for Subsurface Imaging", Dissertation of Jilin University, Changchun, 2004.

[5] H.Li, "Research of Loop Character and Transition Process of Sea-floor Transient Electromagnetic Method and experiment in Lab", 2005 International Conference on Machine Learning and Cybernetics, Guangzhou, China, August 2005,1589-1595.

[6] S.J. Cheesman, R.N. Edwards et al, "On the Theory of Sea-Floor Conductivity Mapping Using Transient Electromagnetic Systems", Geophysics, February 1987, 52(2). 204-217.

Robustness of Workpiece Process Order

Zihou Yuan [1] [2]
[1]School of Civil Engineering, Wuhan University, Wuhan , Hubei Province, 430072, China
[2]Dept. of Mathematics and Physics, Wuhan University of Science and Engineering, Wuhan, Hubei, 430073, China
Email:whuyzh@163.com
Mingxiang Chen
School of Civil Engineering, Wuhan University, Wuhan , Hubei Province, 430072, China
Email:whuyzh@163.com
Yinquan Han
School of Civil Engineering, Wuhan University, Wuhan , Hubei Province, 430072, China
Email:hanyinquan@163.com

ABSTRACT

For the purpose of investigating anti-jamming ability of attempering system we carried through research to workpiece process order in attempering system with induction method. We gained exported expression in stand-alone system by finite recursion, gained and proved robustness criterion of workpiece machining order. This conclusion has a certain contribution to attempering production.

Keywords: Attempering system, Workpiece process order, Induction, Robustness, Condition.

1. INTRODUCTION

Robustness of controlling system is a property that system keeps work performance in gear when there is uncertain factor in system. In this paper, we carried through research to workpiece process order in attempering system with induction method and educe distinguishing condition of workpiece process order robustness.

2. EXPORTED EXPRESSION IN STAND-ALONE SYSTEM n/1/p/C_{max} [ω, a, b]

Let's consider stand-alone system in follow Fig. 1.:

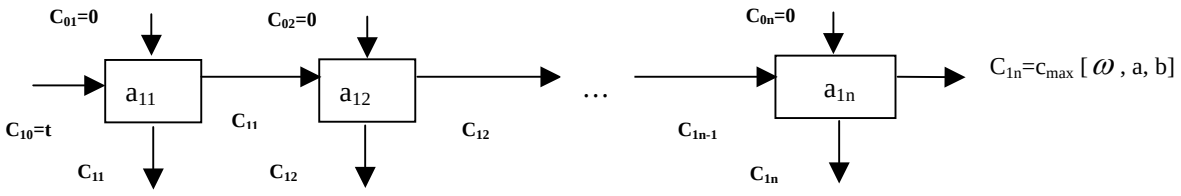


Fig. 1. Stand-alone system

ω =(1,2,…,n) is machining order of workpieces J_1, J_2, ……, Jn in machine M_1.
According to expression (1) of c_{max},

We can receive by finite recursion and trends layout equation group:

$$\text{Cmax}[\omega, a, b] = c_{1n} =$$
$$t + (a+bt)(a_{11} \otimes_b a_{12} \otimes_b \ldots \otimes_b a_{1n})$$
$$= t + (a+bt) \sum_{(1,j)\in L(\omega,a,b)} \otimes_b a_{1j} \quad \ldots\ldots \quad (1)$$

$L[\omega, a, b] = L[\omega]$:

(1, 1) ⟶ (1, 2) ⟶ …… ⟶ (1, n)

is capable line of system $n/1/p/c_{max}[\omega, a, b]$, $\sum_{(1,j)\in L(\omega)} \otimes_b a_{1j}$ is $\otimes_b$-addition.

Rewrite: $\sum_{(1,j)\in L(\omega)} \otimes_b a_{1j}$ is $l[a_{1,1}, a_{1,2}, \ldots, a_{1,n}]$,

Rewrite:

$$t + (a+bt) \sum_{(1,j)\in L(\omega)} \otimes_b a_{1j} \text{ is}$$

$l[t; a_{1,1}, a_{1,2}, \ldots, a_{1,n}]$, then

$l[t; a_{1,1}, a_{1,2}, \ldots, a_{1,n}] = c_{max}[\omega, a, b]$,

$l[0; a_{1,1}, a_{1,2}, \ldots, a_{1,n}] = a\, l[a_{1,1}, a_{1,2}, \ldots, a_{1,n}]$。

$l[t; a_{1,1}, a_{1,2}, \ldots, a_{1,n}]$ is unminus to parameter

a,b,t,$a_{1,j}$（j=1,2,…,n）,and we can get next character:

Character function $l[t;a_{1,1},a_{1,2},...,a_{1,n}]$ is humdrum.

Supposing {{ $i',...,i''$ }, { $j',...,j''$ }} is any plotting to aggregate

$\omega-\{i,j\}=\omega'-\{i,j\}$, if

$\forall a>0,\forall b\ge 0, c_{\max}[\omega,a,b]\le c_{\max}[\omega',a,b]$,

we think workpiece J_j can't process before workpiece J_i, notes $J_i \le J_j$.

3. ROBUSTNESS CRITERION OF WORKPIECE MACHINING ORDER

Foretheorem: condition of $J_i \le J_j$

$l^*(0;a_{1,i},a_{2,j})\le l^*(0;a_{1,j},a_{2,i}))$ Or

$l^*(0;a_{1,j}\oplus a_{2,i},a_{1,i},a_{2,j})\le$

$l^*(0;a_{1,i}\oplus a_{2,j},a_{1,j},a_{2,i}))$ Of $J_i \le J_j$ is equivalence with

$\min\{a_{1,i},a_{2,j}\}\le\min\{a_{1,j},a_{2,i}\}$. [1]

Setting up proposition

$\underset{1\le u\le m}{\Lambda}(l^*(0;a_{u,i},.....,a_{v,j})\le$

$l^*(0;a_{u,j},.....,a_{v,i}))$ … (2)

Theorem 1 (2) is robustness distinguishing condition of $J_i \le J_j$.

First, translate proposition (2) into keeping order S machine condition group of machine order (M_1, M2…Mm).

$P_m[s]: l^*[0;a_{v1,i},...,a_{vs,j}]\le l^*[0;a_{v1,j},...,a_{vs,i}]$,

$v_j=1+v_{j-1}$, $s=2,3...m$

Second, define Yue-Han keeping order: machine condition group

$Q_m[2]$: $\min\{a_{u,i},a_{v,j}\}\le\min\{a_{u,j},a_{v,i}\}$,

$(u,v)\in C_m^{(2)}[1,2,...,m]=$ {(1,2),(1,3), ……,(1,m), (2, 3)… (m-1,m)}

We should prove

$\forall s\in$ {2, 3… m}, $P_m[s]\Leftrightarrow Q_m[2]$.

We have proved theorem2 in [1]:

Theorem2 $P_2[s]\Leftrightarrow Q_2[2]$,

Also $l^*[0;a_{1,i},...,a_{2,j}]\le l^*[0;a_{1,j},...,a_{2,i}]\Leftrightarrow$

$\min\{a_{1,i},a_{2,j}\}\le\min\{a_{1,j},a_{2,i}\}$

Theorem 3 $P_m[s]\Leftarrow Q_m[2]$, s=2, 3… m

prove: universality supposing to take （$v_1,v_2,…,v_s$）= (1, 2… s)

So merely need to prove

$P_m[s]\Leftarrow Q_m[2]$

$l^*[0;a_{1,i},...,a_{s,j}]\le l^*[0;a_{1,j},...,a_{s,i}]$,

s=2, 3… m

by mathematics induction, when s=2, $C_m^{(2)}[1,2,...,m]$

Of $Q_m[2]$ includes the whole 2-machine condition group of $P_m[2]$,to base Theorem 1, we find

$P_m[2]\Leftarrow Q_m[2]$.

Supposing when s<m, $P_m[s]\Leftarrow Q_m[2]$,

s=2, 3… m－1

Wish to prove when s=m, $P_m[s]\Leftarrow Q_m[2]$,

s=2, 3… m－1, m ……(3)

Proof is disparted two steps.

Step 1, supposition $a_{m,j}\le a_{1,i}$, by proposition (2) $\min\{a_{1,i},a_{m,j}\}\le\min\{a_{1,j},a_{m,i}\}$,we receive

$a_{m,j}\le a_{1,i}$, $a_{mj}\le a_{m,i}$, $a_{m,j}\le a_{1,j}$.

Combine $P_m[m-1]$ in (3) with character, we gain:

$l[0;\frac{1}{a}l^*[0;a_{1,i},a_{2,i},...,a_{m-1,j}],a_{m,j}]\le$

$l[0;\frac{1}{a}l^*[0;a_{1,j},...,a_{m-1,i}],a_{m,i}]$ …（4）

We gain by $a_{m,j}\le a_{1,j}$ and character:

$l[0;a_{1,i},a_{2,i},...,a_{m-1,i},a_{m,i},a_{m,j}]\le$

$l[0;a_{1,i},...,a_{m-1,i},a_{m,i},a_{1,j}]$ …（5）

As $\otimes_b-$ addition in key way $\ge$any $\otimes_b-$ addition covered by key road,

$l[0;a_{1,j},a_{1,i},a_{2,i},...,a_{m,i}]\le$

$l[0;\frac{1}{a}l^*[0;a_{1,j},...,a_{m-1,i}],a_{m,i}]$ …（6）

according to （5），（6）,we receive

$l[0;a_{1,i},a_{2,i},...,a_{m,i},a_{m,j}]\le$

$l[0;\frac{1}{a}l^*[0;a_{1,j},...,a_{m-1,i}],a_{m,i}]$ …（7）

Write out its decompound formula:

$l^*[0;a_{1,j},a_{1,i},a_{2,i},...,a_{m,i}]=$

$l[0;\frac{1}{a}l^*[0;a_{1,i},...,a_{m-1,j}],a_{m,j}]\oplus$

$l[0;a_{1,i},a_{2,i},...,a_{m,i},a_{m,j}]$

$\underset{=}{\Delta}A\oplus B$ …（8）

$l^*[0;a_{1,j},a_{2,j},...,a_{m,j}]=$

$l[0;\frac{1}{a}l^*[0;a_{1,j},...,a_{m-1,i}],a_{m,j}]\oplus$

$l[0;a_{1,j},a_{2,j},...,a_{m,j},a_{m,i}]$

This work was supported by the Emphases Project of Hubei Education Dept (No: D200517004), the Emphases Laboratory of Hubei Province weave new material and appliance (No: TKL200401) and the Project of Hubei Education Dept (No: 2004X074) to Mr. Yuan.

$\Delta C \oplus D$ …（9）

we gain A≤C from （4）(8)(9),we gain B≤C from （6）（8）(9),then

$A \oplus B \le C \oplus C \le C \oplus D$

so we gain $P_m[s] \Leftarrow Q_m[2]$ ， s=2,3,…,m－1,m ,under $a_{m,j} \le a_{1,i}$

step 2, supposing $a_{m,j} > a_{1,i}$,we can gain $P_m[s] \Leftarrow Q_m[2]$ ，s=2,3,…,m－1,m, similarly the proof of step 1,

we gain the conclusion $P_m[s] \Leftarrow Q_m[2]$ ， s=2,3,…,m－1,m, according to step 1and mathematics induction.

Theorem 4 $P_m[s] \Leftrightarrow Q_m[2]$, s=2,3…m,

Prove: according to theorem 3, do nothing but prove $P_m[s] \Rightarrow Q_m[2]$ ， s=2,3,…,m－1,m

Using mathematics induction to m, when m=2， $P_2[2]$ equipollence to:

$P_2[2]: l[0; a_{1,j} \oplus a_{2,i}, a_{1,i}, a_{2,j}] \le$

$l[0; a_{1,i} \oplus a_{2,j}, a_{1,j}, a_{2,i}]$

And $Q_2[2]: \min\{a_{1,i}, a_{2,j}\} \le \min\{a_{1,j}, a_{2,i}\}$

According to theorem 1, we have $P_2[2] \Rightarrow Q_m[2]$

supposing m=r－1, $P_{r-1}[s] \Rightarrow Q_{r-1}[2]$ ， s=2,3,…,r-1, namely:

$l^*[a_{1,i},...,a_{s,j}] \le l^*[a_{1,j},...,a_{s,i}]$， $2 \le s \le r-1$

$\Rightarrow \min\{a_{u,i}, a_{v,j}\} \le \min\{a_{u,j}, a_{v,i}\}$,

$(u,v) \in c_r^{[2]} - \{(1,r)\}$ …（10）

And

$l^*[a_{2,i},...,a_{s+1,j}] \le l^*[a_{2,j},...,a_{s+1,i}]$

$2 \le s \le r-1$

$\Rightarrow \quad \min\{a_{u,i}, a_{v,j}\} \le \min\{a_{u,j}, a_{v,i}\}$,

$(u,v) \in c_r^{[2]} - \{(1,r)\}$ ……（11）

if we want to prove $P_r[s] \Rightarrow Q_r[2]$ ， $2 \le s \le r$,

when m=r, only need to prove （10）、（11）

$\min\{a_{1,i}, a_{r,j}\} \le \min\{a_{1,j}, a_{r,i}\}$

When m=r, we can change

$l^*[a_{1,i},...,a_{r,j}] \le l^*[a_{1,j},...,a_{r,i}]$ to

$l[a_{1,i}, a_{r,j}, l^*[a_{2,i},...,a_{r-1,j}]] \oplus$

$l[a_{1,i}, a_{r,j}, a_{1,j}, a_{2,j},...,a_{r-1,j}] \oplus$

$l[a_{1,i}, a_{r,j}, a_{2,i},...,a_{r,i}] \le$

$l[a_{1,j}, a_{r,i}, l^*[a_{2,j},...,a_{r-1,i}]] \oplus$

$l[a_{1,j}, a_{r,i}, a_{1,i}, a_{2,i},...,a_{r-1,i}] \oplus$

$l[a_{1,j}, a_{r,i}, a_{2,j},...,a_{r,j}]$

three order term in both sides notes 1_L,2_L,3_L,1_R,2_R,3_R

in turn, $1_R \oplus 2_R \oplus 3_R$ have three probability:

Case 1 $\boldsymbol{1_R \oplus 2_R \oplus 3_R = 2_R} =$

$l[a_{1,j}, a_{r,i}, a_{1,i}, a_{2,i},..., a_{r-1,i}]$

Because $\boldsymbol{2_L \le 2_R}$ ， we gain

$l[a_{2,j}, a_{3,j},..., a_{r,j}] \le l[a_{2,i}, a_{3,i},..., a_{r,i}]$

…（12）

Because $\boldsymbol{3_L \le 2_R}$ ， we gain

$a_{r,j} \le a_{1,j}$ …（13）

According （11）、（12）、（13）,we gain

$a_{r,j} \le a_{r,i}$ …（14）

According （13）、（14）,we gain

$\min\{a_{1,i}, a_{r,j}\} \le \min\{a_{1,j}, a_{r,i}\}$

Case 2 $1_R \oplus 2_R \oplus 3_R = 3_R \Rightarrow$

$\min\{a_{1,i}, a_{r,j}\} \le \min\{a_{1,j}, a_{r,i}\}$

Case 3 $1_R \oplus 2_R \oplus 3_R = 1_R$

$\min\{a_{1,i}, a_{r,j}\} \le \min\{a_{1,j}, a_{r,i}\}$

The proof of case 2 and case 3 is similar as case 1, and the proof is omitted.

4 ROBUSTNESS OF SYSTEM n/m/p/C_{max} [ω , 1, 0]

According to robustness attemper [2,3], Yue-Han attemper [4], we gained:

Theorem 3 if $\omega^* \in W$ is the most excellent robustness attempering of attempering system n/m/p/C_{max} [ω , 1, 0], it is the most excellent attempering of this system.. If $\omega^* \in W$ is Yue-han most attempering of attempering system n/m/p/C_{max} [ω , 1, 0], it is the most excellent robustness attempering of this system.

Proof:

(1) If $\omega^* \in W$ is the most excellent robustness attempering of attempering system n/m/p/C_{max}[ω ,1,0], to any a>0, any b>0,any $\omega \in W$, C_{max}[ω^* ,a,b]≤ C_{max}[ω ,a,b].

Especially, when a=1,b=0 or b$\to 0^+$, to any $\omega \in W$, C_{max}[ω^* ,1,0]≤ C_{max}[ω ,1,0].

Therefore, $\omega^* \in W$ is the most excellent attempering of attempering system n/m/p/C_{max} [ω , 1, 0].

(2) According to Yue-Han most excellent attempering define, we know that if $\omega^* \in W$ is Yue-Han most attempering of attempering system n/m/p/C_{max} [ω , 1, 0], it is the most excellent robustness attempering of this system.

5 REFERENCES

[1]Mosheiov G, “Sheduling jobs simple linear deterioration”, Computers and Operations Research, 1994, 21(6):653-699.

[2] Bachman A, Janiak A, “ Minimizing maximum lateness under linear deterioration”, European Journal of Operational Research, 2000, 126(4):557-566.

[3]C.L.Zhao, "Sheduling Problems under linear deterioration", ACTA Automatica sinica, 2003, 29(4):531-535.

[4]M.Y.Yue, J.Y. Han, "Sequencing problem for the n jobs on m machines (1)",China Science, 1975,5:462-470.

This work was supported by the Emphases Project of Hubei Education Dept (No: D200517004), the Emphases Laboratory of Hubei Province weave new material and appliance (No: TKL200401) and the Project of Hubei Education Dept (No: 2004X074) to Mr. Yuan.

Unascertained Measure-----a New Tool of Safety Assessment for Platform Installation

Shujing Zhou
College of Civil Engineering, Hebei University of Engineering,
Handan, Hebei056038, China
Email: zhousj@hebeu.edu.cn
And
Yancang LI
College of Civil Engineering, Tianjin University,
Tianjin, Tianjin300072, China
Email: liyancang@163.com

ABSTRACT

To deal with the unascertained information and solve the problem of the data shortage in the safety assessment of the offshore platform installation procedure, a mathematical tool, the unascertained measure, was proposed. First, the basic knowledge of the Unascertained Sets and the essential difference between the Unascertained Sets and the Fuzzy Sets were analyzed systematically. Second, the unascertained measure was defined and the credible identification model was set up. Then the method was introduced into the field of safety assessment of jacket installation procedure. Finally, engineering practices show that it can complete the safety assessment systematically and scientifically without any assumption. The work provides a method for safety assessment and it has significance in theory and practice for offshore engineering.

Keywords: safety assessment, offshore platform, installation, unascertained measure, information entropy.

1. INTRODUCTION

The offshore platform has emerged as a promising structure in offshore engineering. And the catastrophic disasters related to them have caused environment pollution and enormous loss of lives and properties. The situation has signified the need for safety assessment of platforms. Up to date, most safety assessments dealt with the structure of the platform and its maintenance, repair and retirement, not much research has been done on the installation procedure [1, 2, 3, 4]. On the other hand, the methods employed [5, 6], such as fuzzy sets, have many shortcomings. How to solve the problems of uncertainty and the data lack and find out a new method for safety assessment of platform jacket installation are urgent tasks in offshore engineering.

According to this, a new method---the Unascertained Mathematic was introduced to solve the unascertained problem of the assessment. The unascertained measure was proposed and the credible identification model was set up for the safety assessment. Application results show that it can complete the safety assessment systematically and scientifically.

The structure of the paper is as follows. In the introductory part, attention was paid to the basic concepts of the Unascertained Mathematic. Also, the difference between Unascertained Sets and Fuzzy Sets was introduced. In the following part, a safety assessment mode for platform jacket installation based on the unascertained measure was set up. Then, its application in practice was introduced. Finally, compared with the Fuzzy Sets, the advantages of the method proposed here were pointed out.

2. BRIEF INTRODUCTION TO UNASCERTAINED MATHEMATIC

The Unascertained Mathematic, which is a tool to quantitatively describe the subjective uncertainty, was proposed by WANG Guangyuan in 1990 [7]. It mainly deals with the unascertained information which differs from the stochastic information, fuzzy information and grey information. The unascertained information refers to the decision-making-demanded information which has no uncertainty itself, but, because of situation constrain, the decision-maker can not grasp the whole massage of which. Uncertainty was produced by the decision-maker himself. Since 1990s, The Unascertained Mathematic has been used systemically and successfully in many fields [8, 9, 10]. For more knowledge, we can refer to [8]. Here, I will mainly discuss the difference between the Unascertained Set and Fuzzy Set.

1) On the space, the membership function of the Fuzzy Set is a single-variable function defined in the universe of U, while the membership function of the Unascertained Set is a two-variables function defined in $U \times E$, where U is the universe and E is a sigma-algebra.

2) On the definition, the Unascertained Set introduces the measure principles into its definition, and it clearly points out the nature space and measurable space in its definition.

3) On the calculation, the calculation principles of Fuzzy Set are correct for the same nature in sigma-algebra E; for the different nature, the numerical value vibration will cause inconsequence and its logic algorithm will obtain the error results. While, the Unascertained Set overcomes these shortcomings.

3. MODEL BASED ON UNASCERTAINED MEASURE FOR SAFETY ASSESSMENT OF JACKET INSTALLATION PROCEDURE

The system of assessment indexes is shown in Fig. 1.

3.1 Mode based on unascertained measure

We suppose that $x_1, x_2, \ldots, x_n$ are the n procedures of

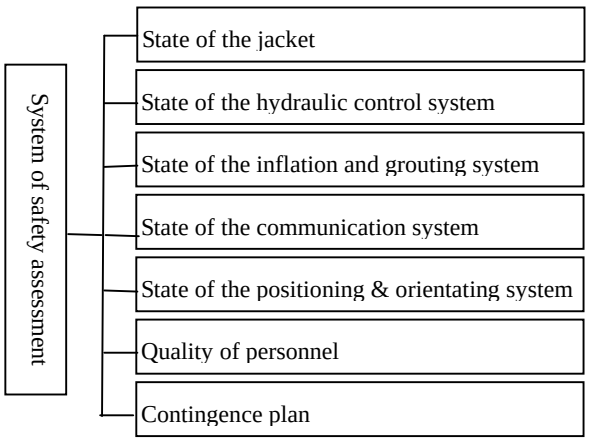


Fig. 1. The index systems of safety comprehensive assessment of process of jacket installation

The jacket installation, $X=\{x_1,x_2,...x_n\}$ is the universe, $I_1,I_2,...,I_m$ is the indexes for the assessment of x_i, $I=\{I_1,I_2,...I_m\}$ is the index space, x_{ij} is the observed value of x_i based on index I_j, $C=\{c_1,c_2,...c_k\}$ is the comment space, and c_k is the k^{th} comment $(1\le k\le K)$.

3.2 Single-index unascertained measure

Let μ_{ijk} be the degree which shows the subject x_i belongs to the assessment rank c_k under the index of I_j. It must satisfy the "three principles for a measure" which means nonnegative and limited principle, additive principle and Convergent principle.

The Delphi method was employed to get the scores of every factor. The number of the specialists is k. Every specialist should rank the degree that the index $I_j(1\le j\le m)$ belongs to the rank c_k $(1\le k\le K)$ using 0-10. If the k th specialist ranks the degree that the index I_j of the object x_i belongs to rank c_k is x_{ijk}, then $\mu_{ijk}=x_{ijk}/10$ is the unascertained measure that the observed value x_{ij} makes x_i located in rank c_k.

The single-index measure assessment matrix of x_i is:

$$\left(\mu_{ijk}\right)_{m\times k}=\begin{bmatrix}\mu_{i11}, & \mu_{i12}, & \cdots, & \mu_{i1k}\\ \mu_{i21}, & \mu_{i22}, & \cdots, & \mu_{i2k}\\ \vdots & \vdots & \vdots & \vdots\\ \mu_{im1}, & \mu_{im2}, & \cdots, & \mu_{imk}\end{bmatrix} \quad (1)$$

where $i=1,2,\cdots,n$.

3.3 Weight of every index

The weight of index means the influence degree of every factor to the classification of the plan. The common methods used to get the weight are Delphi method, Brainstorming, Analytic Hierarchy Process (AHP) and so on. In the paper, the information entropy which is employed to measure the uncertainty roughly [11] is used to give the weight of every factor.

The information entropy has the following characteristics:

1) Symmetrical characteristic: the value of the information does not rely on the order of p_i. That is to say:

$$S(p_1,p_2,\cdots,p_n)=S(p_n,p_{n-1},\cdots,p_1)=S(p_2,p_1,\cdots,p_n)$$

(2) Additive characteristic: the related independent state's information is equal to the total state's information entropy.

(3) Extreme property characteristic: when $p_i=\frac{1}{n}$ the information entropy has its biggest value $\ln n$.

(4) Nonnegative characteristic: $S(p_1,p_2,\cdots,p_n)\ge 0$.

Because $u_i\left(0\le u_i\le 1,\sum u_i=1\right)$ has the similar nature with p_i, let $H(\alpha)=-\sum_{k=1}^{K}u_{ijk}\log u_{ijk}$,

where $u_j^{(i)}=1+\frac{1}{\log k}H(\alpha)$,

$\omega^{(i)}=\left(\omega_1^{(i)},\ \omega_2^{(i)},\ \cdots,\ \omega_m^{(i)}\right)$ is the weight of the $I_1,I_2,\cdots,I_m$, where $\omega_j^{(i)}=\frac{u_j^i}{\sum_{j=1}^{m}u_j^{(i)}}$.

Suppose the measure which an object has some natures is the p_i of the information entropy, then

$$\omega_j^{(x)}=L_j^{(i)}/\sum_{j=1}^{m}L_j^{(i)} \quad (2)$$

Where

$$L_j^{(i)}=1+(1/\log k)\cdot H(\alpha)$$

and $H(\alpha)=-\sum_{\kappa=1}^{k}u_{ijk}\log u_{ijk}$. The weight ω_j^x is obtained.

3.4 Comprehensive assessment system

After the single index matrix and the weight are derived, the comprehensive assessment vector μ^i can be derived.

Let $\mu^i=\omega^i\cdot\left(\mu_{ijk}\right)_{m\times K}$

$$=\left(\omega_1^i,\omega_2^i,...\omega_m^i\right)\begin{bmatrix}\mu_{i11}, & \mu_{i12}, & \cdots, & \mu_{i1k}\\ \mu_{i21}, & \mu_{i22}, & \cdots, & \mu_{i2k}\\ \vdots & \vdots & \vdots & \vdots\\ \mu_{im1}, & \mu_{im2}, & \cdots, & \mu_{imk}\end{bmatrix} \quad (3)$$

Where $\mu^i=\left(\mu_{i1},\mu_{i2},...\mu_{iK}\right)$ describes the unascertained classification. In order to obtain the certainty classification, the identification is needed.

3.5 Principle of identification

Because the classification of the comment ranks is orderly, e.g. c_k is "better" than c_{k+1}, the identification principle of "maximum measure" is not available. The credible identification principle is needed. Let the credible identification be λ, where $0.5\prec\lambda\prec 1$, and it is always 0.6 or 0.7.

If

$$k_0=\min\left[\left(\sum_{l=0}^{k}\mu_{jl}\right)\geq\lambda,k=0,1,\cdots,K-1\right] \quad (4)$$

then x_i belongs to the rank c_{k_0}.

4. ENGINEERING APPLICATION

Using the method above, we finished the safety assessment of a jacket installation procedure of an offshore oil field in South China Sea. The installation procedure contains the lifting installation, jacket launch and upending operations. The scores of every factor given by the specialists are list in Table 1.

Table 1. Scores of the factors given by the specialists

ranks / factors	1	2	3	4	5
State of the jacket	0.3	0.1	0.1	0.2	0.3
State of hydraulic control system	0.2	0.2	0.3	0.2	0.1
State of inflation and grouting system	0.2	0.2	0.3	0.2	0.1
State of communication system	0.1	0.2	0.3	0.2	0.2
State of positioning and orientating system	0.2	0.2	0.2	0.2	0.2
quality of personnel	0.2	0.2	0.3	0.1	0.2
contingency plan	0.1	0.2	0.3	0.2	0.2

In the table shown above, the "rank1" means "best", "rank 2" means "better", "rank 3" means "normal", "rank 4" means "worse" and "rank 5" means "worst".

The single-index unascertained measure $\left(\mu_{ijk}\right)_{m\times k}$ can be obtained:

$$\left(\mu_{ijk}\right)_{m\times k}=\begin{bmatrix}0.3 & 0.1 & 0.1 & 0.2 & 0.3\\ 0.2 & 0.2 & 0.3 & 0.2 & 0.1\\ 0.2 & 0.2 & 0.3 & 0.2 & 0.1\\ 0.1 & 0.2 & 0.3 & 0.2 & 0.2\\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2\\ 0.2 & 0.2 & 0.3 & 0.1 & 0.2\\ 0.1 & 0.2 & 0.3 & 0.2 & 0.2\end{bmatrix}$$

The weight can be derived: ω_j^i=(0.15，0.1，0.1，0.2，0.15，0.1，0.2)

The assessment result can be obtain: μ^i=(0.20，0.22，0.18，0.20，0.20)

Let λ=6, according to format (4), the final assessment results can be obtain: the procedure belongs to the rank of "normal". Using the fuzzy comprehensive assessment, the procedure belongs to the rank of "better". All of the specialists agree that the former is more rational. The reasons are that the unascertained measure pays attention to the order of the assessment space and gives the rational rank and credible identification principles.

5. CONCLUSIONS

The jacket installation is a key procedure of offshore engineering. Its failure will cause catastrophic disasters. The safety assessment, which is an urgent task, can eliminate the possibility of failure, ensure the quality of construction. The method proposed in the paper has following characteristics:

1) Using the qualitative and quantitative analysis, it overcomes the disadvantage of subjective and perceptual assessment and makes the result more objective.

2) It is easy to realize the scientific and rational decision-making. The method can be used in the safety assessment of the loadout, towing, maintenance, repair and retirement.

6. REFERENCES

[1] Aggarwal P K, "Development of a Methodology for Safety Assessment of Existing Steel Jacket Offshore Platforms", the 22nd Offshore Technology Conference, Houston, 1990.

[2] Cole, G.K., Ronalds, B.F. et al, "The Interaction Between Strength and Fatigue Reliability for a Minimum Structure in Shallow Water", Journal of Offshore Mechanics & Arctic Engineering, vol.125(4), 2003, 281-287.

[3] W.L.Jin, J.Song et al, "Evaluation of damage to offshore platform structures due to collision of large barge", Engineering Structures, vol.27(9) , 2005,

1317-1326.
[4] Onoufriou T, “Reliability based on inspection planning for offshore structures”, Journal of Marine Structure, No.12, 1999, 521-559.
[5] Cornell, C.A., Rackwitz, R. et al, Reliability evaluation of tension leg platforms. Proceedings of the ASCE Specialty Conference on Probabilistic Methods and Structural Reliability. Berkeley, CA.1984, 159-162.
[6] G.Chen, S.K.Zhang, “Review of human reliability assessment (HRA) in ocean engineering”, Ocean Engineering, 2000, Vol. 18(4): 6-12.
[7] G.Y.Wang, “Mathematics treatment of unascertained information”, Journal of Civil Engineering Institute of Haerbin, 1990, vol. 23(4): 1-9.
[8] K.D.Liu, Q.K.Cao et al, “A method of fault diagnosis based on unascertained set”, Acta Automatic Sinica, 2004, vol. 30(5) 747-756.
[9] H.Y.Wu, H.Q.Wu, “Application of unascertained mathematics in project management”, Journal of Quantitative & Technical Economics, 2000,No.2, 21-24
[10] TENG Haiwen, HUO Da et al, “Application of unascertained rational numbers in structure reliability analysis”, Mathematics in Practice and Theory, 2005, vol. 35(4): 127-130.
[11] W.H.Qiu, Management decision making and applied entropy. Beijing: China Machine Press, 2002.

DDSS of Anti-Collision Warning of Construction Equipments For Hydropower Projects

Qingming Wu, Huadong Zhao, Zhiqiang Zhang
School of Power and Mechanical Engineering, Wuhan University
Wuhan, Hubei 430072, China
Email: qmwu@whu.edu.cn , haha_huadong@163.com , zhangzhiqiang78@263.net

ABSTRACT

In order to improve calculation velocity and warning effect of anti-collision warning system, which is used to be of stand-alone model, of construction equipment for hydropower station, this paper constructs a DDSS (Distributed Decision Support Systems) model, which is based on multi-agent theory, and realized distributed calculation. This model runs on intranet. After movement parameters of equipments were sent to the model from GPS devices, the interfere examination agents calculate the probability of interference and send the results to the synthesizing agent on the server. Then the synthesizing agent educes the type of warning and sends the warning to the operators. An anti-collision warning system of construction equipments for Longtan hydropower station was constructed based on this model with web service technology. According to test, this model improves the velocity and effect obviously.

Keywords: Multi-Agent; DDSS; Construction Equipment; Hydropower Station; Anti-collision.

1. INTRODUCTION

There are many huge hydropower projects, which are going to be built, or being built in China, such as Three Gorges, Longtan, Yantan, Ertan, etc. In the process of construction, many huge construction equipments are frequently used, such as tower crane, moving tower crane, topbelt, tower-belt and cable crane, etc. Any equipment occupies large work area, which overlaps with others'. Also, these equipments have high work intensity and high moving velocity. For aforementioned reasons, collision is easy to happen and may cause great influence on construction.

In order to avoid collision, we developed an anti-collision warning system of construction equipments. In the primal design, the decision model is designed to run in a single computer, thus, the calculation velocity is limited seriously and the warning effect is not good enough. To get higher velocity and better warning effect, this paper constructs a distributed warning decision model based on multi-agent theory.

2. AGENT AND DDSS

2.1 Agent theory

The word agent has found its way into a number of technologies. It has been applied to aspects of artificial intelligence research and to constructs developed for improving the experience provided by collaborative online social environments (MUDS, MOOs, and the like). It is a branch on the tree of distributed computing. It's difficult to find a succinct definition for agent. Workers involved in agent research have offered a variety of definitions, each hoping to explicate his or her use of the word "agent." This paper just gives two of them.

The Woodridge and Jennings definition is: "... a hardware or (more usually) software-based computer system that enjoys the following properties [1]:
autonomy: agents operate without the direct intervention of humans or others, and have some kind of control over their actions and internal state;
Social ability: agents interact with other agents (and possibly humans) via some kind of agent-communication language.
Reactivity: agents perceive their environment, (which may be the physical world, a user via a graphical user interface, a collection of other agents, the internet, or perhaps all of these combined), and respond in a timely fashion to changes that occur in it;
Pro-activeness: agents do not simply act in response to their environment; they are able to exhibit goal-directed behavior by taking the initiative."

The definition of Stan Franklin and Art Graesser is: an autonomous agent is a system situated within and a part of an environment that senses that environment and acts on it, over time, in pursuit of its own agenda and so as to affect what it senses in the future [2].

A Multi-Agent System (MAS) is a system composed of a population of autonomous agents, which cooperate with each other to reach common objectives, while simultaneously each agent pursues Individual objectives.

2.2 Distributed Decision Support Systems

Decision Support Systems (DSS) has been an active area of research for over twenty years. Early research, focusing on support for an individual decision maker, has been extended to support decision making in groups and organizations: Distributed Decision Support Systems (DDSS) [3, 4, 6, 7]. Jarke[5], for instance, characterized DDSS as an organization of men owning different knowledge and interests with the machine facilitating group problem solving, and Swanson [5] explored the environment of and strategies for DDSS. The key characteristics of DDSS that present challenges for their development are:
1) More than one problem domain, or multiple models of the same domain, may be involved in the decisions to be supported by the system. The construction and management of such knowledge in a centralized system is infeasible and, more than likely, inefficient.
2) Although the construction of a centralized system is infeasible, the relationships between the problems solving of

individual models must be coordinated. For example, the solution produced by one model may depend on that of another.

3) The problem solving is often naturally distributed (multiple human decision-makers are typically involved), and so it is sensible to support distributed decision-making using a system that supports this distribution. However, it is essential for the system to enable each individual to retain an appropriate level of autonomy over his or her knowledge and inference.

3. DDSS MODEL OF ANTI-COLLISION WARNING OF CONSTRUCTION EQUIPMENTS

3.1 Outline of anti-collision warning system of construction equipments

As described in Figure 1: "cable crane", "tower crane" and "belt machine" (topbelt and tower-belt) is the frequently using and most important machine in huge hydropower projects. All these equipments have complex movement styles and their combining movements are even more complex. The position, where the GPS devices were installed, is also described in Fig. 1.

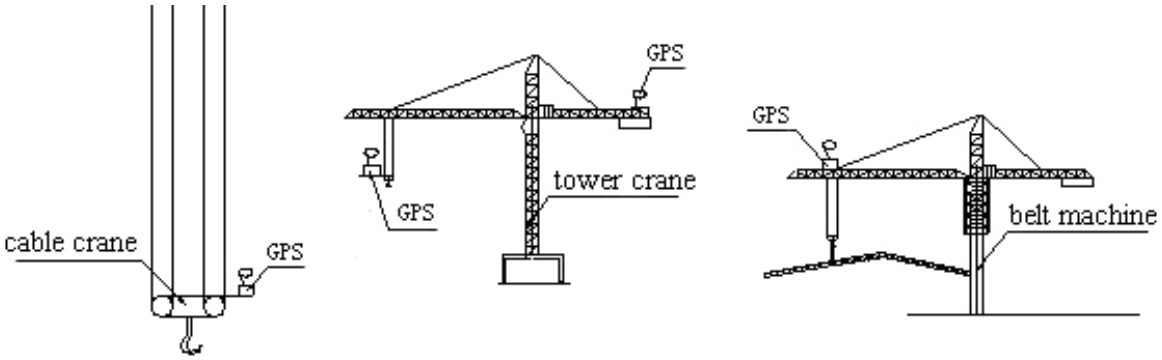


Fig.1. Main equipments of hydropower projects

The DDSS is the hard core of the anti-collision warning system of construction equipment. The method of anti-collision warning system: GPS devices are installed at the key positions of the equipments; they send all position and movement parameters of equipments to the base station in which the DDSS is running; then, base station judges the possibility of the collision according to the received data and sends corresponding warning signal to the related operators.

3.2 The model of DDSS

The model runs on intranet and each module locates corresponding node. The configuration of the system is described as Fig. 2.

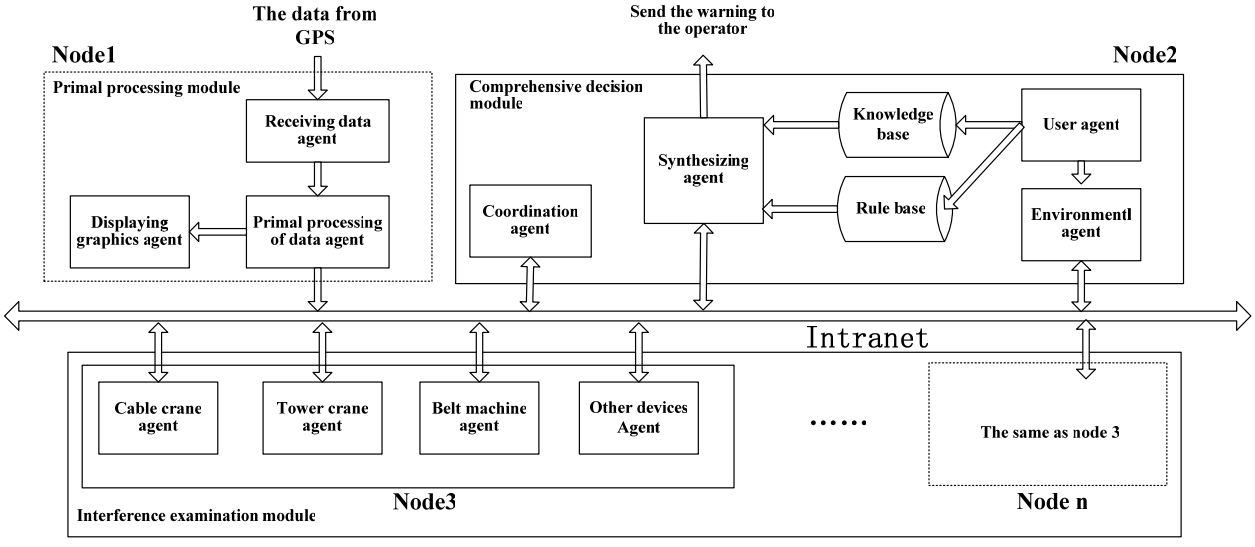


Fig.2. DDSS model of the anti-collision warning system

The system consists of three parts: "primal processing module", "interference examination module" and "comprehensive decision module". They are distributed on different nodes and cooperate with each other while making the anti-collision warning decision.

"Primal processing module" mainly receives data from every GPS device, processes the data primitively and displays the graphics of devices; then send the data to other nodes. For the data that come from GPS are the coordinates of the GPS, "primal processing module" should transform them into key parameters of the equipments.

"Interference examination module" mainly calculates the possibility that certain equipment hits others in coming time *T*. The movements of devices are very complex, so it needs a lot of calculation for each calculation. To accelerate circulation of the system, this paper distributes the module on many nodes and realizes the parallel processing. The "coordination agent" assigns the task of interference examination.

"Comprehensive decision module" mainly judges which type of warning signal should be given according to the knowledge in knowledge base and the rules in rule base after receiving the possibilities of interference from every node. There are a lot of interference styles and more combining interference styles in the process of construction. The warning signals vary greatly according to different interference styles, such as the warning signals for the collisions of two equipments that move in the same direction and in opposite direction. In order to educe the correct warning signal in the complex interference situation, the system must have perfect knowledge base and rule base.

This model runs on intranet. After movement parameters of equipments were sent to the model, the interfere examination agents, such as "cable crane agent", "tower crane agent" and "belt machine agent", calculate the probability of interference and send the results to the synthesizing agent on the server. Then the synthesizing agent educes the last warning and sends the warning to the operator.

3.3 The function of main Agents and communication mechanism

Primal processing agents: "receiving data agent" receives the data from GPS devices and sensors; "primal processing of data agent" transforms the primal data to the key parameters of devices; "displaying graphics agent" displays the position and situation of devices.

"Interference examination agents": mainly includes: "cable crane agent", "tower crane agent" and "belt machine agent"; they check if one device interferes with others in coming time T.

"Coordination agent": it distributes the tasks of interference examination to different nodes in the intranet by Contract-Net Protocol and optimizes whole performance.

"Environmental agent": in the process of construction there are many parameters which need changing, such as the height of tower crane. For one device, the others are also its environment. The "environment agent" saves all the parameters of the environment. And the parameters aforementioned should be input by user through user agent.

"Synthesizing agent": it synthesizes all the interfering results and makes the decision that which type of warning was sent out.

This system use message mechanism of web service to realize communication mechanism.

3.4 Data flow graph of DDSS

The basic system model of DDSS is described in Fig. 3. It shows relations among DDSS and other parts of the anti-collision system.

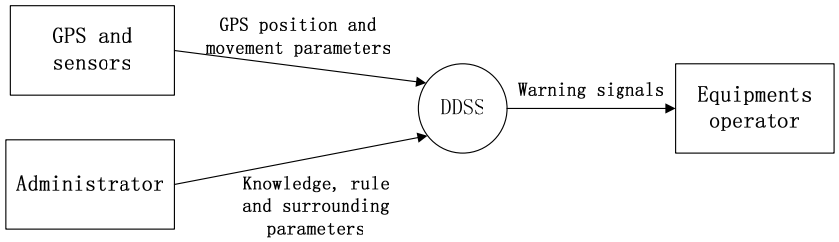


Fig.3. Basic system model of DDSS

The data flow graph of DDSS is described in Fig. 4. It shows transmission and processing of data.

3.5 The realization of agent

There is common ground between AOP (Agent Oriented Programming) and OOP (Object Oriented Programming), such as encapsulation of data and methods, inheritance and polymorphism, etc. But there are obviously differences in some aspects: agent has stronger concept of autonomy; agent can make decision whether perform the application of other agents; agent has autonomy, social ability, pro-activeness and reactivity while standard object of OOP do not has these characteristics; Multi-Agent System has multithreading naturally.

From the comparison above, we can find there are some important differences between AOP and OOP, but the object of OOP can still realize the agent. In fact, there are many agent development tools and applications that are realized on the base of OOP. Moreover the characteristics of agent aforementioned can be improved on object.

3.6 The anti-collision warning systems of construction equipment for Longtan hydropower project

The anti-collision warning systems of construction equipment for Longtan hydropower project was realized on this model. Longtan hydropower project is the highest roller compaction concrete gravity dam in the world. The main construction equipments are: TC2400 Towerbelt,

MD2200 Topbelt, MD1800 Moving tower crane MD2200 tower crane and two cable cranes.

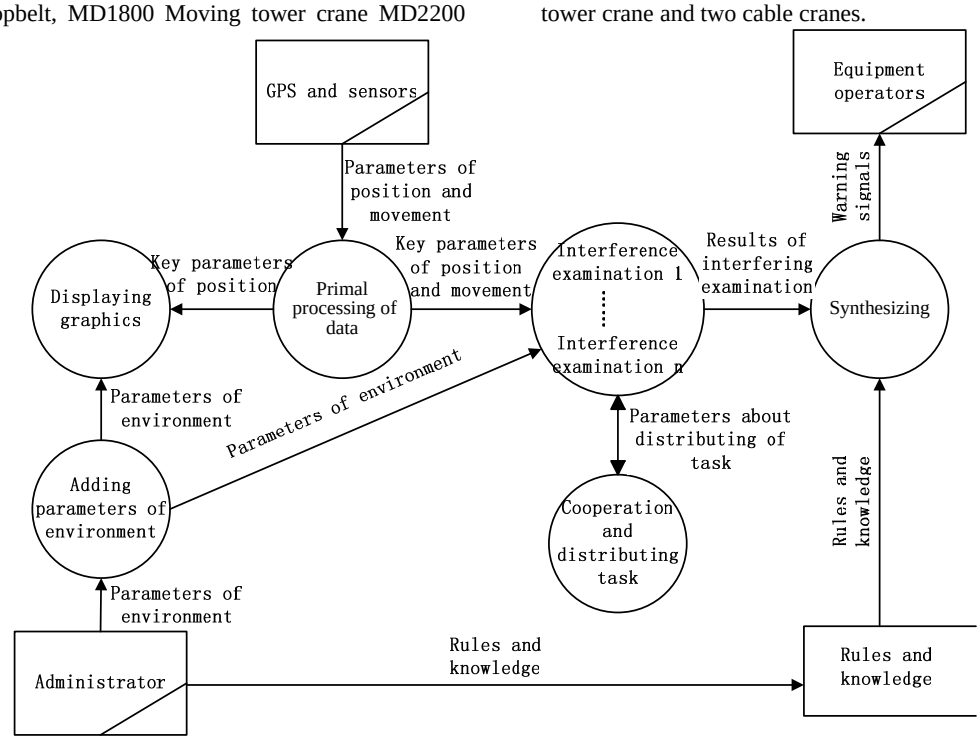


Fig.4. Data flow graph of DDSS

The base station includes an intranet, which has six nodes: node 1 is responsible for primal processing, node 2 is responsible for comprehensive decision and other nodes are responsible for interference examination. The model is described as the Fig. 5.

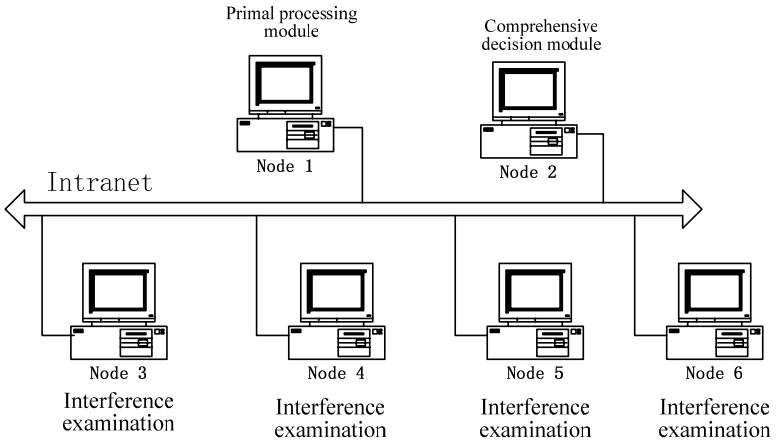


Fig.5. Hardware components of anti-collision warning system for Longtan

This model is realized by C# and Web Service. C# is a new programming language designed for building a wide range of enterprise applications that run on the .NET Framework. An evolution of Microsoft C and Microsoft C++, C# is simple, modern, type safe, and object oriented. It has excellent function for distributed calculation. There is much common ground between Web Service and agent, so it is a suitable choice using Web Service to construct Agent. In simulation test, each time of prediction cost about 0.15 second averagely in stand-alone model and about 0.08 second averagely in the DDSS model.

4. CONCLUSIONS

This paper constructs the DDSS model of the anti-collision warning system of construction equipment for hydropower projects and realizes the model on Longtan hydropower project with OOP and Web Service technology. The future works are optimizing the model to avoid bottleneck; further researching on technology of realizing agent.

5. REFERENCES

[1] M. Wooldridge, N. Jennings, Intelligent agents: theory and practice. The Knowledge Engineering Review. Vol.10, No.2, 1995, pp. 115-152.

[2] S. Franklin, A. Graesser, "Is it an Agent, or just a Program? A Taxonomy for Autonomous Agents" Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, Springer-Verlag, 1996, pp. 21-35.

[3] C.H.Ju, G.M.Wang, "research of decision support system on commerce", Proceedings of the International Conference on Management Science & Engineering, 1999, pp. 156-161.

[4] H. M. Chung, R. Mahapatra, "Distributed decision support system: Characterization and Design Choices", Proceedings of the 26th Annual Hawaii International Conference on System Sciences, 1993, pp. 660-667.

[5] M. Jarke, "Knowledge sharing and negotiation support in multi-person decision support systems", Decision Support Systems, 2, 1986.pp.93-102.

[6] J. M. Scher, "Distributed decision support for management and organization", In Proceedings of the 1st International Conference on Decision Support Systems, 1981, pp. 130-140.

[7] R. C. Thomas, A. Burns, “The case for distributed decision making systems” The Computer Journal, Vol.25, No.1, 1982. pp. 148-152.

Improved GMDH and Its Application

Jie Chen
School of Electronic & Information Engineering, Wuhan Institute of Technology, Wuhan, Hubei 430073, China)
Ling Zhou
School of Energy and Power Engineering, Changsha University of Science and Technology, Changsha, Hunan 410076, China)
Chenghua Yin
Faculty of Engineering, Dalhousie University, Halifax, B3H 1R2, Canada)
E-mail:ch58j@163.com

ABSTRACT

An improved GMDH(group method of data handling) algorithm — GMDH plus model structural optimization was put forward . The variables were self-organized by group and some weak effect was removed to optimize the model structure in the algorithm. When the algorithm was applied into the manufacture of lubricating oil for identifying the mathematical model of the dew axing process of phytol ketone, the acquired model was consistent with the actual operating data. Based on this model, an optimal deduction was made for locating the optimal operating order measured by the biggest absorbing-rate. Theory analysis and simulation result show that the algorithm is effective.

Key words: GMDH, recognition, model structural optimization, identification, lubricating oil production.

1. INTRODUCTION

There are always scholars dedicating their researches to the improvement of GMDH, which was put forward in the late of 1960's and widely applied into the identification and control of uncertain systems with huge structures and numerous variables in the fields of economy, environmental ecology, process control and etc, and some results were satisfactory; the difficulty in modeling brought by uncertain process mechanism was reduced to some extent. Nevertheless, existing GMDH possess some shortages, such as numerous original variables and substitutive variables as well as a lot of calculation, which bring difficulties to the reasonable choice of the substitution. For instance, if a binary quadratic is used as some expressions for a multivariable circumstance, the error will be great; if a lot of variables exist in the final expression, the structural optimization cannot be made(re.[1,2]). Aimed at these kinds of problems, this paper proposes an novel improved method; that is, GMDH plus model structural optimization (GMDH+MSO).

2. IMPROVED GMDH

Basic structure of GM

DH,such as Fig. 1. Show.

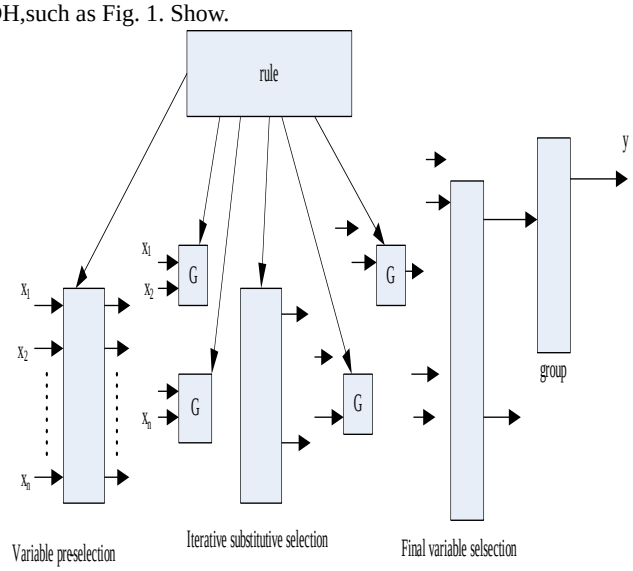


Fig. 1. Basic structure of GMDH

Some improvements are made in this paper to overcome the shortages of basic GMDH, as listed below(re.[3]):

a) The variables are self-organized by group, usually 4 variables a group, to reduce all kinds of difficulties brought by the excessive combination of two variables under the multivariable circumstance.
b) Partial expressions are ternary quadratic.
c) The interactive items among the groups of variables, which are lost during grouping, are stored.
d) The substitutions of weak effect are removed to optimize the structure.
e) The full expression is structurally optimized to get the final model. The concrete method is: first regroup all variable items, which enter the final model after organization and iteration level by level, as partial expressions; then remove the substitutions of weak effect in Z_k and get Z_k^*, meanwhile using least

square algorithm, compare AIC or RSS value of Y and Y^* to decide if the remove is correct; repeat the above steps until the predefined accuracy is reach and obtain the final model.

The structure of the improved GMDH, i.e., GMDH+MSO, is as shown in Fig. 2.

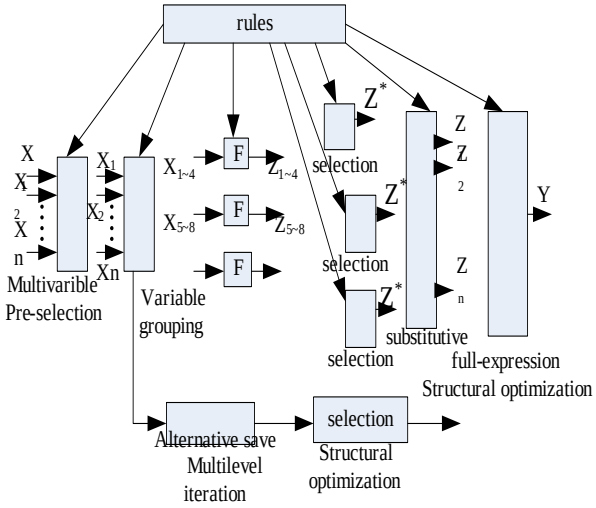


Fig. 2. Structure of GMDH+MSO

3 APPLICATION STUDY

As an application, the method of GMDH+MSO is introduced into the model identification of the process of benzd-acetone in the lubricant production and the calculation is optimized to reach the optimal operating point.

3.1 Brief introduction to the process

The lubricant production is a process obtaining lubricants by cooling, crystallizing and dew axing the raw materials and other technologies. The effect of the dew axing influences directly on the quality of the product and the absorption rate. Fig. 3. Shows the flow graph of dew axing technological process.

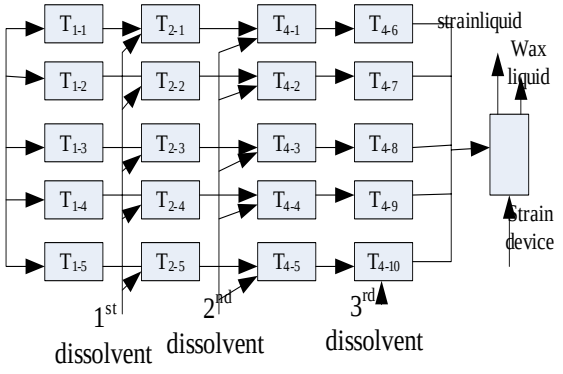


Fig. 3. The flowchart of benzd-acetone dewaxing technology

3.2 Modeling of the process

The key to enhance the rate of absorption is to control the crystallization, provided that the quality of the products is guaranteed. So far however, the mechanism of crystallizing and dew axing can not be put across, and it is difficult to analyze quantitatively the relationship between the rate of absorption and the state of the crystallization, and the relationship between the state of the crystallization and the variables. The process, affected by numerous factors, is a typical nonlinear multivariable "black box" system. In this paper, eight critical variables of the process are chosen, and then the method of GMDH+MSO is used to build the model of the process with the rate of absorption as its output based on the operands. With six organized iterations and structural optimization, the final model is obtained as follow:

$$y=10.51-0.019x_1-28.17x_2-9.18x_3-3.1x_4-0.004x_5+0.003x_6+0.752x_7+0.037x_8+0.95x_2^2+1.988x_3^2+0.109x_2x_7+3.12x_3x_4-0.73x_4x_7-0.1x_2x_8-7.2x_2x_3+25.63x_2x_4 \quad \text{Eq. (1)}$$

3.3 Calculation of the optimal operating point

In the practical industrial process, the steady-state operating point of the production is really not the optimal operating point. However, the optimal operating point can be obtained by optimization according to the mathematical model of the process so that the performance is the best. The rate of absorption is thought as the performance index of the dew axing of benzd-acetone. For the already established mathematical model, x_2、x_3、x_4 are the first, the second and the third dissolvent rate, respectively. When the total dissolvent rate is a constant, the reasonableness of its distribution to the first, the second and the third dissolvent rate is questionable, and the optimal distribution to three dissolvent rates will vary with other variables. Some optimal points are calculated according to the established model, as shown in Table 1. , in which five optimal points are compared with practical operating points. The results indicate that if the calculated optimal values are set as he first, the second and the third dissolvent rate, the rate of absorption will be enhanced obviously, also the commercial benefits (re. [4]).

Table 1. Simulation result

	practical operating point				optimal computing value			
	x_2	x_3	x_4	y	x_2	x_3	x_4	y
1	0.354	0.618	1.023	69.3%	0.36	0.53	1.2	73.68%
2	0.375	0.667	0.936	69.8%	0.357	0.58	1.16	74.38%
3	0.375	0.654	0.974	70.1%	0.358	0.56	1.16	74.5%
4	0.363	0.634	1.057	71.2%	0.346	0.515	1.14	75.4%
5	0.357	0.623	1.037	68.9%	0.36	0.61	1.19	72.5%

4. CONCLUSION

The method of GMDH+MSO proposed in this paper can be applied into the modeling of industrial process of multivariable, nonlinearity and uncertain mechanism. The application results show that it is reliable, convenient, and accurate to some extent.

References

[1] Z.H. Yang, "Original Variable Save Arithmetic", Automation Journal, No. 10, 1986.

[2] C. M. Close, D. K. Frederick, Modeling and Analysis of Dynamic Systems, 2nd ed., Houghton Mifflin, Boston, 1995.

[3] W.J. Palm, Modeling Analysis and Control, 2nd ed., John

Wiley & Sons, New York, 2000.
[4] R.C. Dorf, R.H. Bishop, Modern Control Systems, 10th ed., Science Press, Beijing, 2005

A Combined Filtering Approach for Chinese Papers and Its Empirical Results

Lijun Bai, Yujia Ge, Changbing Jiang
College of Computer and Information Engineering, Zhejiang Gongshang University
Hangzhou, Zhejiang 310035, China
Email: blj@mail.zjgsu.edu.cn, yge@mail.zjgsu.edu.cn, johncabin@mail.zjgsu.edu.cn

ABSTRACT

With the explosive growth of information and the popularization of the Internet, it is necessary to have an efficient approach to facilitate paper searching for researchers. While existing general purpose search engines cannot satisfy the needs of users with specific background and objective, Information Filtering (IF) could be seen as a solution to solve this problem. This paper examines two kinds of information filtering approaches: content-based filtering and collaborative filtering. But each of them has its limitations and filtering results are then affected. Aiming to make up their weaknesses complementarily, a new combined approach of information filtering is proposed. The paper also reports the experiments to compare the filtering results of these three approaches on Chinese technical papers. The experiment results show that the combined filtering approach can provide better recommendation accuracy than content-based filtering or collaborative filtering.

Keywords: Information Filtering, Content-based Filtering, Collaborative Filtering, User Profile, and Filtering System.

1. INTRODUCTION

With the rapid development of IT and the Internet, people always feel helpless to find what they exactly need before the huge amount of data (digital magazines, periodicals, and technical reports and so on), which leads to a serious problem called "information overload" or "information confusion". Therefore, research on retrieving data by users' preference while hiding irrelevant information, is recently becoming an important research area. While existing general purpose search engines cannot satisfy the needs for the inquiries of users with specific background and objective, Information Filtering (IF) could be seen as a solution to solve this problem [1, 2].

Online papers, including Chinese papers, have become a major resource for Chinese researchers to learn about the status of their fields. It presents a need for an efficient information filtering approach to prioritize Chinese papers so that researchers can spend less time searching for papers of interest. Two of the most representative techniques in IF are (1) Content-Based Filtering (CBF), and (2) Collaborative Filtering (CF), also known as "Social Filtering". This paper proposes a novel filtering approach, which makes use of the outstanding characteristics of CBF and CF, and overcome their limitations complementarily. The approach then improves the performance of an information filtering system.

2. CONTENT-BASED FILTERING AND ITS LIMITATIONS

2.1 Content-Based Filtering

CBF is the most basic method among filtering techniques [3]. It mostly uses natural language processing, artificial intelligence, probabilistic statistics, and machine learning techniques. The information needs for each user is represented as a user interests profile, i.e., a user vector. The process of word segmentation, indexing and counting word frequencies in papers generates text vectors. Then we can compute the degree of match (M) between a user vector and a text vector as Eq. (1).

$$M = \frac{2|D \cap Q|}{\min(|D|, |Q|)} \quad (1)$$

where D is the elements in a text vector and Q is the elements in a user's profile.

The texts with relatively high degree of match will be recommended to the registered user. After the user receives the text item, he/she can rate the item by some criteria, such as relevant or irrelevant. The feedback rating information can be utilized for maintaining or revising the user profile. The interaction between users and the System consequently improves the efficiency of the filtering process and the quality of retrieved information [4].

2.2 Limitations of Pure Content-Based Filtering

Content-based filtering is based on the similarity of information resources with the user's interests. Each user operates separately by him and is not related to others' interest. Although content-based filtering is simple and effective, it still has several problems:

(1) It is difficult to evaluate the quality of filtering results from the same subject. In a long run, the efficiency and quality of CBF will be decreased with the increase of items in one subject.

(2) It is not able to exploit new interests of a user and can only give the user the answer derived from his/her known interests.

(3) Choosing an inappropriate CBF algorithm can result in the inaccurate computation, which leads to retrieving irrelevant items.

3. COLLABORATIVE FILTERING AND ITS LIMITATIONS

3.1 Collaborative Filtering

Collaborative Filtering is becoming a very popular technique among information filtering approaches and recommendation systems [5, 6]. Collaborative filtering is to analyze the user's interest and to find other users who have similar interests. By statistically analyzing the rating information from this group of users, the user's preference to a specific information item can be predicted. The process

of CF can be illustrated in a user-item matrix. Table 1 shows an example matrix.

Table 1. Example of User-Item Matrix

	Item1	Item2	Item3	Item4
User1	?	2	3	4
User2	3	?	2	2
User3	3	4	1	?
User4	1	2	4	3

In Table 1, the values in the matrix are preference ratings for an item by a specific user (1 means the lowest preference and the highest is 4). All the unavailable values marked by "?" need to be predicted later. The keystone process of CF is to predict these unknown preference ratings by known ones.

So the procedure of CF has two steps: First, analyze the similarity (S_{xy}) between *user x* and *user y* using constrained *Pearson correlation coefficient* [7] as Eq. (2).

$$s_{xy} = \frac{\sum_{j \in I_{xy}} (r_{xj} - \bar{r}_x)(r_{yj} - \bar{r}_y)}{\sqrt{\sum_{j \in I_{xy}} (r_{xj} - \bar{r}_x)^2} \sqrt{\sum_{j \in I_{xy}} (r_{yj} - \bar{r}_y)^2}} \quad (2)$$

where I_{xy} represents the item set rated by both *user x* and *user y*. $\bar{r}_x$ is the mean rating of *user x*, r_{xj} is the rating by *user y* to *item j*. Secondly, choose users with similarity higher than a certain value to get rid of noise brought by individuals with different interest. Then compute the weighted average of the ratings from these users. A general formula to get a preference prediction (r_{xj}) by *user x* to *item j* is as Eq. (3).

$$r_{xj} = \bar{r}x + \frac{\sum_{y=1}^{n} s_{xy}(r_{yj} - \bar{r}_y)}{\sum_{y=1}^{n} |s_{xy}|} \quad (3)$$

3.2 Limitations of Pure Collaborative Filtering

Compared to CBF, collaborative filtering has its advantages as follows:

(1) It can filter some information, which is difficult for automated content analysis, such as arts, music, films, etc.

(2) Filtering information items can be based on complicated concepts, such as quality of information, taste.

(3) It is outstanding in recommending a novelty to the user.

On the other side, CF has its own limitations:

(1) Early Rater Problem: At the initial stage of a System, there are only a few users to give their ratings. A large number of items have not gained any evaluation, which is Early Rater Problem. For instance, within a collaborative filtering system of technical papers, if one paper has not been evaluated, the system is not able to make a recommendation of the paper to a user by prediction.

(2) Sparse Matrix Problem: Since the number of information items is much more than what the user needs. He/She may not be willing to give reviews. A sparse matrix will lead to difficulty in computation to get similar interests and search sources.

(3) Scalability Problem: The prediction algorithm is designed to search all of the users with similar interests. The computation burden will accordingly be increased and the system performance will obviously be worse when users and information resources are increased.

4. COMBINED FILTERING APPROACH

4.1 Our Objective

Analysis on complementary of CF and CBF are furthered proceeded. As to *Limitation 1* and *2* of CBF techniques as we mentioned previously, CF can mitigate them in some extent by making use of users' preference rating information extensively. If users with similar interests usually give high ratings to certain items with high quality and give low ratings to items with bad quality, users can tell the difference between items with different quality easily. And the system can recommend good information items to users with similar interests later. For *Limitation 3*, if joined with collaborative filtering and after the rating level is increased, accuracy of CF results will be enhanced. Herewith the system with the combination of both techniques may have better performance.

To solve the problems related to *Limitation 1* and *2* of CF, we can integrate it with CBF. CBF is based on the similarity of information resources and users' interest. It is not related to any rating information but the degree of similarity. So an item with high similarity can be recommended to the user even before any one reads and evaluates the information. On the other side, users can evaluate the content to increase the amount of rating information. These characteristics will help the filtering system overcome the early rater problem and the sparse matrix problem.

In one word, to make the filtering system more accurate and provides better-personalized results to users. Combining these two techniques in a good way can be seen a solution to improve the system performance.

4.2 Combining Filtering Techniques

There are already some researches of combining CBF and CF techniques, the performance of those systems all increased in some extent [8, 9]. The key point is to select a good combination method. In this paper, we define the prediction of interests (F) as a weighted average of CBF and CF. The formula is as Eq. (4).

$$F = \alpha * F_{CBF} + \beta * F_{CF} \quad (4)$$

where α and β satisfies the conditions $\alpha+\beta=1$ and $0<\alpha<1$, $0<\beta<1$, F_{CBF} is the weight of CBF, F_{CF} is the weight of CF. To gain the best combination value of α and β, this paper represents a prototype of an automated filtering system for Chinese texts. It consists three main parts: client, server and database, as shown in Fig. 1.

The "user profile generation" module forms user profiles of their interests by analyzing users' needs. The "match" module is in charge of content-based filtering. The "browse or evaluate" module let user browse or evaluate the filtering results in a visualization interface. The "evaluated item sets" includes the texts with user evaluations, which can be used by the "feedback" module to refresh "user interest profiles", or be cooperated with the "collaborative prediction" module to make better recommendations to users. "Filtering engine" is the kernel of the system. It forwards the results to users from the "match" module or the "user

interests profile".

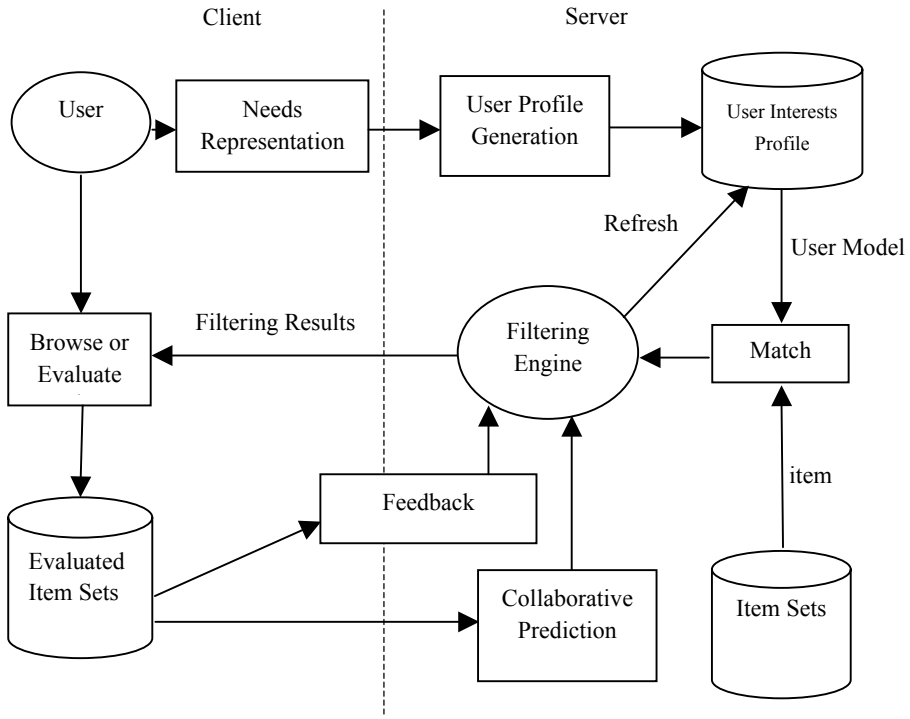


Fig. 1. Architecture of a Prototype Filtering System

4.3 Preliminary Weight Setting

To determine the best value of α and β in $F=\alpha*F_{CBF}+\beta*F_{CF}$, we choose 400 Chinese papers in Computer Science and 5 user interests profiles as training data sets.

In the server side, the procedure of the experiment is as follows.

(1) Upload 50 papers each time.

(2) Process a series of initial operations:

- File format transition: from html files to txt format;
- Word segmentation in Chinese: using a component of Word Segmentation 2000 [10];
- Vector generation: Retrieve key words in files, count and generate text vectors.

(3) Do content-based filtering: send files filtered by CBF to every relevant user.

(4) Do collaborative filtering: recommend files to different users by CF.

In the client side, do the combined approach to every user and analyze recommendation accuracy (P) by different α and β. Recommendation accuracy is defined as Eq. (5).

$$P = \frac{total\ number\ of\ relevant\ papers}{total\ number\ of\ recommended\ papers} \quad (5)$$

where total number of relevant papers refers to the paper received rating as 5 (very relevant) and 2 (relevant).

The experiment process in the client side is:

(1) Given specific values of α and β, let each user evaluate the filtering results. Users can choose one of these 4 ratings: 5 (very relevant), 2 (relevant), -2 (irrelevant), -5 (totally irrelevant). The one who has the highest recommendation accuracy (P) is selected as the first user. Repeat this step for several times. Then we can get the average recommendation accuracy under the status for this user.

(2) Delete the user' interests profile and his/her evaluation history. This step is to re-initiate the filtering system and make the filtering accuracy only affected by the values of α and β.

(3) Change the values of α and β, the user re-registered the same interest again. Server then re-distributes the same bunch of papers to the users.

Continue step (1) (2) (3). Then we can gain the lines of recommendation accuracy with different users against α/β depicted in Fig. 2. For each β/α, the mean recommendation accuracy of those 5 users is calculated. The result is listed in Table 2. From Fig. 2 and Table 2, we can see that when α = 0.4 and β = 0.6, i.e., β/α = 1.5, the average recommend accuracy is the highest one with the value of 0.82. So it is reasonable to be considered as the best weight for our later experiments.

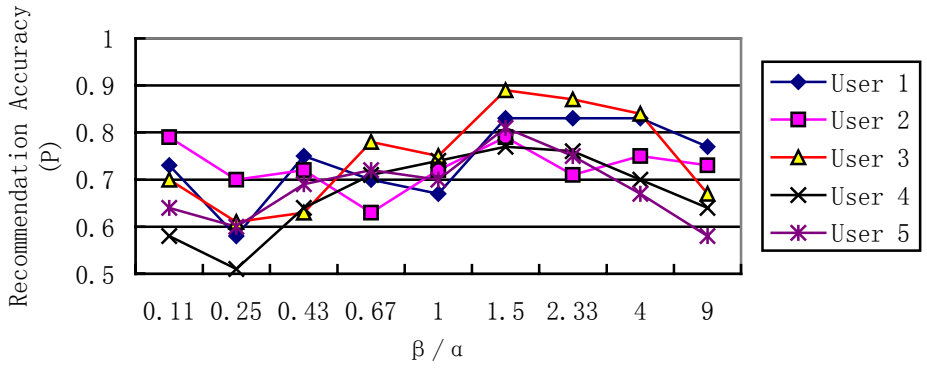


Fig. 2. β/α versus Recommendation Accuracy P

Table 2. Average Recommendation Accuracy of 5 users

β/α	0.11	0.25	0.43	0.67	1	1.5	2.33	4	9
Average Recommendation Accuracy	0.69	0.6	0.69	0.71	0.72	0.82	0.78	0.76	0.68

5. EXPERIMENTAL RESULTS

To test the performance of our approach, an experiment with 700 Chinese Computer Science papers was conducted to compare CBF, CF with the combined one. During the experiment, new users are registered randomly.

In the server side, upload 60 papers each time, process content-based filtering and collaborative filtering respectively, and save interests prediction to an interest prediction table.

In the client side, we compare pure content-based filtering, pure collaborative filtering and combined filtering. We recommend papers with high values of interest prediction F that is calculated by $F=\alpha*F_{CBF}+\beta*F_{CF}$. When $\alpha=1$ and $\beta=0$, CBF is processed; When $\alpha=0$ and $\beta=1$, CF is conducted; In other status as $\alpha+\beta=1$ ($0<\alpha<1$, $0<\beta<1$), the combined filtering is conducted. The comparison results of average recommendation accuracy of these three filtering techniques are shown as Fig. 3.

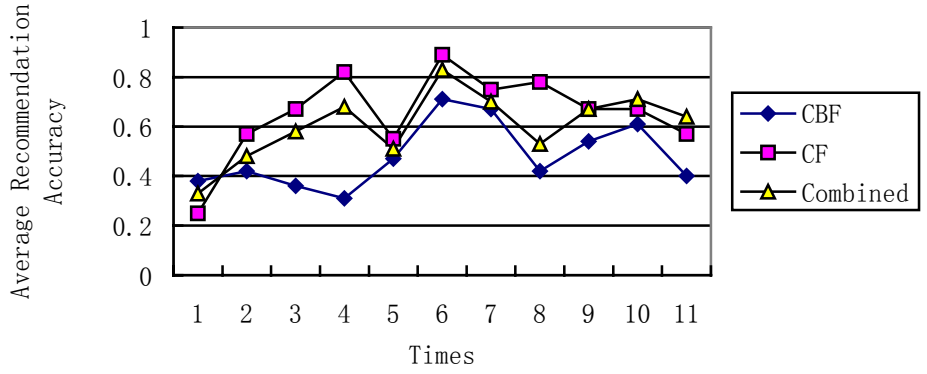


Fig. 3. Comparison of Average Accuracy of CBF, CF and Combined Approach

From Fig. 3, we can see that CBF has better performance than the other two only in the first time of filtering. In the other rounds, the performance of CF and combined filtering are both over CBF. In general, CF has better performance in later rounds. And the composite filtering is close to the performance of CF, and in the last three times, composite filtering is more accurate than CF. In this circumstances, composite approach is an improved filtering method compared to simple CBF and CF.

The approach proposed in this paper is not a hybrid one. It is a combined approach by processing CBF and CF respectively in the server and doing combined recommendation in the client. This approach is simple and extendible by integrating other filtering techniques with reasonable weights.

6. CONCLUSION

With the popularity of the Internet, information filtering

technology has been applied more widely in the application of on-line libraries and e-commerce. Especially for researchers, retrieving relevant ones out of numerous on-line papers accurately and efficiently becomes very important. This paper reports a combined filtering approach for solving this problem. The contributions of this work include:

(1) A novel combined filtering approach is proposed based on collaborative filtering and content-based filtering.

(2) Experiments were conducted on a filtering prototype system for Chinese computer science papers.

(3) Preliminary results analysis shows a better performance compared to pure CF and CBF.

But the combined approach also has some limitations. For example it cannot solve the scalability problem of CF. Additionally, due to processing CBF and CF in the server at the same time, and recommending the user items, the computation burden of server is relatively high. So the future work is to improve the performance of the Filtering System by designing optimized algorithm. Additionally, the focus of our future research is to integrate other information technology for better performance and to generate more friendly human machine interface.

7. REFERENCES

[1] N. J. Belkin, W. B. Croft, “Information Filtering and information Retrieval: Two Sides of the Same Coin,” Communications of the ACM, 35(12), 1992,29-38.

[2] P. W. Foltz, S. T. Dumais, “Personalized Information Delivery An Analysis of Information Filtering Methods,” Communications of the ACM, 35(12), 1992,51-60.

[3] V. L. Centeno, C. F. Panadero, et al, “Personalizing Your Electronic Newspaper,” In Proceedings of the 4th Euromedia Conference (WEBTEC), April 1999,26-28.

[4] H.F.Lin, Y.L.Li, et al, “An Information Diffluence Mechanism for Chinese Text Filtering,” Journal of Computer Research and Development, 37(4), 2000,470-476.

[5] P. Resnick, N. Iacovou, et al, “GroupLens: An Open Architecture for Collaborative Filtering of Netnews,” Proceedings of the Conference on Computer Supported Collaborative Work, 1994,175-186.

[6] U. Shardanand, P. Maes, “Social Information Filtering: Algorithms for Automating ‘Word of Mouth’, " Proceedings of the ACM CHI conference, 1995,210-217.

[7] D. Billsus, M. J. Pazzani, “Learning Collaborative Information Filters,” In Proceedings of ICML’98, 1998, 46-53.

[8] M. Balabanovie, Y. Shoham, “Fab: Content-Based, Collaborative Recommendation,” Communications of the ACM, 10(3), 1997,66-72.

[9] M. Claypool, A. Gokhale, et al, and “Combining Content-Based and Collaborative Filters in an Online Newspaper,” In Proceedings of the SIGIR-99 Workshop on Recommender Systems: Algorithms and Evaluation, 1999.

[10] J.H.Zhen, F.Li, “Component Extraction for Automatic Chinese Character Segmentation and Pos Tagging System,” Chinese Journal of Computer Development and Application, 7(14), 2001,11-13.

Method of Measuring the Capacity Flexibility of a Transportation Network

Mingyu Li
Xi'an Research Inst. of Hi-tech HongQing Town
Xi'an, P. R .China, 710025
Email:limingyu800824@126.com
Yiming Bi
Xi'an Research Inst. of Hi-tech HongQing Town
Xi'an, P. R .China, 710025
Email:biym8888@126.com
Rui Zhang
Xi'an Research Inst. of Hi-tech HongQing Town
Xi'an, P. R .China, 710025
Email:ruirui821026@126.com

Abstract

In the light of changes in supply chains and traffic patterns, and in the vulnerability of the network to both natural disasters and terrorist actions, flexibility was gradually desired as a characteristic of transport network. Concepts and techniques of the flexibility of transportation network were described to measure and accommodate the changing demands and traffic systems. It was emphasized on the graph network, which was used to measure the capacity of individual links, and nodes of a network in terms of volume-delay functions with methods already existed. Some approaches are described to measure the flexibility of the network. Applying a penalty to the model when it deviates from the underlying traffic pattern solves the issue of uniqueness of the available capacity. The result is shown to be valuable and applied.

Keywords: network capacity reliability, node capacity, network reserve capacity, flexibility.

1. INFORMATION

With the rapidly increasing and shifting demands for transportation network services, flexibility is increasingly seen as an important characteristic of transportation networks. This paper defines flexibility as the ability of a network to adapt to external changes, while satisfactory network performance should be maintained. Network performance consists of factors, such as capacity, level of service, maintainability and profitability. External changes are uncontrolled conditions, which affect the network. These changes include varied aspects such as the level of demand or use, shifts in spatial traffic patterns, infrastructure loss and degradation. They also include changes in the price and availability of important resources such as fuel, etc. Thus the concept of the network capacity flexibility is defined as the ability of a transport network to accommodate variations or changes in traffic demand that should maintain a satisfactory level of performance.

2. BASIC APPROACHES TO MEASURING FLEXIBILITY

In traditional concept, flexibility is the idea of sensitivity analysis, especially the particular form used in breakeven analysis. The essential idea is to determine whether a productive unit (factory, etc.) will be profitable or not facing an uncertain sales volume of its product. That volume can vary from zero up to the capacity of the production unit. This method of measuring the flexibility is availability because it provides the range of demand over which the network will have satisfactory performance such as a positive profit. However it is difficulty to use this method in the transportation context because the demand of a transportation networks is multi-dimensional, reflecting many O-D pairs, commodities, quality of service, etc. Performance may also be multi-dimensional. However, the fundamental idea of identifying the range of demand over which the network will perform satisfactorily is relevant.

3. ADJUSTED TRAFFIC PATTERN APPROACH TO MEASURING FLEXIBILITY

It is for the adjusted traffic pattern approach to measure flexibility (ADJUSTED) model that it is an estimator of the maximum demand that can be accommodated by the network. It is can estimate how much more traffic volume could added to the MAXC [1,2] estimate for the network by the ADJUSTED model when deviations from the base traffic pattern are permitted. It is meaning that the fraction of total traffic that moves between each OD pair is now a variable. Consequently, there are many more chances to add traffic to the network than the MAXC approach permitted. It includes converting empty container moves to loaded moves and adding traffic to arcs that are not constrained. The ADJUSTED model is more flexible to achieve a maximum cargo flow of traffic on every traffic lane.

In this paper, we add two equations below (equation (1) and equation (2)) to the ADJUSTED model from the MAXC model [3]. It can seek to add as much cargo traffic as possible over by the ADJUSTED model. The base capacity obtained from the MAXC models are as follows:

$$\max x_A = \sum_{r=1}^{n} (Y_r + \overline{y_r}) \tag{1}$$

Subject to arc vehicle and cargo flows:

$$\sum_{p \in Pr} \mu_p f_p = Y_r + \overline{y}_r, \forall r \tag{2}$$

Where two new variables and one new parameter are added:

$\overline{y_r}$: Additional cargo volume on OD pair r (tons/month);

x_A : Estimated maximum system capacity for ADJUSTED model;

Y_r : Base traffic pattern system capacity for each OD pair r;

Therefore, the new model maximizes overall cargo volume, but no longer subject to the base traffic pattern. The resulting capacity is called the adjusted traffic pattern capacity with the format used earlier. The reserve capacity is called the adjusted traffic pattern reserve capacity. As will be shown in table 1, the adjusted pattern capacity is a function of the magnitude of the deviation from the base traffic pattern.

Table 1. Maximum traffic increases by lane allowing deviations from the base traffic pattern and with the shortest path + one node disjoint path set routing option (ADJUSTED model results) (FEUs/month)

lane	1	3	5	7
Base traffic	14720	14892	14720	4370
Base pattern capacity (y_r)	41038	41519	41038	12183
Adjusted pattern capacity (Y_r+y_x)	45006	45527	53553	24748
Adjusted pattern reserve capacity	30286	30635	38833	20378

But the ADJUSTED model also has a problem that virtually always will give a degenerate solution because there are many possible combinations of traffic in individual lanes, and it will yield the same total volume for the system. To solve this problem, we developed various methods to achieve a unique solution. We adopt the minimizing deviation between the additional traffic volume for each lane and the volume, which result were the flows to adhere to the original base unit OD matrix. This is a reasonable objective. Traffic patterns reflect the spatial pattern of production and consumption, and would not change arbitrarily. Some limitation of possible patterns is reasonable and appropriate.

4. PENALTY ADJUSTED MODEL

In this paper, we solve the issue of uniqueness of the ADJUSTED traffic volumes by applying a penalty to the model when it deviates from the underlying traffic pattern. The penalty increases with the deviation increases. In this way, we can constrain the traffic pattern to be as close to the base or underlying one as desired. We try many possible measures of deviation that generally based on standard deviation concepts, but also including other measures such as absolute value of deviation [4].

Equation (3) is the measure of deviation that emerged as preferred is the following:

$$FSD = \sum_{r=1}^{n} \left[\frac{[(Y_r + \overline{y_r}) - (\frac{Y_r}{X_m} \cdot \sum_{r=1}^{n} [Y_r + \overline{Y_r}])]}{X_m} \right]^2 \quad (3)$$

We chose this penalty function because it captures the deviation of the resulting or new traffic pattern from the base traffic pattern but normalized by the base traffic pattern capacity. The numerator of the penalty function is the difference between the new traffic on each OD pair and the traffic that would be on that pair if there were compliance with the base traffic pattern. The denominator is the estimated maximum system capacity from the MAXC Model. Thus this is the fractional deviation of the new pattern from the base traffic pattern. It is squared to avoid the problem of positive and negative deviations canceling one another. The sum of the squared fractional deviations gives the Fractional Squared Deviation (FSD).

The penalty function chosen has other advantages. The deviation measure is fairly easy to understand, and it has definite upper and lower limits. The lower limit is zero, when the new traffic conforms to the base traffic pattern. The upper limit is the square of the additional traffic as a fraction of the MAXC solution.

This deviation measure is added to the ADJUSTED objective function and weighted by beta to make it a penalty function (4). The constraints in the original ADJUSTED model remain the same.

$$\max x_{po} = x_p - \beta \sum_{r=1}^{n} \left[\frac{[(Y_r + \overline{y_r}) - (\frac{Y_r}{X_m} \cdot \sum_{r=1}^{n} [Y_r + \overline{y_r}])]}{X_m} \right]^2 \quad (4)$$

Where:

$$x_p = \sum_{r=1}^{n} [Y_r + \overline{y_r}])$$

x_{po} : Estimated maximum system capacity for the modified ADJUSTED model with fractional squared difference penalty function added [5];

β : Traffic pattern deviation penalty factor;

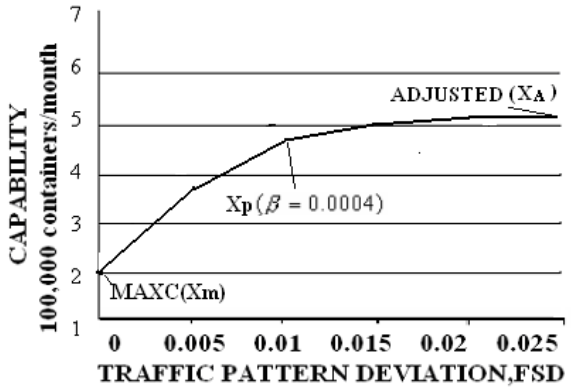


Fig. 1. System capacity versus allowable deviation from the base traffic pattern (FSD) using shortest + one node disjoint path traffic routing option.

Figure. 1. presents the result of using this on the example system (with the shortest path + one node disjoint paths option, as used earlier). It plots system capacity (in units of one-hundred thousand containers/month) versus the traffic pattern deviation measured by FSD. As the deviation increases, the capacity of the system increases. At zero deviation[6], the capacity is that given by base pattern capacity x_M. But as deviations from the base traffic pattern appear, the capacity increases. The increase is steady but at a declining rate, as might be expected. Once the deviation is sufficiently large, the capacity is the maximum value of the adjusted pattern capacity x_A. Above this no further increase in capacity is possible, regardless of the magnitude of the deviation; that is, even complete freedom as to traffic pattern will not enable the network to accommodate more traffic[7].

The curve of Figure. 1. is generated by varying β. When β is small, the estimated capacity and the deviation (FSD) are large. When β is large, the estimated capacity and the deviation is small. For large β, the penalty is so large that, in certain instances, it is better not to add traffic between certain OD pairs and incur the "cost" of increasing the penalty (in the objective function). So, as β is increased to successively larger values, the optimization algorithm ultimately chooses not to ship additional cargo and hence the traffic pattern conforms to the base pattern. Similarly, smaller values of β have almost no effect on the estimated system capacity, as the penalty for deviations from the base pattern are small compared to the gain in objective function from more traffic carried [8,9]. The values for β are determined through experimentation, in this case from 10^{-6} to 10^{-3}.

5. CONCLUSION

This paper defines and formulates the concepts of flexibility for a transportation network. We develop two approaches to measure the capacity (or demand) flexibility. The adjusted traffic pattern approach allows for deviations from the base traffic pattern, starting from the MAXC solution. Thus it provides greater latitude for traffic or demand changes, as from shifts in the location of production or consumption patterns. The measures are applied to a freight transportation network and results show that these measures show promise as means to quantify flexibility. The research also illustrates the increased network capacity flexibility by providing path options significantly [10, 11].

There are many avenues for future research. There are also a number of technical issues related to the measures and the models that should be addressed, such as the multiple optimal solutions issue discussed in the ADJUSTED Model sections. One example is the effect of losses of elements of the system, such as links lost due to natural or manmade disasters (e.g., vulnerability to terrorism). An important exogenous factor is change in the relative cost of resources, or in their availability[12].Another avenue of research would be to consider a broader range of performance measures, not simply the ability of the system to carry traffic at a satisfactory level of service, was done here. Another direction for future work is to quantify the more likely uncontrolled changes that may occur, including variations in demand and traffic patterns. Thus we can produce a better traffic network to accommodate better service.

6. REFERENCES

[1] S.C. Wong, "On the reserve capacity of priority junctions and roundabouts", Transportation Research B, 1996, 30 (6): 441-453.

[2] S.C. Wong, H. Yang, "Reserve capacity of a signal2 controlled road network", Transportation Research B, 1997, 31 (5): 397-402.

[3] A. Chen, H. Yang, et al,"A capacity reliability for networks", Journal of Advanced Transportation, 1999, 33 (2): 183-200.

[4] H. Inoue ,"Traffic equilibrium considering congestion at junctions and its computational solution", Infrastructure Planning Review , 1988 , 6 : 169-176

[5] Bell M G H. "Stochastic user equilibrium assignment in networks with queues". Transportation Research B, 1995, 29 (2): 125-137.

[6] H. Inoue, "Traffic assignment on a network with congested flows". Proceedings of JSCE, 1986, 24: 125-133.

[7] R. E. Allsop, "Estimating the traffic capacity of a signalized road junction". Transp. Res., 6, 245–255.

[8] Arnott, R., de Palma, A., and R. Lindsey, ''Does providing information to drivers reduce traffic congestion?'' Transp. Res., Part A, 25A, 309–318.

[9] T. Basar, G.L. Olsder, "Dynamic on cooperative game theory", 2nd Ed., Academic, London.

[10] Y. E. Ge, "The combined traffic assignment and signal control problem: A game-theoretical approach. PhD dissertation", Tongji University, Shanghai, China.

[11] Koutsopoulos, H.N. Lotan, "Motorist information systems and recurrent traffic congestion: Sensitivity analysis of expected results", Transportation Research Record 1281, Transportation Research Board, Washington, D.C., pp.148–158.

[12] M. J. Maher, P. C. Hughes, "Estimation of the potential benefits from an ATT system using a multiple user class stochastic user equilibrium assignment model", Applications of advanced technologies in transportation engineering, pp.700–704.

IPO Pricing Model based on Game Analysis between Offering Firms in IT Industry and Underwriters

Xusong Xu, Jianmin Sun
Economics and Management School, Wuhan University
Wuhan, Hubei 430072, P.R.China
Email: xuxusong@whu.edu.cn Email: alexsun2000@163.com

ABSTRACT

It is well known that firms in IT industry must issue their shares for the further development when growing large enough scale. Stock market in china, as a new stock market, has stronger speculation and more randomly fluctuant prices than those in any developed countries, which also expresses underpricing clearly. Firms in IT industry shows unique characteristic, which is different from the companies in other industries because of their operation and development pattern. Analyzing the relation among offering firms in IT industry, underwriters and new investor in IPO, this paper deduces a cooperative game model.

Keywords: Game Analysis, Initial Public Offerings (IPO), IT Industry, Offering Firms, and Underwriters.

The price behavior of Initial Public Offerings (IPO hereafter) has long been the focus and perplexity of theorists' research. IPO underpricing refers that the price of new shares is lower than the closing price of first exchange, and high average initial return result from the under pricing behavior. Because offering firms in IT industry own core techniques and all investors expect better foreground than those firms in any other field. In contrast to the long-term return, the abnormally high initial return attracts theorists' extensive attention, and has been confirmed by the market test of almost all the countries.

1. MODEL DEFINITION

Suppose that offering firms in IT industry and underwriters are neutral risk, that one underwriter can only serve one offering firm during a certain period, that offering firms will use their credit resources and have to issue equity to increase their fund, that there is not secondary market, that underwriters must underwrite all stocks having not been sold. Simply put, this process of confirming pricing is exogenous.

Of the two issuing firms, F_1 and F_2, searching for public capital by IPO, F_1 is the smaller one. Of the two underwriters for IPO, U_1 and U_2, supervised by two syndicate and many potential investors, U_1 is the one with lower reputation. Different characteristic of two firms and two underwriters will be listed below.

Suppose $y(k;\theta)$ is NPV of the enterprise while k new capital is injected to support its business growth, $\theta \in [0, 1]$ denotes the characteristic of the enterprise' production plans. Assume that, once θ is known, the asset-pricing model determines the value of enterprise. We omit the detail of the model. For fixed quality θ, $y(k;\theta)$ is strictly increasing and strictly convex in k; for fixed k, $y(k;\theta)$ is strictly increasing in θ and $y(k;\theta)>0$. However, the quality index θ cannot be directly observed and can be defined as a random variable. Its corresponding unconditional distribution function $p(\theta)$ is considered to be public information, θ_i denotes the quality of firm F_i being the unconditional distribution.

Assuming that F_1 has a vector of factors unrelated with its quality index θ_i including the recent gains, the firm's current book value, the relationships between the firm and venture capital. It is assumed that the vector is mapped into a one-dimension variable x. Because $x' > x''$, the distribution of θ under the first-order derivative x' is stochastically dominant over that under the second-order derivative x'', and a higher x stochastically corresponds to a higher θ [1]. Therefore, the conditions $x_1<x_2$ defines the difference between F_1 (characteristic x_1) and F_2 (characteristic x_2).

2. GAME ANALYSIS

The problem that issuers should confront is that the public cannot accept their information easily. For the investor, the reputation information x_i of that firm depends on the reputation of the underwriter [2]. The same value of the x_i given by different underwriters is the realization of two different random variables. We take $z_j \in [z_0, \bar{z}]$ as the reputation of the underwriter U_j. The larger z_j is, the higher reputation of the underwriter U_j is, and the more credible the guarantee given by underwriter to investors is.

Suppose $P(\theta \mid x, z)$ has a distribution with condition x and z. x is the guarantee given by the underwriter and z is the reputation of underwriter. Suppose that the condition distribution is differentiable to x and z. Then the properties of the two firms and two underwriters are as follows:

$$P_x(\theta|x,z) \equiv \frac{\partial}{\partial x} P(\theta|x,z) < 0 \quad (1)$$

$$Pz(\theta|x,z) \equiv \frac{\partial}{\partial z} P(\theta|x,z) < 0 \quad (2)$$

For the sake of completeness, we take z_0 as the reputation level of the underwriter when they don't carry out their guarantees. So $P(\theta \mid x, z)$ is the distribution function without guarantees of underwriter. For (1), the factors concerning about enterprise quality make investors observe nonexistent reputation problem.

We can define the follow consistent conditions in the relativity analysis about distribute game.

$$P_{xz}(\theta|x,z) \equiv \frac{\partial^2}{\partial x \partial z} P(\theta|x,z) < 0 \quad (3)$$

In order to explain this condition, take $\bar{\theta}_{ij} \equiv E(\theta|x_i, z_j)$ as the expected value θ of F_i when x_i is ensured by guarantee U_j whose reputation value is z_j. We can get:

$$\bar{\theta}_{12} - \bar{\theta}_{11} < \bar{\theta}_{22} - \bar{\theta}_{21},\ x_1 < x_2, z_1 < z_2 \qquad (4)$$

The model focuses on the results of conference between issuer and underwriter. The demand of investors on new shares is a random variable. Suppose that the firm having $\bar{q}_i$ old shares, issues q_i new shares, and the issuing price is p_i. The underwriter takes underwriting method and a fixed price difference $\delta \in$ (0, 1). So they can get compensation of $\delta p_i q_i$. The true investment the firm acquires is k_i=(1-δ) $p_i q_i$. Each investor makes choices based on the information x_i ensured by U_j and other information he possesses. The amount of gross demand of the new shares of the F_i underwritten by U_j is s_{ij}. Suppose that the gross demand gathers all the related information about θ,then the market value of the firm is

$$y_{ij}(k_i | x_i, z_j, s_{ij}) = E[y(k_i, \theta_i) | x_i, z_j, s_{ij}] \qquad (5)$$

Considering the conditional distribution $P(\theta | x_i, z_j, s_{ij})$, the market value of the firm is

$$p_i^a(k_i) = y_{ij}(k_i | x_i, z_j, s_{ij}) / (\bar{q}_i + q_i) \qquad (6)$$

The gross demand s_{ij} is known by issuer and underwriter, then the expected value of firm F_i under conditions x_i, z_j is

$$y_{ij}(k_i) = E[y_{ij}(k_i | x_i, z_j, s_{ij}) | x_i, z_j] \qquad (7)$$

$y_{ij}(k_i)$ does not include x_i, z_j which are known by F_i and U_j. As we can see, the work to simplify the form of $y_{ij}(k_i | x_i, z_j, s_{ij})$ is not so important that it is inefficient to the market value and value expectation.

We begin with considering an issuing firm $F_i \in \{F_1, F_2\}$ and an underwriter $U_j \in \{U_1, U_2\}$. Before IPO, the firms holding $\bar{q}_i$ issue its shares. For a fixed δ,the key variable of issuer and underwriter is q_i and p_i. The value of issued stock right will increase as follows

$$f_{ij}(p_i; k_i) = \frac{\bar{q}_i}{\bar{q}_i + k_i / p_i(1-\delta)} y_{ij}(k_i) - y_{i0}(0)$$

$$i = 1,2; j = 1,2 \qquad (8)$$

Here $y_{i0} = \int y(0, \theta) dP(\theta | x_i, z_0)$, $f_{ij}(p_i, k_i)$ is increasing p_i if k_i is fixed. The expectation income of the investor is

$$u_{ij}(p_i; k_i) = \frac{q_i}{\bar{q}_i + q_i} y_{ij}(k_i) - \frac{k_i}{1-\delta}$$

$$= \frac{k_i / p_i(1-\delta)}{\bar{q}_i + k_i / p_i(1-\delta)} y_{ij}(k_i) - \frac{k_i}{1-\delta}$$

$$i = 1,2;\ j = 1,2 \qquad (9)$$

When the issuing price is fixed, issuers buy all new shares and sell them to investors. They have to buy all the unsold shares [3,4]. The lower p_i is, the easier the sale will be. So the total expected return of underwriters and investors is

$$\hat{u}_{ij}(p_i; k_i) = \frac{q_i}{\bar{q}_i + q_i} y_{ij}(k_i) - k_i$$

$$= \frac{k_i / p_i(1-\delta)}{\bar{q}_i + k_i / p_i(1-\delta)} y_{ij}(k_i) - k_i$$

$$i = 1,2;\ j = 1 \qquad (10)$$

The cost of underwriter is k_i = （1-δ） $p_i q_i$ and they get q_i shares. The expected return of shares is $[q_i / (\bar{q}_i + q_i)] y_{ij}(k_i)$. To fixed k_i , $\hat{u}_{ij}(p_i; k_i)$ is decreasing in p_i .

Therefore, once total amount of new capital k_i come to a decision, the benefits of firms and underwriters are opposite. There must be trade-off between risk-neutral underwriters and investors. We also have $f_{ij}(p_i, k_i) \geq 0$, $\hat{u}_{ij}(p_i, k_i) \geq 0$ and can get a (p_i, k_i) based on individual-rational.

3. PRICING MODEL

Suppose that the future of the issuing firm and the investors' attitude are known. They choose to increase their shared benefits. Then for a fixed portfolio, firm F_i and underwriter U_j, increment of the best issue quantity will maximize expected NPV when they have same information.

We emphasize that the optimal quantity of the issue also depends on underwriter [5]. That is

$$k_{ij}^* = \arg\max_k [y_{ij}(k) - y_{i0} - k],$$

Let: $$v_{ij} = \int y(k_{ij}^*, \theta) dP(\theta | x_i, z_j) - y_{i0}(0) - k_{ij}^* / (1-\delta) \qquad (11)$$

The V_{ij} of a firm is positive. The x of other firms is too low to get a higher expected NPV than inflation cost.

We take the same definition,

$$v_{i0} = \max\{0, \max_k [\int y(k_{ij}^*, \theta) dP(\theta | x_i, z_0) - y_{i0}(0) - k/(1-\delta)]\} \qquad (12)$$

Suppose that the firm maximizes their NPV of IPO and there is a comparable issuing cost δ . Clearly, $v_{10} \leq v_{20}$.

According to above analysis, we suppose the maximum is feasible. $k_i = k_{ij}^*$ and δ is fixed. The underwriter negotiates with firm as an investor. An individual-rational trade-off p_i is the trade-off between $f_{ij}(p_i, k_i) \geq 0$ and $\hat{u}_{ij}(p_i, k_i) \geq 0$ .We find $v_{11} < v_{12} < v_{22}$, $v_{11} < v_{21} < v_{22}$, $v_{10} < v_{11}$ and $v_{20} < v_{21}$ (i); $v_{20} - v_{10} < v_{21} - v_{11} < v_{22} - v_{12}$ (ii) describing the relation between v_{ij} and s. $v_{21} - v_{11} < v_{22} - v_{12}$ means it is better to contact with big firms than small ones. Underwriter with higher reputation is better than that with lower one.

Let $u_i(p_i^o) = 0$, $p_i^o = \frac{y_{ij}(k_{ij}^*)(1-\delta) - k_{ij}^*}{\bar{q}_i(1-\delta)}$,

$$q_i^o = \frac{k_{ij}^*}{p_i^o(1-\delta)} = \frac{\bar{q}_i k_{ij}^*}{y_{ij}(k_{ij}^*)(1-\delta) - k_{ij}^*}$$

The expected price per share is

$$E p_i^o(k_{ij}^*) = \frac{y_{ij}(k_{ij}^*)}{\bar{q}_i y_{ij}(k_{ij}^*)(1-\delta) / [y_{ij}(k_{ij}^*)(1-\delta) - k_{ij}^*]} = \frac{y_{ij}(k_{ij}^*)(1-\delta) - k_{ij}^*}{\bar{q}_i(1-\delta)} = p_i^o$$

That is, the expected price per share is equal to the issue price.

Under the assumption that issuer and underwriter are risk neutral, it is called under-pricing if the issuing price is lower than the expected price per share [6]. If the underwriter is risk neutral, the only way to avoid underpricing is that the issuer has complete pricing ability when the underwriter is

also the investor. When the issue price p_i satisfies $U_j(p_i)=0$, there is no underpricing in IPO with individual-rational trade-off price. So underpricing happens in IPO as the issuer lacks of pricing ability. Though the outcome of the negotiation will be changed because of the experience and technology of the two parties' negotiators, there are still some principles and social standards to supervise negotiation. On the other hand, the competition between underwriters and offering firms in IT industry is the most important driven factor among all factors that influence bargaining results.

4. CONCLUSION

Information asymmetry between issuers and investors results in an inadequate competitive IPO market. Investment banks will get more profits for their own good reputation. Underwriters would like to negotiate issuing price and profit which issuing firms gain from new investors with them. In conclusion, IPO under pricing is a fruit of this negotiation.

In a word, the average underpricing of IPO is not wealth transfer from original equity owners to new owners but reflects re-allocation of the increase of the surplus equity between original and new equity owners; The cooperation of investors and underwriters intents to deplete absolute monopolization of offering firms in IT industry; Underpricing degree depends on the bargaining ability of offering firms in IT industry to underwriters in the whole IPO market.

5. REFERENCES

[1] Ritter, R. Jay, "Initial public offerings [J]", Contemporary Finance Digest, 1989, 2:5-30.
[2] R.G. Ibbotson, "Initial public offerings [J]", "Journal of Applied Corporate Finance", 1988, 1: 37-45.
[3] T. Loughran, "Initial public offerings: international insights [J]", Pacific-Basin Finance Journal 1994, 2: 165-199.
[4] N. Barber, A.R. Shleifer, "A model of investor sentiment [J]", Journal of Financial Economics 1998, 49:307-345.
[5] H.F. Leland, "Information asymmetries, financial structure, and financial intermediation [J]", Journal of Finance, 1997, 52:37-387.
[6] D. Baron, "A model of the demand for investment banking advising and distribution services for new issues [J]", Journal of Finance, 1982,2:955-976.

Real-Time Simulation of Control System for Tall Building with Podium Structure

Jing Chen
College of Automation, Wuhan University of Technology
Wuhan, Hubei, P.R. China
Email: cje@mail.whut.edu.cn or Jingchen680@163.com
And
Chaozhi Ren
College of Automation, Wuhan University of Technology
Wuhan, Hubei, P.R. China
Email: zhaoyangstar@sohu.com or ren_hk@163.com

ABSTRACT

To introduce the software and hardware resources of dSPACE real-time simulation system, and takes a slender 12-story building model and a relatively stiff 3-story building model as control target, constructs the simulation model based on the graphic modelling tool Simulink , and gives out the methods to achieve real-time simulation based on dSPACE simulation environment. Experimental investigations were performed in the Structural Dynamics Laboratory in the Hong Kong Polytechnic University, which greatly reduced the development time.

Keywords: dSPACE, structural vibration control, Matlab.

1. INTRODUCTION

Structural vibration control is the burgeoning subject that developed in 30 years ago, synthesizing the newest research of many subjects such as control theory, computer technology, new material and structural vibration theory. It is a new field of automatic application research, and now is still in developing stage, which can effectively reduce the response and injury of the powerful winds and earthquakes, increase the capacity of withstanding vibration and disasters. As for the control is based primarily on the actual structure, non-structural components and the changes of the mass which influence the calculation model very much. Therefore, the research does not rely on the precise calculation, the simple control algorithms are of great significance.

According to the features of the tall buildings linked by MR damper, abandons the search for accurate mathematical control model, develops a semi-active logic control method [1]. In order to testify its effectiveness, and promote the application, it is very necessary for real-time simulation.

DSPACE system currently has been widely applied to robotics, aerospace, automobile, engine, power train, industrial fields, and etc; but few application introduction on the field of structural vibration control. This paper aims to introduce a new type of real-time simulation system, which takes a slender 12-story building model and a relatively stiff 3-story building model as structural model and uses simulink to construct the simulation models, gives out the methods to achieve real-time simulation and testing results based on dSPACE simulation environment.

2. REAL-TIME SIMULATION CONTROL SYSTEM

Real-time simulation control system is made up of signal detection and transducers, A/D converters, signal controllers, D/A converters, operation unit and computers, and the constitute diagram is shown in Fig. 1.

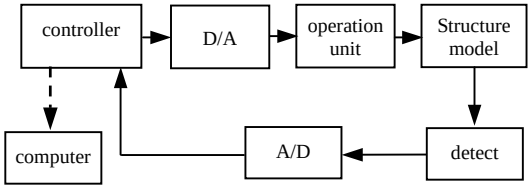


Fig. 1. Simulation Control System Constitute

A slender 12-story building model and a relatively stiff 3-story building model linked by MR dampers is the control target[2] and chose dSPACE real-time simulation system as hardware and software platform, which is a control system based on MATLAB/Simulink and achieved completely seamless connectivity with MATLAB / Simulink / RTW (Real- Time-Workshop).

DSPACE hardware system [3] is made up of DS1005 PPC processor boards, DS2003DS, DS2102, DS4003 and their interfaces, as is shown in table 1.

Table 1. dSPACE Hardware Parameter

Number	Model number	Specification
1	DS1005	PowerPC 750，480MHz
2	DS2003	32 A/D
3	DS2102	6 D/A
4	DS4003	96 I/O
5	PX10	10-slot box （connect to Host）
6	PHS-Cable	Electric cable
8	CP2102	DS2102 port panel
9	CP4003	DS4003 port panel
10	CP-Desk	Wiring desk

The hardware of dSPACE real-time simulation system, including processors, I/O, and etc, has high speed coping capability. And the software environment could achieve following functions, such as convenient automatic code generation, downloading and testing/debugging. With this

system, it can commendably solve two major application problems, which are RCP (Rapid Control Prototyping) and HILS (Hardware in the Loop simulation).

The aim of dSPACE system is completely separating the real-time system and user interface system. Processor boards and various I/O boards have high operational speed, and they connect through PHS bus, which is designed for real-time application and always open to I/O and processor. There is no waiting time existed as other CPU bus systems has, and not like the bus systems which has internal software but transmits through external agreements, therefore ensures the I/O visiting time. The interface between dSPACE board and computer is connected by ISA bus or Ethernet. The computer is able to visit all the memories in the dSPACE processor board, and realizes the data transmission between dSPACE hardware and PC, which works as a smart terminal and a visualization control platform.

3. DSPACE DEVELOPMENT PROCESS

DSPACE provides a computer-aided control system for RCP and HILS, which is named CDP (Control Development Package). The development processes are depicted in fig.2.

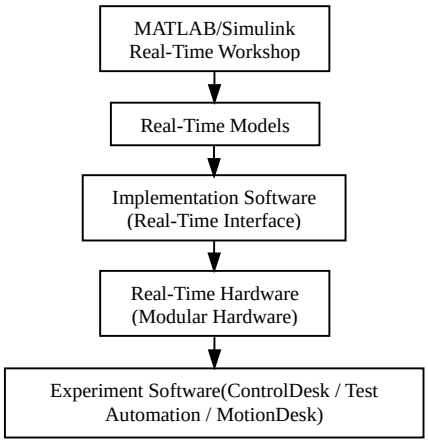


Fig. 2. DSPACE Development

MATLAB can be used for analysis, design, optimization and off-line data processing, and Simulink has the function of off-line simulation which based on block diagram, and programming (block diagram); RTW can transfer the program into C code. The Real-Time Models, including A/D converter, D/A converter, I/O interface, can be downloaded and edited in the environment of simulink; Real-Time Interface(RTI) transfer the C code to machine code which ran in processors of dSPACE. Real-Time hardware indicates the whole hardware system of dSPACE, and the Experiment Software is used to operate closed-loop experiments (automatic/manual), online detection display, control parameters set and debugging.

Development methods: system modeling, analysis, off-Line simulation, real-time simulation.

(1) The theory model (linear or nonlinear equations) of objects, control algorithms and the original control scheme are designed by Simulink.

(2) Simulink the control scheme off-line, optimize the control algorithm and parameters in order to guide the experiment.

(3) Drag the A/D converter, or D/A converter, or I/O interface, to the block diagram, which is needed for real-time testing, from the library of RTI, and set the parameters, in the aim of using real experiment control objects to replace the theory model in (1).

(4) Click the build command in the Simulink/RTW, automatically realize C code generation, translation, connecting and downloading of the target DSP system.

(5) Establish a new experiment; cover it in Control Desk, which is a testing tool software environment. First, design the platform (*.lay) as required; second, connect the platform with available variables from variable file (*.trc) generated in the (4), at last, link the object file (*.ppc) with hardware system.

(6) Run. Modify the parameters online according to the results. Return to (1) if the control algorithm should be changed, repeat the above process.

4. THE REALIZATION OF STRUCTURAL CONTROL SIMULATION SYSTEM

Configuration

This paper applies the rapid control model of dSPACE system, which is not only a development testing platform, but also a practical high-speed and high accuracy controller, to design a slender 12-story building model and a relatively stiff 3-story building structural model system. The real-time hardware system could separate from PC host when the control algorithms are successfully certified, and run as an independent structural controller. Structural measurements

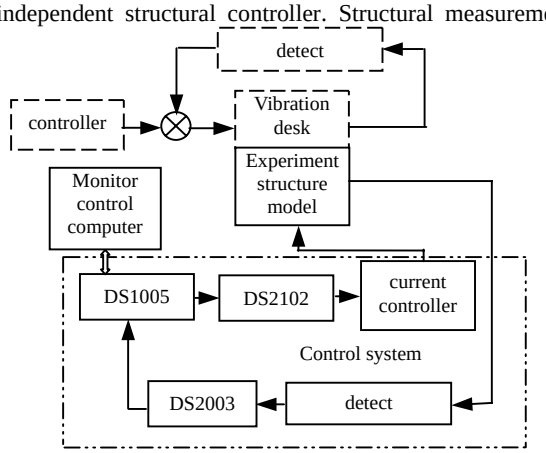


Fig. 3 Control System Diagram

and control systems are depicted in Fig. 3.

Acceleration sensors can be used to measure structural response, an acceleration sensor is lying on the vibration table, and respectively, each lies one acceleration sensor on the 1st, 2nd, 3rd, 4th, 6th, 8th, 10th, 12th floor of the high building and the 1st, 2nd, 3rd of the relatively stiff 3-story building. The sensors and amplifiers (2635) which were made in Denmark are used to detect and transmit. RD-1097-01X dampers are the chosen MR dampers produced in the Lord Corporation of American, and the continuous input current is 0.5A as well as the non-continuous input current is 1.0A, piston movement distance is 50mm, and the dampers will produce the most vigorous over 100N when the input current is 1.0A. The current controller of MR dampers is RD-3002-03 controller of the Lord Company.

Construct Rapid Control Model

DSPACE company has designed different block diagram models for different resources of real-time hardware, which can be used as the models of simulink. Use the model to build robot rapid control model diagram, take the semi-active logic control of the slender 12-story building and a relatively stiff 3-story building as an example, and the control model is depicted in fig.4.

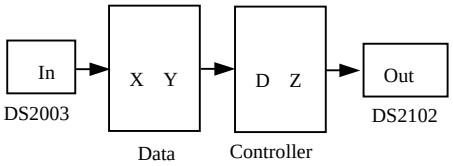


Fig. 4. Fast Structure Control Model

The signal detected by sensors will be transfer from analog to digital after amplifying and filtering through DS2003 module, the result of the conversion is the value between 0~1; Data module completes project value (EV) conversion and signal conversion. The function of the project value is to change the result of A/D converter to real value, and the transferring equation is as follows:

$$EV = V_{ra} / Unit \qquad (Project\ values) \qquad (1)$$

In this equation, V_{ra} is the range of A/D converter (+5000mV or +10000mV), $Unit$ is engineering unit (mV/ project value); Controller module is the core of the design, where only semi-active logic control is edited by simulink; DS2102 module completes the digital/analog conversion of control value according to the relations between the input and the output of current controller.

5. RUN & TEST

The code run in the dSPACE real-time controller after complying and downloading, the ControlDesk provided by dSPACE can be used to real-time watching, modify the control parameters and manage the testing process. Using ControlDesk, it is easy to realize diagram management for real-time hardware, establish user virtual devices, manage the variables and parameters visually, and automatically control the process. Only through the image of memory, visit the parameters and the result variables of the process of example, and not to stop the process, keep the real-time control of the system.

The operation platform primary completes the choice of control methods, simulation control, start/stop of data acquisition, parameters set, operation display, and so on. Take semi-active logic control as an example to design the operation platform, as is shown in fig.5.

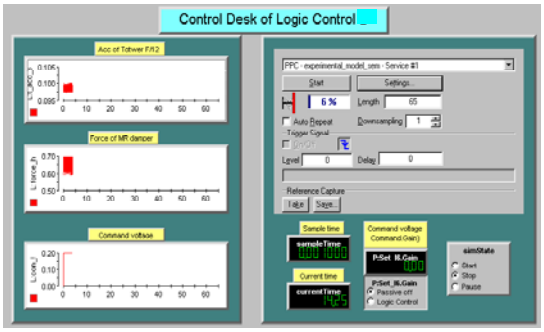


Fig. 5. Semi-active logic control operation platform

Under the EI-Centro earthquake wave, of which the amplitude value is 0.13g and the time compression ratio is 1/3, carry on the vibration table experiment on the structure, the time and displacement curve of the top layer are shown in fig.6.

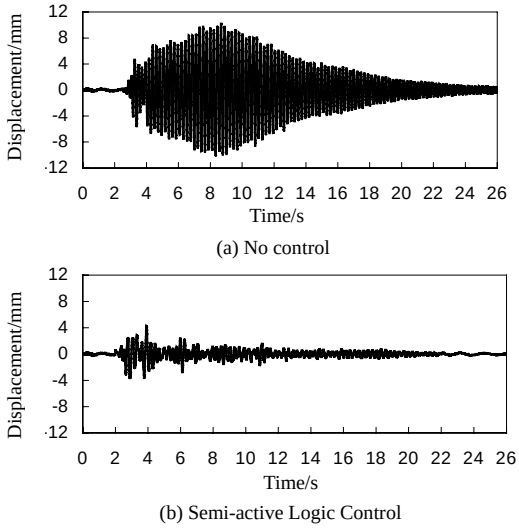


Fig. 6. Time histories of displacement responses at the top floor of the 12-story building under the excitation

6. CONCLUSION

Following conclusions are based on the former study and test:

(1) The software and hardware of dSPACE system use the modular design method, which is much reliable and possible to facilitate quickly construct the rapid control prototype through the integrated development & debugging platform, complete the design and debug of the control system, save a great of time.

(2) The seamless connection of dSPACE system and MA -TLAB has enabled the general MATLAB user to be possible with ease to grasp dSPACE. Thus it is convenient to transfer from the non-real-time to the real-time analysis and design.

(3) The structure controller design is a quite complex work, often needs to compare many different control algorithms and adjust the controlled variables ceaselessly. The rapid control prototype has simple structure and convenient adjustment, extremely suitably for study on the structure control algorithm.

REFERENCES

[1] J. Chen, W.L. Qu, et al, Structural Control Based on Panboolean Algebra. The 6th World Multi-Conference on Systemics, Cybernetics and Informatics. July 14-18, 2002 Orlando, Florida, USA. pp.199-202. ISNB: 980-07-8150-1

[2] J. Chen, W.L. Qu, et al, “Semi-active Logic Control of Tall Buildings with Podium Structure connected to MR Dampers”, Earthquake Engineering and Engineering vibration, in Chinese, 2003. 23(2): 150-153

[3] DSPACE. Real-Time Interface Implementation Guide (Version 6.2). [Z] Paderbom: dSPACE GmbH, 2003

Nonlinear Finite Element Buckling Analysis of Rectangular Reinforced Concrete long Columns Confined with Carbon Fiber Reinforced Plastic Sheets under Axial Compression *

Qingxin Ren, Tingguo Chen, Chengkui Huang
State Key Laboratory of Coastal and Offshore Engineering, Dalian University of Technology, Dalian 116024, China
Email: renqx_007@163.com, chentg@dlut.edu.cn, huangck@dlut.edu.cn
And
Yanhua Liu
College of Water Resource, Shenyang Agricultural University, Shenyang 110161, China
Email: xmliuyh@163.com

ABSTRACT

Owing to some problems in their stability of building structures in service, the general software ANSYS is adopted to establish the finite element models of rectangular reinforced concrete long columns confined with CFRP sheets under axial compression and used to analyze the influence of CFRP sheets to these structural buckling loads. Height of columns, layers of CFRP sheets, concrete grade, longitudinal rebar diameter, stirrup diameter and stirrup spacing are some factors considered. Based on the nonlinear finite element buckling analysis of these columns, some major conclusions are summarized as follows: (1) The enhanced level of buckling load descends with columns height's increasing; (2) The buckling load rises along with the increasing number of layers on a nonlinear manner; (3) The buckling load's increment is hardly dependent of concrete grade, longitudinal rebar diameter, stirrup diameter and stirrup spacing; (4) The improvement is not very obvious to use CFRP sheets to strengthen the structure when it's stability goes wrong. Influences of CFRP sheets to rectangular reinforced concrete long columns under axial compression can be found by analyzing the stability of these columns, and the analysis results are available for the strengthening of some structures in practice.

Keywords: Nonlinear finite element, stability, reinforced concrete long column, carbon fiber reinforced plastic (CFRP), strengthening, axial compression.

1. INTRODUCTION

Advanced composites such as CFRP sheets have received great attention as materials for choice in repair and strengthening projects. The area of composites in construction, and in particular for strengthening, has been one of the fastest growing new areas within civil engineering during the last ten years [1]. CFRP sheets have a lot of advantages comparing to traditional construction materials such as steel, wood and concrete. CFRP sheets offer excellent corrosion resistance to environmental agents as well as the advantages of high stiffness-to-weight and strength-to-weight ratios, Perhaps the biggest advantage of CFRP sheets is tailorability. Reinforcement can be arranged according to the loading conditions so that a CFRP structure or component can be optimized for performance. Many researchers have found that CFRP sheets are efficient, reliable and cost-effective in strengthening short columns [2, 3]. Whereas, the documents of researches on long columns are few. In order to understand the characteristic of reinforced concrete long columns confined with CFRP sheets under axial compression, this paper presents an analytical model using finite element method (FEM) and analyses the influence of CFRP sheets to these structural buckling loads.

2. FINITE ELEMENT MODELING

A three-dimensional (3D) finite element solid model of rectangular reinforced concrete long column confined with CFRP sheets is developed in ANSYS [4, 5]. The rectangular reinforced concrete long column is modeled as 3D solid (SOLID65). In concrete applications, for example, the solid capability of the element may be used to model the concrete while the rebar capability is available for modeling reinforcement behavior. The element is defined by eight nodes having three degrees of freedom at each node: translations in the nodal x, y, and z directions. The most important aspect of this concrete element is the treatment of nonlinear material properties. The concrete is capable of cracking (in three orthogonal directions), crushing, plastic deformation and creep.

A membrane shell element, SHELL41, is used to model the CFRP sheets. SHELL41 is a 3D element having membrane (in-plane) stiffness but no bending (out-of-plane) stiffness. It is intended for shell structures where bending of the elements is of secondary importance. The element has the same degrees of freedom as SOLID65 at each node. It has variable thickness, stress stiffening, large deflection, and a cloth option.

The translational degrees of freedom of all nodes are coupled, the foundations of the column are modeled as fixed end supports. Fig.1. shows the full 3D view of the finite element model of the rectangular reinforced concrete column confined with CFRP sheets under axial compression [6].

3. NUMERICAL RESULTS

The model proposed here has been implemented in the ANSYS program for the nonlinear and buckling analysis of

*Project (Nos. LP0402) supported by the State Key Laboratory of Coastal and Offshore Engineering of Dalian University of Technology of China.

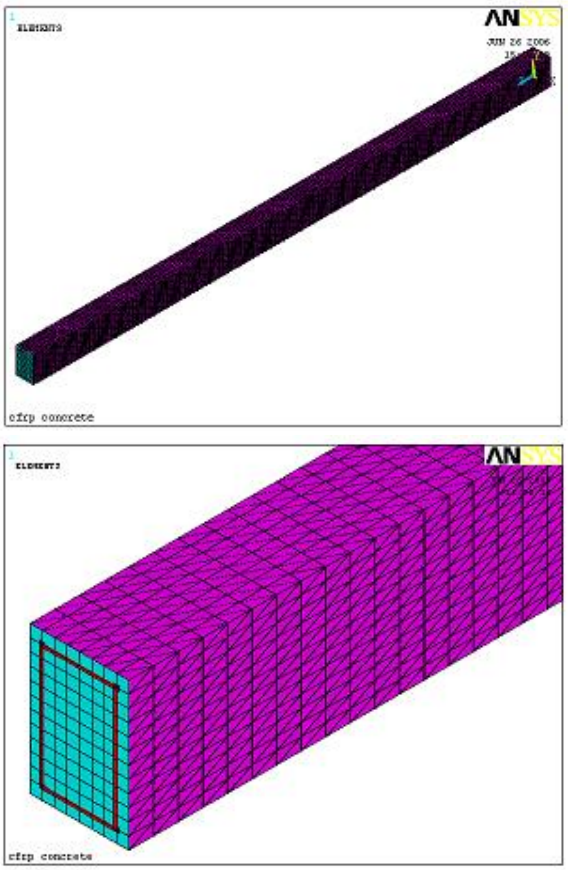


Fig. 1. A 3D Finite element model of reinforced concrete column strengthened with CFRP sheets

rectangular reinforced concrete column confined with CFRP sheets under axial compression. Some factors affecting the collapsing load are considered, such as height of reinforced concrete columns, layers of CFRP sheets, concrete grade, longitudinal rebar diameter, stirrup diameter and stirrup spacing. Nonlinear buckling analysis provides the critical buckling load, and the results are presented in Table 1.-Table 6.. In Table.1., the concrete grade is C30[7], the sectional dimension is 200mm×300mm, each layer of CFRP sheets is 0.167mm thickness, longitudinal rebar diameter is 12mm, stirrup diameter is 6mm, stirrup spacing is 100mm, and the thickness of covering layer is 25mm. In other tables, the height of columns is 6000mm, only one factor is variable, other factors are same as Table 1..

Results of buckling load varying with the height of columns are presented in Table 1.. The buckling load increment is descending while the height of columns is increasing. As visible in Table 2., the buckling load increment is increasing on a nonlinear manner with the increasing layers of CFRP sheets. Other results are presented in Table 3.- Table 6., from these tables, it is clear that the buckling load increment is almost independent of concrete grade, longitudinal rebar diameter, stirrup diameter and stirrup spacing.

4. CONCLUSIONS

Based on the numerical results from the analysis, the following conclusions may be drawn:

(1) The enhanced level of buckling load by CFRP sheets is descending with the increasing of columns' height;

Table 1. Results of buckling load varying with the height of columns

Height of columns(mm)	Buckling load (kN)		
	Concrete(1)	One sheet (2)	(2)-(1)
5000	645.59	646.59	1.00
6000	448.42	449.12	0.70
7000	329.49	330.01	0.52
8000	252.28	252.67	0.39

Table 2. Results of buckling load varying with the layers of sheets

Layers of sheets	Buckling load (kN)		
	Concrete(1)	Sheets(2)	(2)-(1)
0	448.42	-	-
1	-	449.12	0.70
2	-	449.76	1.34
3	-	450.36	1.94

Table 3. Results of buckling load varying with concrete grade

Concrete grade	Buckling load (kN)		
	Concrete(1)	One sheet(2)	(2)-(1)
C20	386.70	387.39	0.69
C30	448.42	449.12	0.70
C40	482.70	483.41	0.70
C50	510.13	510.85	0.71

Table 4. Results of buckling load varying with longitudinal rebar diameter

Longitudinal rebar diameter(mm)	Buckling load (kN)		
	Concrete(1)	One sheet(2)	(2)-(1)
12	448.42	449.12	0.70
14	461.60	462.31	0.70
16	476.87	477.57	0.70
18	494.12	494.82	0.70

Table 5. Results of buckling load varying with stirrup diameter

Stirrup diameter(mm)	Buckling load (kN)		
	Concrete(1)	One sheet(2)	(2)-(1)
6	448.42	449.12	0.70
8	448.76	449.44	0.68
10	449.15	449.81	0.66
12	449.58	450.22	0.64

Table 6. Results of buckling load varying with stirrup spacing

Stirrup spacing(mm)	Buckling load (kN)		
	Concrete(1)	One sheet(2)	(2)-(1)
50	448.90	449.58	0.68
100	448.42	449.12	0.70
150	448.25	448.97	0.71
200	448.18	448.89	0.71

(2) The buckling load increment is rising along with the increasing layers of CFRP sheets on a nonlinear manner;

(3) The buckling load increment is almost independent of concrete grade, longitudinal rebar diameter, stirrup diameter and stirrup spacing;

(4) The improvement is not very obvious to use CFRP sheets to strengthen the structure under axial compression when it's stability goes wrong.

Influences of CFRP sheets to rectangular reinforced columns can be found by analyzing the stability of these columns and the analysis results are available for the strengthening of some concrete structures in practice.

5. REFERENCES

[1] T. Björn, “Strengthening of beams by plate bonding [J]”, Mater Civil Engineering, 1997, 9(4): 6-212.

[2] T. Zhao, J.M. Zhang, et al, “Experimental study on the application of continuous carbon fiber sheet to improve the ductility of reinforced concrete short columns [J]”, World Information on Earthquake Engineering (in Chinese), 2001, 17(1):82-86.

[3] Z.X. Li, C.X. Xu, et al, “Experimental research on seismic behavior of reinforced concrete short columns strengthened with carbon fiber sheets [J]”, Journal of Building Structures (in Chinese), 2002, 23(6):41-48.

[4] H. Liu, Y.H. Zhao, “A finite element study on the load carrying capacity of CFRP strengthened concrete columns [J]”, Journal of Shenyang Architecture and Civil Engineering University (Natural Science), 2002, 18(3):161-163.

[5] K. Damian, M. Thomas, Solomon Yim, “Finite element modeling of reinforced concrete structures strengthened with FRP laminates (final report)”, Oregon Department of Transportation , 2001.

[6] S.S. Wang, H.B. Liu, et al, “Finite element nonlinear analysis for the flexural strengthening of carbon fiber reinforced concrete beams [J]”,Journal of Highway and Transportation Research and Development (in Chinese), 2005 , 22(4):61-64.

[7] China Academy of building research, Design Code for Concrete Structures (GB50010-2002) [S], China Building Industry Press, 2002.

A Cooperative Design Platform for Architectural System Based on Network

Shanshan Yu [1] Jindian Su [2] Xiaoxia Xin [1]
[1.] Dept. of Hospital Information Management, The information and Network Center, Sun Yat-sen University, Guangzhou 510080, China
[2.] Dept. of Computer Science and Engineering, South China Univ. of Tech., Guangzhou 510641, China
Email：shanshan_ssyu@hotmail.com, Brisk_su@hotmail.com

ABSTRACT

Based on Web and CSCW (Computer Supported Cooperative Work) technologies, a cooperative design platform for architectural analog system for heat, electricity and cold supplies and evaluation was proposed to support those remote and distributive design tasks effectively. We mainly focused on the architectural design and implementation methods, and also presented the system structure and pointed out some key problems in detail. This platform can help people in different places join the project effectively so that they can allocate resources reasonably, strengthen the communication, decrease the cost and improve their efficiency.

Keywords: Cooperative Design, CSCW, Analog System, Conflict Control, Architectural Design.

1. INTRODUCTION

Architectural design is a kind of cooperative work involving many subjects. Beginning with users' requirements, architects produce an initiative solution, and structure engineers finish the structure design. Then, engineers who take charge of equipments begin to design systems for heat, electricity and cold supplies. Finally, all solutions are analyzed and controlled in term of costs. This multi-subject cooperation is not only a simple ordinary workflow, but also needs iterative modification and cooperation to obtain the optimized design to satisfy users.

In fact, being confined by many factors, the actual architectural design process doesn't include the idea of cooperative design, which results in many collisions of design drawings and solutions. Usually, engineers have to change them so frequently that the quality of the project is badly affected. Nowadays, with the actual development of international cooperation, architectural design is no longer limited to some design studio or a small group, but becomes complicated and cooperative in various fields. Under this condition, it is of great importance to build an architectural design model and its supporting design platform.

Using CSCW (Computer Supported Cooperative Work) to build a cooperative work platform based on network, experiments, designs and supervisions can provide a networked cooperative environment for people in different places and at different time. In this way, this platform can help workers and experts in each place join the project fully [1]. They can allocate resources reasonably and strengthen the communication and cooperation, and thus decrease the cost greatly and improve the efficiency.

Because of its characteristics of colony, interaction and distribution, the process of architectural design can adapts to CSCW. For these reasons, a cooperative design platform of architectural system for heat, electricity and cold supplies based on network is proposed in this paper.

2. FUNCTIONS OF COOPERATIVE DESIGN PLATFORM

This cooperative design platform of architectural system composed of heat, electricity and cold supplies includes the following functions:

1) Integrating an analog system with heat, electricity and cold supplies. Each member can design and analyze on the analog system like using a traditional analog simulation.

2) System users can analyze the project documents created by the analog system and give their advice, so that the design process can be controlled synchronously.

3) Mechanism of data description, exchange and share. This is one of the most important factors in the cooperative solution of architectural design.

4) Promulgating up-to-date messages about the design and the review. It can harmonize the whole process of the design activity and supply some message communication tools such as message exchange and file transfer.

5) System management. It consists of user rights' settings (deviser, supervisor, browser and so on), and all kinds of strategic management during the design process.

6) Supporting applications in the heterogeneous computing environment**.**

3. FRAME DESIGN

This design platform should have cross-platform ability, so that it can run on most of the current platforms. Java is the programming language we choose for this cooperative platform. Its platform independence characteristic shields actual requirements of the running environment and various java applications can run on it only if it supports the Java Virtual Machine. Moreover, it also has many other characteristics such as distribution, interpretation and transplantable feature [2].

The system architecture, as what is described in Fig. 1., has three layers: the representation layer, the business logic layer and the data layer.

Representation Layer： It includes an analog subsystem on the client and some client applications based on the browser. In detail, the first one is an analog subsystem, which takes charge of design, analysis and computational work using the share mechanism, creates project documents and commits them to the business logic layer. The other is some simple client application based on browser, which help users view the project documents and advise modifications. After that, the server transfers them to the clients.

Business logic layer： It includes the Web server and the business logic server. Its main responsibility is to solve the problems of how to manage project documents submitted by clients, transform their formats (extract valuable information

and create corresponding XML documents) and provide cooperative work control.

1) Web Server: It listens and responds to HTTP requests from clients, establishes connections between server and clients, provides services and issues project documents.

2) Business Server

Document scheduling: It is a scheduler that manages the workflow of the design. All system modules are controlled by the scheduling automation control.

Document management: It includes two key problems: data description and data exchange. Data description confirms the data formats and describes three-dimensional models and planar pictures during the architectural design process, while data exchange mainly deals with the interface realization in the process of data and model conversion.

Cooperative work control: It monitors, archives and eliminates all conflicts during the design according to some prescriptive strategies.

Database interface access and stored procedure call.

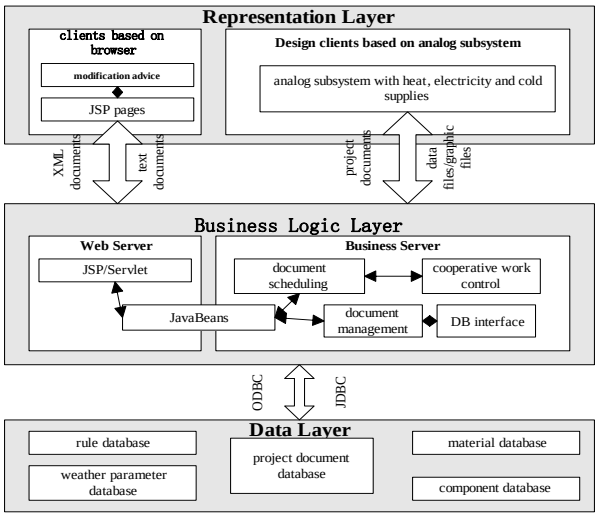


Fig. 1. Layers of cooperative design platform

Data layer： The database system includes the rule database, the weather parameter database, the material database, the component database and the project document database. With the help of stored procedures, the system performance can be greatly improved. All similar operations are encapsulated in one procedure so that they can be altered together.

4. DETAILED DESIGN

General Design

The cooperative design platform includes analog subsystem, business process subsystem and system management subsystem. Users of the system are divided into two types: ordinary users and system administrator. The use case figure is shown in Fig. 2..

1) Ordinary Users

Designing analog subsystems cooperatively: It allows users to choose data files containing load information and create interrelated load parameters, which can help users build up analog systems using analysis and statistic functions, and submit project documents.

Browsing documents: It enables users to browse documents during the design and give their modification advices or the plan of the project.

2) System Administrator

User rights' configuration.

Setting the temperature parameters of air outdoors, device parameters and price parameters of public projects that are used by analog subsystems.

Managing the whole design process.

Managing all types of databases.

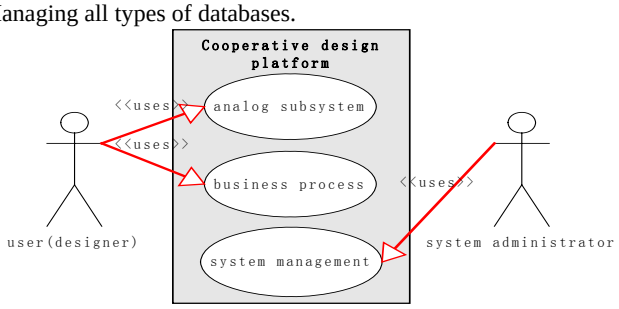


Fig. 2 .Use case figure of system users

Functional Module Designs

This design platform emphasizes particularly on the development of cooperative environment. The designs of analog subsystems are implemented on the client, and the server takes charge of cooperative controls and system management. The structure and relationship diagram is shown in Fig. 3..

1) Analog Subsystem Module **(**It includes five sub modules):

Load selection

The third-party software selects data files of heat and water loads. These data files are in text format. Electricity loads are created automatically in two ways: an automatic load calculation according plans or manual inputs. The loads of daily hot water are also created automatically according to hot water plans.

Building up analog subsystems

There are two ways to construct the form of an analog subsystem: crossing windows and visual components. The flow diagram comprising all kinds of equipments will be created after the forenamed work is finished.

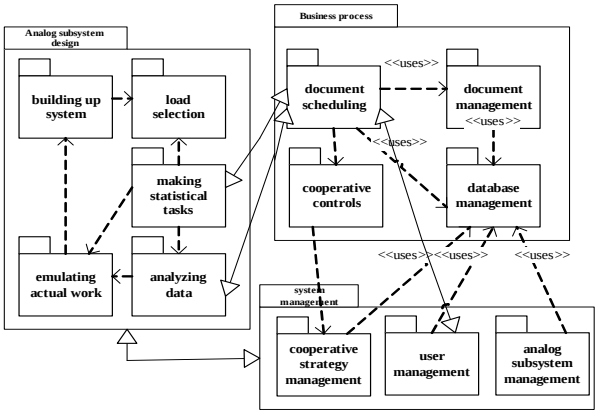


Fig. 3. Structure and relationship diagram

Emulating some actual work

It allows users to choose equipments in the following ways. Users can either choose to click equipment boxes and choose equipment types from dropdown list boxes in the attribute editor, or input the key parameters by themselves.

When all key equipment parameters are inputted, the subsystem begins to run the simulation program with all kinds of loads and the time parameter is set as 8760 hours. During the runtime, all parameters of input and output

equipments should be displayed, and the default sampling time is 3 hours.

Analyzing data

The analog subsystem can set capacities of some equipment such as default electricity quantity of gas turbines and capacities of cold storage equipments. Some key parameters like proportions of power generation will be analyzed.

Doing statistical tasks

Statistical work can show the logistics changes for the whole year (8760 hours), which consists of fuel traffic, power production, cooling capacity and cold storage. Curves are also shown according to the changes of time. It also figures out whole year economic benefits such as fuel cost, electricity cost, electricity benefits and the dynamic recycle period.

2) Business Process Module

Document scheduling

It mainly deals with how to submit, call, edit and send files such as graphics data, related project documents and texts concerning modified opinions, and controls all system modules to work together in phase.

Document management

This module extracts information from project documents submitted by designers with different identities and creates the corresponding XML documents. It includes data extraction, data reorganization and document mode conversion.

Cooperative controls

This function monitors and solves conflicts in the design work. It includes conflict monitoring (catcher and logger), conflict archives and conflict elimination functions based on knowledge reasoning and constraint relaxation.

Database Management

Interactive functions are realized between software and databases such as the rule database, the weather parameter database and the material database, etc.

3) System Management Modules

User Management

Analog subsystem management

Cooperative strategy management

5. SOME KEY PROBLEMS

5.1 Cooperative Design Strategy

During the process of the networked and cooperative architectural design platform, conflicts often take place in each design phases. To some extent, the design process is the process of enouncing conflicts and solving them. The conflicts are inconsistent and opposite during the design process, and can be divided into two types: result conflicts and target conflicts[3]. The former one refers to the opposite relationships in the solution design scheme or attributes design and the latter one means the conflicts existing among design targets when the design solution cannot meet the functional requirements. The conflict monitoring mechanism in parallel designs is not only a kind of activity to solve conflicts, but also a kind of activity to manage and harmonize the design procedure based on conflicts.

The cooperative work mechanism has four primary problems: workflow management, constraint management, work record and conflict coordination. Workflow involves some tasks such as task decomposition, modeling, analysis and control. Constraint management check and judge automatically whether conflicts can meet the requirements of constraint conditions by creating a constraint network, search and broadcast using some algorithms, and can get the values of system parameters and so on. The work record provides records about the reasoning processes and the decision processes during the design and enables members to understand how the current results are achieved. Conflict coordination accounts for how to detect or coordinate conflicts with the support of group decision tools, or how to provide relevant means.

5.2. Mechanism of Data Description, Data Exchange and Share

Because of the inconsistency of data formats from all kinds of architectural analog software, share mechanism is established on the basis of data description and data exchange and consummate project documents, which enables those professional workers to communicate.

Part of the DTD definition of project document is as follows:

```
<! ELEMENT DOCUMENT (PROJECT_INFO MODEL_FILE, DRAWING_FILE, COMMENTS, VEDIO_DEMO, OTHER_MESSAGE)>
<! ELEMENT PROJECT_INFO (PROJECT_NAME, PROJECT_NUM, PROJECT_PERIOD)>
<! ELEMENT MODEL_FILE (SPECIALITY, PATH, CREATE_DATE, LAST_MODIFY_DATE, LAST_ACCESS_DATE)>
…
<!ELEMENT OTHER_MESSAGE (SENDER_NAME, CONTENTS)>
```

6. CONCLUSIONS

In this paper, we propose a system structure of distributed design based on network cooperative work based on the characteristics of processes in district planning and architecture designs. By building the data exchange mechanisms based on XML documents between architectural analog software and coordination and using the flow management technologies to guarantee the high efficiency and coordination of the whole design process, the platform we proposed make it feasible to cooperate among many subjects in district planning and architectural design. Further, we also implement a cooperative design platform of architectural system with heat, electricity and cold supplies based on network. It can support the requisition cooperation of designers and their cooperative designs, and help review experts to make evaluations. In this way, all the designers and review experts can monitor the design process of the architectural system and give modification advice in time.

7. REFERENCES

[1] S. Chunhe, C. Peizhong, Research and Development on a Virtual Laboratory Supporting Computer Cooperative Work, Computer Engineering and Applications, 2003, 1002-8331-(2003)30-0150-03.

[2] Java Language and Technology, http://www.ahjmw.gov.cn/ cit/199910/02.htm.

[3] X. Wensheng, X. Guangleng, Z. Peisi, “Conflict Management in Concurrent Engineering”, Computer Integrated Manufacturing System, 2001,7(1):1-2.

Construction Program Optimization Based On Analytic Hierarchy Process

Tinghua Xu
College of Civil Engineering, Qingdao Technological University
Qingdao, Shandong 266033, China
Email: xuth@qtech.edu.cn
Weizhong Qu
Supervision Company of Qingdao Technological University
Qingdao, Shandong 266033, China
Email: qwzh-jyjl@163.com
Jingru Chen
College of Civil Engineering, Qingdao Technological University
Qingdao, Shandong 266033, China
Email: jingru_chen@tom.com
Zhenhui Liang
Yongxian Li
Qingdao No.1 Construction Engineering Company
Qingdao Shandong 266031, China
College of Civil Engineering, Qingdao Technological University
Qingdao, Shandong 266033, China
Email: lzh3658@163.com

ABSTRACT

The construction program is of great importance to quality, safety and cost of engineering projects. Optimum construction program can be determined according to multi-objective evaluation method. Analytic Hierarchy Process is an effectual method for construction program option. Three aspects including construction period, risk and construction cost were evaluated step by step in the proper order. Analytic Hierarchy Process was employed to choose optimum construction program by the result of evaluation for three objectives.

Keywords: Multi-objective Evaluation, Analytic Hierarchy Process, Construction Period Objective, Cost Objective, Project Risk Objective, Project Network.

1.INTRODUCTION

There are several methods to optimize the construction program of engineering project. Multi-objective evaluation methods are effective and precise. In this paper Analytic Hierarchy Process method was chosen to provide evaluation to the construction program of housing project. Because Analytic Hierarchy Process method can divide complex engineering decision into single objective, and then analyze each intention individually, eventually, leading to the total engineering construction program [1].

The project was a residential district situated in the city of Qingdao, a beautiful city on the east coast of China. There was a steep cliff in the middle of the site from east to west. The project was divided into three parts according to consulting and construction. The first part was made up of two high-rise apartments with a two-story underground park. The second part was a four-story apartment with two-story underground amusement room. The third part was made up of five townhouses with a one-story underground park. The geological conditions of each part were as follows. Part I was thin backfilling soil on surface, below was granite. Part II was backfilling soil between the elevation of 69.76 and 67.66 meter, while granite was below 67.66 meter. Part III was backfilling soil with little granite.

According to characteristics of land and all buildings, construction order was an important parameter to construction period, cost and project risk. Construction order and technology should be determined beforehand. The elevation of construction site was different, in addition, both south and west were one-way roads. Therefore, the construction condition was unsuitable to do earthwork, stack materials and layout construction site facilities. In order to optimize the construction program and choose the best construction scheme, the influence to project period, sales risk and construction cost caused by construction program

Deal with three objectives including construction period, risk and cost simultaneously was complex, optimum construction program was chosen by using Analytic Hierarchy Process method. The decision (parameter) was $f_1(x)$, $f_2(x)$ and $f_3(x)$, according to the importance of three parameters.

2.CONSTRUCTION PERIOD OBJECTIVE

First parameter $f_1(x)$ should be chosen properly. If three parts were built at the same time, there would be no place to stack members, because of three big substructure works. Under such circumstances, material stacking and processing areas should be rented. Therefore, materials should be moved frequently. If three parts were built sequentially, members and processing could be solved in site. However sales risk would be increased. Consequently, construction period will be delayed. Thus all objectives should be studied carefully. Three parts of this project could be built in seven schemes as Table 1. Feasible scheme assembly was: $U=\{A,B,C,D,E,F,G\}$.

Construction duration of earthwork, foundation and main structure of each part could be decided in advance according to construction technology, labor efficiency and construction management. The earliest structure sealing time of each scheme was calculated by project network.

Table 1. Feasible schemes and codes

codes	scheme contents
A	I、II、III built simultaneously
B	III starts after I and II foundation finished
C	II starts after I and III foundation finished
D	I starts after II and III foundation finished
E	II and III start after I foundation finished
F	I and III start after II foundation finished
G	I and II start after III foundation finished

Note: other schemes were omitted because of delaying beyond contract period.

Planning construction process by network would provide order and logic of all works properly. Double code network method was used in this project to map the main structure construction program because of its direct perception and convenience.

Construction period of each line could be calculated as formula (1) and Formula (2).

$$T_L = \sum_{i=1}^{n} t_i \tag{1}$$

$$T_M = Max(\sum_{i=1}^{m} t_i) \tag{2}$$

Where T_L is whole period of single line, n is process number of line, T_M is sealing time of main structure, m is process number of single line before sealing, t_i is duration time of number "i" process.

Each whole period and sealing time could be calculated at Table 2. All period and sealing times could be displayed in Figure 1.

Table 2. Sealing time for each scheme, unit: days

code	part I	part II	part III
A	123	134	168
B	123	134	260
C	123	232	168
D	221	144	168
E	123	209	243
F	215	134	260
G	221	232	168

Note: the processes of calculation were omitted for short.

Sealing commodity housing could be purchased in advance according to present policies, so the first objective was period. It's evaluation basis was formula (3).

$$f_1(x^0) = \min_{x \in R_0 \subset R} f_1(x) \tag{3}$$

Sealing time of D-scheme was 168 days with more risk and financial pressure. So, this scheme should be abandon. Other schemes could be divided into two sets. First set was scheme A,B,C and F. their first sealing time was all 123 days. Second one was scheme E and G. Their sealing times were all 134 days. Delay of 11 days might be remedied by cost. Feasible schemes according to the first objective were as follows. $U_0=\{A,B,C,F,E,G\}$[2,3].

Six feasible schemes could be chosen according to second objective $f_2(x)$.

3.RISK OBJECTIVE

Risk decision was an important parameter in construction program decision. Risk analysis should be made in order to ensure the total engineering project objectives. For this project, there were position, building and elevation differences among three parts. The first part was composed of villas and multi-storied apartments. It was high elevation with hills behind and the sea in front. The project was huge. Hence, it was easy to sale the apartment in this area. Part II was composed of multi-storied apartments in front of a hill and the sea in front too. Its position was perfect. it was easy for sales as well. Part III was composed of multi-stored and high-rise apartments. Its position was unfavorable because of low elevation and closing to the cliff. It was hard to sale the apartments in this area. Important risk parameters and their influences were list in table 3. 0-1 score method was used to assess every parameter in different scheme. Risk result of each scheme was calculated by formula (4).

$$R_j = \sum_{i=1}^{n} W_i \bullet r_{ij} \tag{4}$$

Where R_j is risk result of number "j" scheme, n is numbers of risk factor(n=4), w_i is weight of number "i" risk factor, r_{ij} is score of number "i" factor to number "j" scheme.

Risk assessment result displays risk size. If the score was close to 1, risk of corresponding scheme was low and vice versa. Assessment process was as formula (5). Schemes of A, B and C were feasible schemes according to second objective. $U_1=\{A,B,C\}$.

$$f_2(x^0) = \max_{x \in R_1 \subset R_0} f_2(x) \tag{5}$$

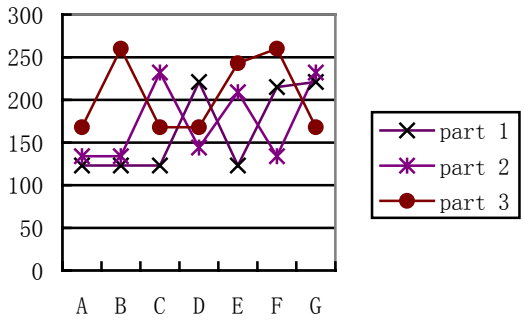


Fig. 1. All period and sealing times

Table 3. Risk analysis

risk factor	weight (W_i)	scheme A	B	C	E	F	G
price falling	0.10	1	1	1	0	0	1
subsoil water	0.30	1	0	1	0	0	1
construction yard	0.20	0	1	1	0	0	0
sales risk	0.40	1	1	1	0	0	0
sum	1.0	0.8	0.7	1	0	0	0.4
assessment		√	√	√	×	×	×

Note: "√" is feasible,"×" is unfeasible.

4.COST OBJECTIVE

Among four selected schemes, member stacking areas and the construction site facilities were in different utilization ratio, because of different construction programs. As a result, costs were different as well. Earthworks in this project include foundation pit cutting, backfilling, and earth transportation to outside. Square grid method was used to survey and calculate earthwork.

All Three parts were divided into square grids in 5 meters. Each square corner was surveyed by surveyor's level. After survey, cutting elevation graph could be draft.

Cutting and backfilling volumes of soils and rocks would be calculated based on geological prospecting report and cutting elevation graph. All volumes were calculated and listed in Table 4.

Table 4. Cutting and backfilling volumes unit: m^3

Part	Cutting volume			Backfilling volume
	Total	Earth	Rock	
I	32984	20120.24	12863.76	6496.00
II	35460.2	25531.34	9928.86	5976.00
III	38684.7	12379.10	26305.60	7264.00

Note: the calculation processed is omitted for short.

Earthwork and construction site facility costs could be decided according to construction program and current market prices in Table 5 for short. Other costs were omitted because of sameness for each scheme.

Seeking for the third objective might be displayed as formula (6). The lower the cost is, the better the scheme.

$$f_3(x^0) = \max_{x \in R_2 \subset R_1} f_3(x) \tag{6}$$

Analysis to the third objective was as follows. The cost scheme A was the highest, and it's affected by outside with more uncertain factors. So it was unfavorable to either construction management or period. We should omit it. Scheme C was low in both cost and risk. It was the most suitable scheme. $U_2=\{C\}$.

Table 5. Earthwork costs calculation unit: RMB

scheme	earthwork cost (1)	facility cost (2)	costs sum (1)+(2)
A	3591683.6	724600	4316283.6
B	3056322.2	455000	3511322.2
C	2644761.2	455000	3099761.2

Unit price and cost were listed as followings.

4.1 Unit price of earth or rock cutting was 28 RMB/m^2;

4.2 Unit price of earth cutting and transportation in site was 10RMB/m^2;

4.3 Unit price of transportation and backfilling in site was 15RMB/m^2;

4.4 Unit price of transportation from outside and backfilling was 30RMB/m^2;

4.5 Construction site facilities cost for stage construction was as follows.25000000×(1.3+0.52)%=455000(RMB);

4.6 Construction site facilities cost for simultaneous construction was as follows.

1)Dormitory cost: 480×2×15×12=172800(RMB);

2) Processing area cost was ¥80000;

3) Warehouse cost was 200000RMB;

4)Twice transportation cost: 4530×60=271800(RMB);

4.7 Total cost was 724600(RMB).

5.CONCLUSIONS

Construction management was complex system engineering. After consideration of construction site, construction period, project risk and cost, construction program decision can be executed according to Analytic Hierarchy Process method. We can conclude from this paper: As one of multi-objective decision methods, Analytic Hierarchy Process method can achieve construction program decision effectively. The most feasible scheme was fast, low risk and cost. Comparing to regular A-scheme, it can save 1216522.4RMB. The decision process is scientific, comprehensive, just and precise. It can satisfy the total objective of engineering construction.

6.REFERENCES

[1] S.Y. Wang, G.H. Liu, System Analysis, Hangzhou: Zhejiang University Press, 2001.

[2] J.B. Jiang, Theory and Usage of Network, Shanghai: Tongji University Press, 1998.

[3] Z.Y. Lin, Engineering Construction Plan and Management, Tongji Shanghai: University Press, 2002.

Modeling J2EE Applications USING RUP

Sheng Ding, Li Mao
School of Information Technology, Southern Yangtze University
Wuxi, Jiangsu, 214112, China
Email: Sandyding2004@gmail.com

ABSTRACT

The Rational Unified Process (RUP) is a comprehensive and detailed software development process framework, which can be customized for a wide range of software projects. The modeling method for developing J2EE applications, discussed in this paper, is a customized methodology based on RUP. The modeling method is used to build a B2B online ordering system for textile industry. The result shows that this method helps developing robust, efficient, reusable J2EE applications quickly.

1. INTRODUCTION

RUP [1] is an architecture-centered, use-case-driven, iterative and incremental process framework. In RUP, the software development lifecycle is divided into four phases: inception, elaboration, construction and transition. Each phase includes one or several successive iterations. Each iteration is a complete development process, which results in improvements in meeting user needs, robust architecture and implementation. The software development becomes closer to the goal of a complete system after each iteration process.

But RUP is a comprehensive and detailed software development process framework, and hence no single project will make use of all RUP content. Therefore customizing RUP to fit a particular software project is necessary [2]. In this paper, a modeling method for J2EE applications using customized RUP is presented, and an example of developing J2EE project in textile industry is used to demonstrate the feasibility and success of this modeling method.

2. The Modeling Method FOR J2EE Applications with RUP

The modeling method used in this paper is a customization of RUP. In this method, RUP has been tailored specifically for building J2EE applications. Based on the needs of the J2EE developer, it focuses on providing guidance on requirements gathering, analysis, design, implementation, test, etc. Each phase is a refinement of, and based on the previous one, and is an iterative and incremental development process. Among these phases, requirements gathering phase builds use-case model and acquires all the requirements that significantly effect software architecture. Analysis phase transforms the requirements into an analysis model. Design phase uses the results of analysis to produce a design model, a specification that can be efficiently implemented. Implementation phase creates source and executable code to implement the results of design, and makes J2EE modules for deployment. Test phase verifies that the resulting system works as designed and meets the user requirements. These phases focus on the development of the following models: use-case model, analysis model, design model, implementation model and test model. The use-case model is the foundation of the other models, and used as an essential input to activities in analysis, design, implementation, and test.

3. Example: B2B online ordering system

The B2B online ordering system is a subsystem of Wuxi public online information services developed specifically for textile industry. The online ordering system provides fast, effective and safe solutions for ordering textile products for a large number of the enterprise customers.

3.1 Requirements Gathering Phase

Requirements gathering phase is essential to software development. A complete and accurate understanding and gathering of the requirements is the key to the success of system development.

Requirements gathering phase results in the following key artifacts: use-case model, supplementary specification, and glossary. The use-case model defines system behavior from an external perspective. It includes the actors that interact with the system, and the use-cases that describe these interactions. A good analysis of the use-cases, which describe user's explicit requirements, is the key to the success of the project. The supplementary specification captures system-wide requirements that are unable to be captured by the use-case model, e.g., system reliability and performance, operating systems and environment requirements. The glossary defines important terms used by the system, which serve as a contract between the customer, the end users and the system developers.

According to the business features and functional requirements, the system contains on-line shopping, backstage management, and negotiating, on-line help subsystems. Due to the length limit of this paper, only the on-line shopping subsystem is discussed. There are two kinds of actors in the on-line shopping subsystem: active actor and passive actor. Active actors, including enterprise customers, system operators and managers, input data to the system. Passive actors, such as banks, receive data from the system. And the use-cases in this subsystem include login, registration, inquiring product, placing order, shopping cart management, checking order, on-line payment, etc.

It is insufficient to describe system requirements only by using a simple diagram. After gathering the use-cases, a more detailed description, using written documents, or dynamic and interactive diagrams, is necessary for explaining the meaning and concrete procedures of each use-case. By describing each use-case in detail, we obtain a use-case model that meets the system requirements as the starting point and foundation of the further development processes.

3.2 Analysis Phase

The objective of analysis phase is to obtain the first approximation of the system architecture and behavior framework and to lay the groundwork for design phase.

There are two major activities in this phase: defining an initial architecture and analyzing the system functional behavior.

It is usually straightforward to define an initial framework with J2EE technologies, because there are already a number of standard J2EE deployment configurations on which we can base our initial architecture framework. For example, a multi-tier J2EE deployment configuration, which contains separate Web and EJB servers, is selected in this online ordering system.

Analyzing system behavior focuses on identifying the analysis classes, their responsibilities and interaction relations, in order to get a more accurate understanding of the system requirements. The analysis classes typically include boundary classes, control classes, and entity classes. A boundary classes is used to model the boundary and interactions between the system and its surroundings (including users, external systems or devices). Control classes encapsulate use-case-specific behavior and handle the main control flows. Entity classes encapsulate information represented within the system, together with any attributes and associated behavior. Entity classes usually represent persistent system elements.

In addition to analysis classes, the analysis model also includes a series of analysis-level use-case realizations. Each analysis-level use-case realization has a structural view and a dynamic view. The structural view, represented as a class diagram, describes the analysis classes that involve in the use-case realization and the relationships between them. The dynamic view, represented as a set of interaction diagrams, describes the analysis class interactions that occur during the execution of the use case.

After describing each use-case with structural and dynamic views, we need to incorporate these individual user-case realizations as a whole. For example, we incorporate and refine the analysis classes with consistent responsibilities, attributes and relationships. The ultimate goal is to build an analysis model that is consistent across all the use cases.

3.3 Design Phase

The objective of design is to take the models obtained from the analysis phase and produce a specification that can be practically and conveniently implemented by software developers with programming code. The major task of design is to transform analysis classes in analysis model into design elements of the system.

With the guidance of the system architecture, we can identify design elements in different J2EE layers (presentation layer, business layer, and integration layer) by analyzing characteristics and functions of analysis classes.

First, we identify design elements in the business layer (e.g. enterprise components, EJBs and Java classes). We organize the system business layer as a set of enterprise components. An enterprise component [2] is a software implementation of a business concept or process. It is used to encapsulate business behavior, and to access or manage system data. Some general guidelines for identifying enterprise components are as follows:

1) Divide the entity classes into groups such that members of a group can be managed together by an enterprise component that handles responsibilities involving manipulating these entities.

2) By looking into analysis-level realizations of interrelated use cases, find and group similar responsibilities of the participating control classes in those realizations, and use an enterprise component to handle these responsibilities.

3) Consider evolving boundary analysis class representing a passive actor into an enterprise component.

After identifying the enterprise components, we organize the internal structure and access control of the enterprise components using J2EE multi-patterns. These patterns [3] include Business Delegate, Session Façade, Application Service, Business Object, Service Locator and Transfer Object. The patterns for the internal structure and access control of the enterprise components are shown in Figure 1.

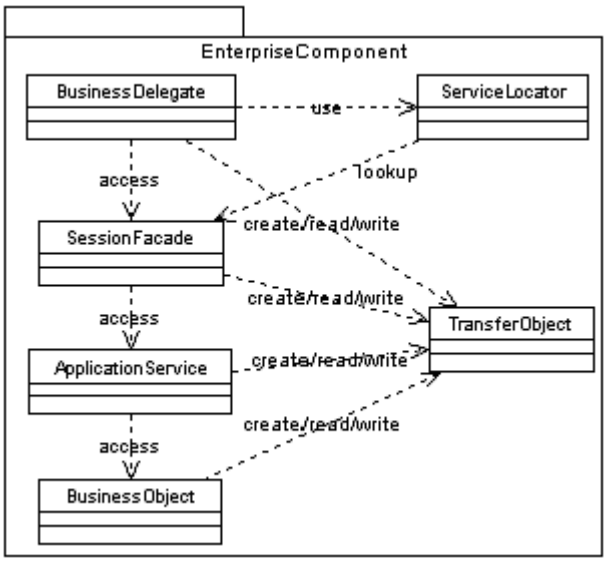


Fig. 1. An Internal Perspective of an Enterprise Component

The internal business accessing process of enterprise component is as follows:

1) Clients access an enterprise component via one or more Business Delegates;

2) Business Delegate uses service locator to find and instantiate Session Facade, then transfer the client's service request to this Session Facade.

3) The Application Service, which is invoked by the Session Façade, completes client's service request by interacting with several Business Objects.

4) During business data transfer process; the Transfer Object is used to transfer data over multi-layers.

We use a session EJB to implement the Session Façade of enterprise component. Entity classes of analysis model are mapped into Business Objects of enterprise component, implemented with entity EJBs or Java classes. Control classes of analysis model (responsible for interval interactions, processing the business logic, calculating, and interacting with entity classes.) are mapped into Application Service of enterprise component, implemented with session EJBs or Java classes. We also need to define additional design elements such as Business Delegate Java classes, Service Locator Java classes, Value Object Java classes etc.

Next, we define design elements in the presentation layer such as Servlets, JSPs and Java classes. We map boundary classes, which are responsible for interacting with active actors, into JSPs in the presentation layer. We map control classes, which are closely related to boundary classes and responsible for interacting with external, into servlets in the presentation layer (e.g. Front controller servlet). In the integrate layer, the major design elements that need to be defined are data access object Java classes.

After identifying the design elements, we also need to refine and integrate them. Finally, all the analysis classes (as

described in class diagram and interaction diagrams of the analysis-level Use-Case Realizations.) are replaced by the design elements, which result in the design-level Use-Case Realizations.

3.4 Implementation and Test Phase

The major objectives of implementation and test are to implement software elements based on the design, to test the product quality and performance, and to release the final product. Due to the length limit, the discussion on this phase is omitted here.

4. CONCLUSION

In this paper, RUP has been well designed and customized for J2EE applications, and successfully applied to model a B2B online ordering system. The result shows that modeling J2EE applications with RUP can reduce the risks in software development, improve development efficiency, shorten the developing time significantly, and enhance numerous aspects of the system, including maintainability, extensibility, and reusability.

5. REFERENCES

[1] P. Kruchten, “The Rational Unified Process”, An Introduction, Second Edition, Addison Wesley, 2000.

[2] P. Eeles, K. Houston, et al, “Building J2EE Applications with the Rational Unified Process”, Addison Wesley, 2002.

[3] D. Alur, J. Crupi, et al, “Core J2EETM Patterns: Best Practices and Design Strategies”, Second Edition, Prentice Hall, 2003.

Traffic flow forecasting of Intersection based on a novel Associative Memory System

Chao Xie, Luyuan Liu
Department of Automation, Tianjin University, Tianjin, China
Department of Automation, Tianjin University of Technology and Education, Tianjin , China
Email: wjsong2004@126.com, **xiechao6@126.com**

ABSTRACT

A novel high-order Associative Memory System was firstly proposed based on Newton's Forward Interpolation (NFI-AMS), which is capable of implementing error-free approximations to multi-variable polynomial functions of arbitrary order. The advantages that NFI-AMS offers over conventional CMAC-type neural network are: high-precision of learning, much smaller memory requirement without the data-collision problem and also much less computational effort for training and faster convergence rates than that attainable with multi-layer BP neural networks. Secondly, The on-line traffic flow intelligent rolling predictive method was designed based on NFI-AMS, and the simulation results show that this method is effective.

Keywords: Associative Memory System, Newton's Forward Interpolation, Traffic flow prediction.

1. INTRODUCTION

The study on traffic flow prediction plays an important role in improving traffic safety and road capacity with the rapid development of intelligent traffic system. In general, it is very difficulty to establish precise traffic flow model with definite mathematic function or model. Recently, Back-propagation (BP) neural network, a typical case of neural networks, is used most widely and is more mature than other networks. It can approximate the specific input and output relationship without a certain model. Therefore, it is usually used to set up a BP ANN model for traffic flow, Unfortunately, traditional BP network has some weaknesses that it is easy to relapse into local minimum point and computation convergence speed is somewhat slower, which affect the enhancement of reliability and accuracy of prediction model. In this paper, to found traffic flow mathematic model accurately, a novel high-order Associative Memory System is proposed based on the Newton's Forward Interpolation (NFI-AMS), which is capable of implementing error-free approximations to multi-variable polynomial functions of arbitrary order. The advantages it offers over conventional CMAC-type neural network are: high-precision of learning, much smaller memory requirement without the data-collision problem and also the advantages of much less computational effort for training and faster convergence rates than that attainable with multi-layer BP neural networks, and it is capable of overcoming the inherent weaknesses of BP algorithm that training processes often settle in undesirable local minima of the error surface. Based on the novel AMS, the author designs a type of traffic flow intelligent predictive strategy, namely, use the measured sample data to train NFI-AMS. Thus, the network is capable of predicting the future traffic flow. The simulation results verify that it is effective.

2. A NOVEL ASSOCIATIVE MEMORY SYSTEM BASED ON NEWTON'S FORWARD INTERPOLATION (NFI-AMS)

A. A NEWTON'S FORWARD INTERPOLATION POLYNOMIALS

Consider the one-dimensional case. If $\mu+1$ sampling data pairs $\{s_i, f(s_i)\}$, $i=1,2,\cdots,N$ are given, the μ -th order Newton's forward interpolation polynomial can be expressed as [1]

$$\begin{aligned}\phi_\mu(s) &= \Delta^{(0)} f(s_0) + (s-s_0)\Delta^{(1)} f(s_0) + \cdots \\ &\quad + (s-s_0)(s-s_1)\Delta^{(2)} f(s_0) \\ &\quad + (s-s_0)(s-s_1)\cdots(s-s_{\mu-1})\Delta^{(\mu)} f(s_0) \\ &= \sum_{l=0}^{\mu} \Delta^{(l)} f(s_0) \qquad (1)\end{aligned}$$

Where

$$c_l = \prod_{j=0}^{l-1} t_j = \begin{cases} 1, l=0 \\ t_0 t_1 \cdots t_{l-1}, l \geq 1 \end{cases} \quad t_j = s - s_j \quad .$$

And $\Delta^{(j)} f(s_i)$ is the j -th order discrete difference of $f(s)$ given by

$$\Delta^{(j)} f(s_i) = \begin{cases} f(s_i), j=0 \\ \dfrac{1}{s_{i+1}-s_i}[\Delta^{(j-1)} f(s_{i+1}) - \Delta^{(j-1)} f(s_i)], j \geq 1 \end{cases}$$

To approximate a μ -th order N -variables dimensional polynomial, the Newton's forward interpolation polynomial is given by [2]

$$\begin{aligned}\phi_\mu(s_1, s_2, \cdots s_N) &= \Delta^{(0,0,\cdots,0)} f(s_{1_0}, s_{2_0} \cdots, s_{N_0}) \\ &+ [(s_1 - s_{1_0})\Delta^{(1,0,\cdots,0)} f(s_{1_0}, s_{2_0} \cdots, s_{N_0})\end{aligned}$$

$$+[(s_2 - s_{2_0})\Delta^{(0,1,\cdots,0)} f(s_{1_0}, s_{2_0} \cdots, s_{N_0}$$
$$+\cdots+(s_N - s_{N_0})\Delta^{(0,0,\cdots,1)} f(s_{1_0}, s_{2_0} \cdots, s_{N_0})$$
$$+\cdots$$
$$+\sum_{l_1+l_2+\cdots+l_N=0}^{\mu}(\prod_{j=0}^{l_1-1} t_{1_j}\prod_{j=0}^{l_2-1} t_{2_j}\cdots\prod_{j=0}^{l_N-1} t_{N_j})\Delta^{(l_1,l_2,\cdots,l_N)} f(s_{1_0}, s_{2_0} \cdots, s_{N_0})$$
$$=\sum_{l_1+l_2+\cdots+l_N=0}^{\mu} c_{l_1 l_2 \cdots l_N}\Delta^{(l_1,l_2,\cdots,l_N)} f(s_{1_0}, s_{2_0} \cdots, s_{N_0}) \quad (2)$$

where

$$c_{l_1,l_2,\cdots l_N}=\prod_{j=0}^{l_1-1} t_{1_j}\prod_{j=0}^{l_2-1} t_{2_j}\cdots\prod_{j=0}^{l_N-1} t_{N_j}$$
$$t_{i_j}=s_i - s_{i_j}, i=1,2\cdots,N$$

And $\Delta^{(l_1,l_2,\cdots,l_N)} f(s_{1_0}, s_{2_0} \cdots, s_{N_0})$ is the $(l_1+l_2+\cdots+l_N)$-th order discrete difference of $f(s_1, s_2, \cdots, s_N)$.

B. INTERPOLATION ALGORITHM

Tolle et al. indicated that the weights number strictly depended on the coefficient of a N-variable polynomial function in the form of [3]

$$f(s)=f(s_1, s_2, \cdots, s_N)$$
$$=a_0+\sum_{j=1}^{N} a_j s_j+\sum_{j=1}^{N}\sum_{j_2=j_1}^{N} a_{j_1} a_{j_2} s_{j_1} s_{j_2}+\cdots$$
$$+\sum_{j=1}^{N}\sum_{j_2=j_1}^{N}\cdots\sum_{j_\mu=j_\mu-1}^{N} a_{j_1} a_{j_2}\cdots a_{j_\mu} s_{j_1} s_{j_2}\cdots s_{j_\mu} \quad (3)$$

where $s=[s_1, s_2, \cdots, s_N]$. Thus, the total number of its coefficients is

$$n_\mu=1+C_N^1+C_{N+1}^2+\cdots+C_{N+\mu-1}^{\mu}+=C_{N+\mu}^{\mu}$$

So n_μ pieces of independent information is enough to approximate μ-th order polynomial function in a specified N-dimensional input space.

In order to reduce the computational effort, it is necessary to develop an implicit expression, in which all the discrete differences (instead of the function values themselves) of $f(s_1, s_2, \cdots, s_N)$ are directly used in the interpolation algorithm [4].

From Eqn. (1), a single-variable High-order polynomial function can be approximated as follows

$$f(s)=\phi_\mu(s)+R_\mu(s)\approx\phi_\mu(s)=\sum_{l=0}^{\mu} c_l\Delta^{(l)} f(s_0) \quad (4)$$

Where $R_\mu(\bullet)$ is the residual composed of ail differences with order higher than μ.

A N-variable high-order polynomial function can be approximated by

$$f(s_1, s_2, \cdots, s_N)$$
$$=\phi_\mu(s_1, s_2, \cdots, s_N)+R_\mu(s_1, s_2, \cdots, s_N)$$
$$\approx\phi_\mu(s_1, s_2, \cdots, s_N)$$
$$=\sum_{l_1+l_2+\cdots+l_N=0}^{\mu} c_{l_1,l_2\cdots l_N}\Delta^{(l_1,l_2,\cdots,l_N)} f(s_1, s_2, \cdots, s_N) \quad (5)$$

It can be easily proved that the total number of its coefficients is

$$n_\mu=1+C_N^1+C_{N+1}^2+\cdots+C_{N+\mu-1}^{\mu}+=C_{N+\mu}^{\mu} \quad (6)$$

The distribution of the active cells of the two-dimensional cases is illustrated as Fig.1.

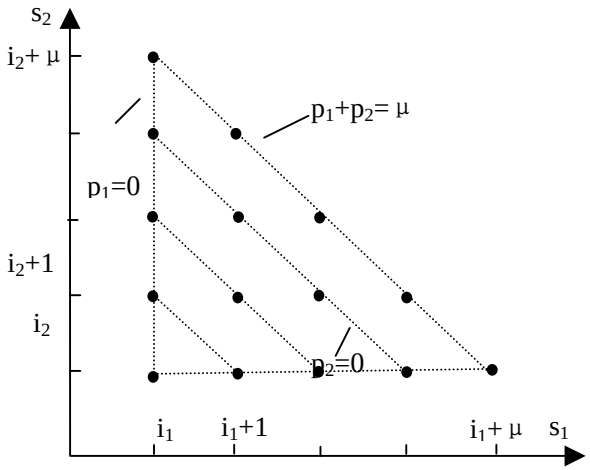


Fig. 1. Distribution of the active cells of the two-dimensional cases

C. TRAINING ALGORITHM

The training of the proposed NFI-AMS is performed by using the sampling data pairs in the form of [5]

$$s=(s_{1_0}, s_{2_0}, \cdots, s_{N_o})$$
$$\Rightarrow f(s)=f(s_1, s_2, \cdots, s_N)=p$$

During the training process, the actual output of the NFI-AMS should be firstly calculated as follows

$$\hat{p}=\sum_{l_1+l_2+\cdots+l_N=0}^{\mu} c_{l_1 l_2 \cdots l_N} w^{(N)}(\alpha_{l_1 l_2 \cdots l_N}) \quad (7)$$

Where $w^{(N)}(\alpha_{l_1 l_2 \cdots l_N})$ is the original weight stored in the memory cell addressed by $\alpha_{l_1 l_2 \cdots l_N}$ in the last training run. Furthermore, all the weights of the conceptual receptive field are updated according to the error between the teaching signal ρ and actual output

$\hat{\rho}.e = p - \hat{p}$ as follows

$$w^{(N+1)}(\alpha_{l_1 l_2 \cdots l_N}) = w^{(N)}(\alpha_{l_1 l_2 \cdots l_N}) + \frac{c^2{}_{l_1 l_2 \cdots l_N}}{C_e} e \quad (8)$$

where

$$C_e = \sum_{l_1 + l_2 + \cdots + l_N = 0}^{\mu} c^3{}_{l_1 l_2 \cdots l_N} \quad (9)$$

It is worth to note that in the irregular training the positions of the N terminal points $(s_{1_0}, s_{2_0}, \cdots, s_{N_0})$, $i = 1,2,\cdots,N$ do not affect the values $c_{l_1 l_2 \cdots l_N}{}'s$ used in Equations (10), so they can be arbitrarily chosen [6].

D. CONTENT-ADDRESSING MECHANISM

The content-addressing mechanism is implemented such that the address of the cell storing the function value, $f(i_1, i_2, \cdots, i_N)$, directly depends on the contents of the information carried by the input vector [$i_1, i_2, \cdots, i_N$], $i_j = 0,1,\cdots,M_j \quad j = 1,2,\cdots,N$ in the following simple way.

$$\alpha_{i_1, i_2, \cdots, i_N} = \sum_{j=2}^{N} i_j \prod_{j=1}^{j-1} (M_j + 1) + i_1 + 1 \quad (10)$$

where $i_j = 0,1,\cdots,M_j$, $j = 1,2,\cdots,N$, and M_j is the segmentation number of the variation range of the j-th component s_j.

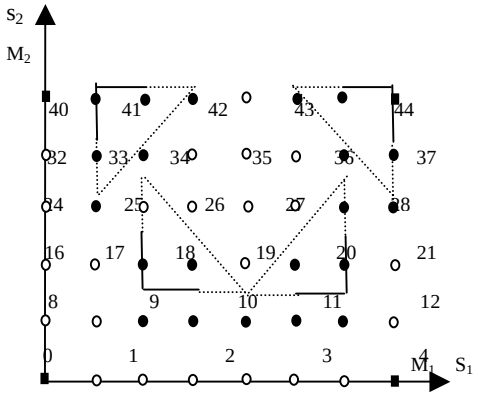


Fig. 2. Input space and memory cell of NFI-AMS with two-dimension input

In a practical use, if the active cells in a receptive field are restricted to a hyper lower triangular sub-domain, the following problem will be encountered with the content-addressing method: Once the input vector, enters into the right upper marginal region:

$$D'_{i_1, i_2, \cdots, i_N} = (M_1 - s_1 \le \mu) \cap (M_2 - s_2 \le \mu) \cap \cdots \cap (M_N - s_N \le \mu) \quad (11)$$

Then the information stored in the corresponding memory cells will no longer be enough for recalling the desired output [7].

To overcome this limit, a novel strategy for addressing the active cells in the receptive field is chosen as follows: The input space is uniformly divided into 2N parts, and thus each part must share one comer of the input space. Within each part, the right angle of a hyper triangular sub-domain should always be pointed to the comer of input space, and thus the corresponding active cells can be determined with easy. The Input space and memory cell of NFI-AMS with two dimension input is shown as Fig.2.

3. INTELLIGENT PREDICTIVE MODEL of TRAFFIC FLOW BASED ON NFI-AMS

A. STRATEGY of TRAFFIC FLOW FORECASTING BASED ON NFI-AMS

The intersection of which traffic flow need to be forecasted is shown as Fig.3. The traffic flow $y(t)$ of the intersection is function of $w(t)$, $r(t)$ and $p(t)$ which is the traffic flow of the neighbor one, the function relationship is as follows:

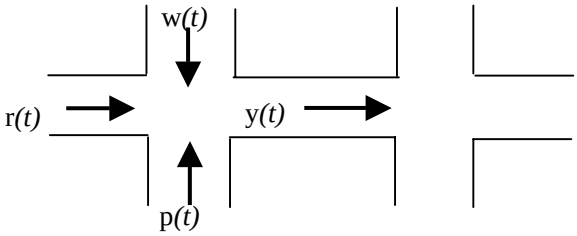


Fig. 3. The forecasted traffic intersection

Based on the relationship between the two intersections, we can assume that $y(k)$ mainly depends on the $w(k-1), w(k-2), r(k-1), r(k-2), p(k-1)$ and $p(k-2)$, in which look back interval is '1'. Provide the assumption above in a equation. We get:

$$y(k) = f[w(k-1), w(k-2), r(k-1), r(k-2), p(k-1), p(k-2)] \quad (12)$$

Where $y(k)$ is the predicted traffic flow of the intersection at *k* instant.

Generally speaking, it is difficult to obtain accurate function relationship of Eqn.12.Based on the novel NFI-AMS, this paper proposes a type of efficient predictive model, the implementation of which needs no model knowledge, To use a NFI-AMS network to model the Eqn.12, we set the inputs of the NFI-AMSS as $w(k-1), w(k-2), r(k-1), r(k-2), p(k-1)$ *and* $p(k-2)$, the output as $y(k)$. Based on the historical data $w(i)$, $r(i)$ and $p(i)$, . $(i = 1,2,\ldots)$.

B. DESIGN of NFI-AMS

The structure of predictive model based on NFI-AMS is

illustrated as Fig.4.
The parameters of NFI-AMS is as follows:

Input dimension: $N = 6$

Output dimension: $N_p = 1$

Segmentation number: $M_i = 8,$ *(j=1,2,…,6)*

Approximation order: $\mu = 2$

Memory size required: $N_a = (8+1)^3 = 729$

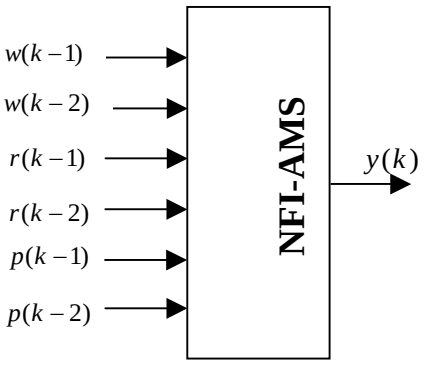


Fig. 4. The structure of predictive model based on NFI-AMS

we can train the NFS as described in the last section. After training, we set the input as $D_{i+1}(t)$ … $D_{i+1}(t-a)$; the output of NFS will be the forecasted $F_{i+1}(t+1)$.

Each predictive cycle consists of a training period and a recalling period. At the beginning of each predictive cycle, a training step is executed; the observed state of the system during the previous predictive cycle is used as input to the NFI-AMS network, which produces $\hat{f}(\bullet)$, so the training data pairs for the NFI-AMS are respectively chosen as follows:

$$\text{w(k -1)}, w(k-2),\ r(k-1), r(k-2),\ p(k-1),\ p(k-2) \Rightarrow y(k) \qquad (13)$$

The different between the predictive traffic flow $\hat{f}(\bullet)$ and the actual traffic flow $y(k)$ during the previous cycle is used for adjusting the weights stored in the NFI-AMS cells. After the training period, the recalling step is executed. According to the w(k), w(k -1), r(k), $r(k-1)$, p(k) and $p(k-1)$, the traffic flow of the next time can be forecasted, so we can implement the on-line intelligent rolling predictive of the traffic flow [8,9].

As the NFI-AMS networks can be continually trained on successive predictive cycles, the NFI-AMS function $\hat{f}(\bullet)$ gradually forms an approximation of the corresponding function $f(\bullet)$ over particular regions in the state space [7]. Finally, if the future situation of the traffic flow is similar to previously trained one, then the NFI-AMS network can output a suitable predictive traffic flow. As a result, the state errors will be small and the NFI-AMS based predictive model will finally approximate the Eqn.12 with wonderful performance.

4. SIMULATION AND CONCLUSION

We applied respectively the NFI-AMS and the conventional CMAC-AMS neural network to forecasting the traffic flow. The average error and memory size required of the two predictive approach is shown as Table 1. The on-line intelligent rolling predictive method of traffic flow is proposed based on a novel high-order Associative Memory System (AMS) which is designed via the Newton's Forward Interpolation, and is capable of implementing error-free approximations to multi-variable polynomial functions of arbitrary order. The advantages of the predictive model based on the novel AMS offers over that based on conventional CMAC-type neural network are: high-precision of learning, much smaller memory requirement without the data-collision problem and also the advantages of much less computational effort for training and faster convergence rates than that attainable with multi-layer BP neural networks. The simulation results show that the predictive method is effective. The future work is the hardware realization of the predictive algorithm.

TABLE 1 COMPARISON BETWEEN NFI-AMS BASED APPROACH AND CMAC-AMS BASED APPROACH

Types of AMS	Memory size required	Errors
CMAC-AMS	305200	0.02350
NFI-AMS	38700	0.00191

REFERENCES

[1] W.T. Miller, F.H. Glanz, “CMAC: An Associative Neural Network alternative to Back-Propagation”, Proc. IEEE. 1990,1561-1567.

[2] J.S. Wang, “Associative Memory Systems-based Robot Intelligent Control System”, Bei-Jing Poly-technique University Master Paper, 1998.

[3] N.S. Xu, “Generalized SHSF-type Associative Memory System”, Proc. of The 13th IFAC World Congress, USA, 1996, 775-780.

[4] N.S. Xu, “A New Content-addressing Mechanism of CMAC Associative Memory System”, Proc. of The 13th IFAC World Congress, USA, 1996, 435-439.

[5] N.S. Xu, “A Novel Associative Memory based on Sharp angled Hat Spline Functions”, The 1996 IEEE Int. Conference on Systems, Man and Cybernetics, Beijing China, 1996, 381-386.

[6] N.S. Xu, “An Innervated Structure System Scheme Using Associative Memory Systems”, Research on Complex System Control Problems. Academia Simca, 1995, 763-768.

[7] N.S. Xu, “Associative Memory-based Intelligent Control of Ship-steering Systems”, Proc. European Control Conference, ECC95. Rome. Italy. 1995, 546-560.

[8] J. Chiu, “Fuzzy Model Identification Based on Cluster Estimation”, Journal of Intelligent & Fuzzy Systems, Sept. 1994, 2(3).

[9] J.S. Jang, “Structure determination in fuzzy modeling: a fuzzy CART approach”, Proc. Of IEEE international conference on fuzzy systems, Orlando, Florida, June 1994.

Generalized E – Convex Sets and Generalized E – Convex Functions

Zuhua Liao [1] **, Hui Gu** [1,2]
[1] **Department of Mathematics, Southern Yangtze University, Wuxi 214122, P.R.China**
[2] **QingShan Senior Middle School, Southern Yangtze University, Wuxi 214062, P.R.China**
Weiyan Fang
DonLing Junior Middle School, Wuxi 214001,P.R.China
E-mail: liaozhuhua@yahoo.com.cn

Abstract

The concept of convexity is one of the most important aspects in solving fuzzy decision-making problems. While making a decision, the maximum and the minimum value always have been a distinctive factor that might drive the result away from the track. To make the result more sound, we defined a class of sets and a class of functions called generalized $E-$ convex sets, $E-$ convex functions. This definition slightly modified the interval $[0,1]$ with $(\lambda,\mu]$ instead. The generalized convexity is based on the effect of an operator E and cut away the extreme situations. And some fundamental properties are obtained with the limitation of this concept.

Keywords: Convex sets, Convex functions, generalized convexity, control theory.

1. INTRODUCTION

Since Zadeh [5] introduced the notion of a fuzzy set in 1965, convex fuzzy set has been further studied in detail by many scholars. Some of its properties can be referred to [1,4,6,7,10,11,12,14]. Based on the concept of $(\alpha,\beta)-$ fuzzy subgroup introduced by Bhakat S K [8] in 1992 and generalized fuzzy subgroup by Xuehai Yuan [9], we have extended the concept of convex fuzzy set, which was accepted by the journal Fuzzy Systems and Mathematics [3]. And recently the class of convex sets has been extended to the class of $E-$ convex sets by Youness[13] and X. Chen [2].

In Section 3 of this paper, we extend the notion of $E-$ convex sets on $[0,1]$ to $(\lambda,\mu]$ and study some properties of this class sets. In Section 4, we extend the class of $E-$ convex functions to the new class of $E-$ convex functions. This class is more general, as it cuts away the extreme point 0 and 1, excluding that can be more soundness. And the results we have in this paper can be better applied to the control theory research.

2. PRELIMINARIES

Definition 2.1 X is a linear space. A map A: $X\to[0,1]$ is said to be a fuzzy subset of X. And all the fuzzy sets of X is denoted by $F(X)$.

Definition 2.2 A set $C\subset R^n$ is said to be $E-$ convex iff there is a map $E: R^n\to R^n$ such that $(1-l)E(x)+lE(y)\in C,$ for each $x,y\in C$ and $0\le l\le 1.$

Definition 2.3 $A\in F(X)$ and A is a fuzzy set. If $A(lx+(1-l)y)\vee\lambda\ge A(x)\wedge A(y)\wedge\mu$, $\forall x,y\in X$ and $\forall l\in[0,1]$, then A is said to be a generalized convex fuzzy set.

3. $E-$ CONVEX SETS

Definition 3.1 Let $\lambda,\mu\in[0,1],$ and $\lambda<\mu$, $A\in F(X)$, $\forall x,y\in X$, $\forall l\in[0,1]$.

If there is a map $E: R^n\to R^n$ such that $A(lE(x)+(1-l)E(y))\vee\lambda\ge A(x)\wedge A(y)\wedge\mu$, then A is said to be a generalized $E-$ convex fuzzy set of X.

Proposition 3.1 Every generalized convex fuzzy set A is generalized $E-$ convex.

The proof is clear by taking a map $E: R^n\to R^n$ as the identity map.

Theorem 3.1 For a fuzzy subset A of X, the following statements are equivalent.

(i) A is a generalized $E-$ convex fuzzy set of X;

(ii) $\forall\alpha\in(\lambda,\mu]$,

nonempty set $A_\alpha=\{x\in X\mid A(x)\ge\alpha\}$ is a $E-$ convex set.

Proof. (i) $\Rightarrow$ (ii)

If A is a generalized $E-$ convex fuzzy set, $\forall\alpha\in(\lambda,\mu]$, $\forall x,y\in A_\alpha$, then $A(x),A(y)\ge\alpha.$

So $A(lE(x)+(1-l)E(y))\vee\lambda$
$\ge A(x)\wedge A(y)\wedge\mu\ge\alpha\wedge\mu=\alpha,$

Since $\lambda<\alpha,$

so $A(lE(x)+(1-l)E(y))\ge\alpha.$

Therefore $lE(x)+(1-l)E(y)\in A_\alpha$.

Thus A_α is a $E-$ convex set.

(ii) $\Rightarrow$ (i)

Assume that, there exist $x_0,y_0\in X$,

Such that $A(lE(x_0)+(1-l)E(y_0))\vee\lambda$
$<\alpha_0=A(x_0)\wedge A(y_0)\wedge\mu,$
So $\alpha_0\in(\lambda,\mu]$, and
$\forall x_0,y_0\in A_{\alpha_0}$, $A(lE(x_0)+(1-l)E(y_0))<\alpha_0$.
Therefore, $lE(x_0)+(1-l)E(y_0)\notin A_{\alpha_0}$
On the other hand, A_{α_0} is a $E-$convex set,
and we have $lE(x_0)+(1-l)E(y_0)\in A_{\alpha_0}$,
So $A(lE(x_0)+(1-l)E(y_0))\geq\alpha_0$,
a contradiction. Therefore, $\forall x,y\in X$, we have
$A(lE(x)+(1-l)E(y))\vee\lambda\geq A(x)\wedge A(y)\wedge\mu.$
That is, A is a generalized $E-$convex fuzzy set.

Theorem 3.2 Let $A,B\in F(X)$, and they are both generalized $E-$convex fuzzy sets. Then $A\bigcap B$ is also a generalized $E-$convex fuzzy set.

Proof. $(A\bigcap B)(lE(x)+(1-l)E(y))\vee\lambda$
$=[A(lE(x)+(1-l)E(y))\wedge B(lE(x)+(1-l)E(y)]\vee\lambda$
$=[A(lE(x)+(1-l)E(y))\vee\lambda]\wedge[B(lE(x)+(1-l)E(y))\vee\lambda]$
$\geq(A(x)\wedge A(y)\wedge\mu)\wedge(B(x)\wedge B(y)\wedge\mu)$
$=(A(x)\wedge B(x))\wedge(A(y)\wedge B(y))\wedge\mu$
$=(A\bigcap B)(x)\wedge(A\bigcap B)(y)\wedge\mu$
From the Definition of 3.1 $A\bigcap B$ is also a generalized $E-$convex fuzzy set.

Corollary 3.1 Let $A_i(i=1,2,...,n)$ be generalized $E-$convex fuzzy sets, then
$\bigcap_{i=1}^{\infty}A_i$ is also a generalized $E-$convex fuzzy set.

Theorem 3.3 Let $A\in F(X)$ be a generalized E_1- and E_2-convex fuzzy set.
Then A is a generalized $(E_1\circ E_2)$ and $(E_2\circ E_1)-$convex fuzzy set.

Proof. Assume that $x,y\in X$,
Then
We have $A(lE_1(x)+(1-l)E_1(y))\vee\lambda$
$\geq A(x)\wedge A(y)\wedge\mu$
$A(lE_2(x)+(1-l)E_2(y))\vee\lambda$
$\geq A(x)\wedge A(y)\wedge\mu$
So $A[l(E_1\circ E_2)x+(1-l)(E_1\circ E_2)y]\vee\lambda$
$=\{A[lE_1(E_2(x))+(1-l)E_1(E_2(y))]\vee\lambda\}\vee\lambda$
$\geq[A(E_2(x))\wedge A(E_2(y))\wedge\mu]\vee\lambda$
$=(A(E_2(x))\vee\lambda)\wedge(A(E_2(y))\vee\lambda)\wedge\mu$
$\geq A(x)\wedge A(y)\wedge\mu$
Therefore,
A is a generalized $(E_1\circ E_2)-$convex fuzzy set.
Similarly,
A is a generalized $(E_2\circ E_1)-$convex fuzzy set.

4. $E-$ CONVEX FUNCTION

Definition 4.1 A function $f:R^n\to R$ is said to be generalized $E-$convex on a set $C\subset R^n$ if there is a map $E:R^n\to R$ such that C is a generalized $E-$convex set and $f(lE(x)+(1-l)E(y))\vee\lambda$
$\geq f(E(x))\wedge f(E(y))\wedge\mu,$ for each $x,y\in C$ and $0\leq\lambda<\mu\leq1,0\leq l\leq1.$

Remark 4.1 Obviously, when $\lambda=0$ and $\mu=1,f$ is an ordinary $E-$convex function.

Definition 4.2 Let $S\subset R^n\times R$ and $E:R^n\to R$. A set S is said to be $E-$convex if $(x,\alpha),(y,\beta)\in S$
Imply
$(lE(x)+(1-l)E(y),\alpha\wedge\beta)\in S,0\leq l\leq1.$

Definition 4.3 Define an $E-$epigraph $E-e(f)$ of f as follows:
$E-e(f)=\{(x,\alpha):x\in C,\alpha\in(\lambda,\mu],\alpha\leq f(E(x))\leq\mu\}.$

Lemma 4.1 If $E(C)$ is a convex subset of C, and then C is a $E-$convex set.

Proof. Let $x,y\in C$, then $E(x),E(y)\in E(C).$
Since $E(C)$ is a convex subset of C, $lE(x)+(1-l)E(y)\in C$
Hence, C is a $E-$convex set.

Theorem 4.1 Let $E(C)$ be a convex subset of C, and $E(x)=x,f(E(x))\leq\mu$ when $x\in E(C)$. Then a numerical function f defined on an $E-$convex set $C\subset R^n$ is $E-$convex on C if $E-e(f)$ is $E-$convex in $R^n\times R.$

Proof.
From Lemma 4.1, C is $E-$convex.
If f is an $E-$convex function on C, then for any $(x,\alpha),(y,\beta)\in E-e(f),0\leq l\leq1,$
We have $f(E(lE(x)+(1-l)E(y)))\vee\lambda$
$=f(lE(x)+(1-l)E(y))\vee\lambda$
$\geq f(E(x))\wedge f(E(y))\wedge\mu$
$\geq\alpha\wedge\beta\wedge\mu=\alpha\wedge\beta$
Since $\lambda<\alpha,\lambda<\beta,$ then $\lambda<\alpha\wedge\beta,$
so $\mu\geq f(E(lE(x)+(1-l)E(y)))\geq\alpha\wedge\beta$
Therefore,
We have

$(lE(x)+(1-l)E(y),\alpha\wedge\beta)\in E-e(f)$, that is, $E-e(f)$ is $E-$ convex set.

Conversely, assume that f is not a $E-$ convex function on C, and then there exist

$x_0,y_0\in C,0\le l\le 1,\alpha_0=f(E(x_0)),\beta_0=f(E(y_0))$

Such that $f(lE(x_0)+(1-l)E(y_0))\vee\lambda$

$<\alpha_0\wedge\beta_0=f(E(x_0))\wedge f(E(y_0))\wedge\mu$

So,

$\alpha_0,\beta_0\in(\lambda,\mu],\ f(lE(x_0)+(1-l)E(y_0))<\alpha_0\wedge\beta_0,$

that is, $f(E(lE(x_0)+(1-l)E(y_0)))$

$<\alpha_0\wedge\beta_0.$

Therefore,

$(lE(x_0)+(1-l)E(y_0),\alpha\wedge\beta)\notin E-e(f)$. On the other hand,

$(E(x_0),f(E(x_0))),(E(y_0),f(E(y_0)))\in E-e(f)$

, since $E-e(f)$ is $E-$ convex,

$(lE(E(x_0))+(1-l)E(E(y_0)),f(E(x_0))\wedge f(E(y_0)))\in E-e(f),$

that is $(lE(x_0)+(1-l)E(y_0),\alpha\wedge\beta)\in E-e(f),$

a contradiction.

Thus, f is a $E-$ convex function on C.

Remark 4.2 A map E is not necessarily an identity map.

Corollary 4.1 Let $E(C)$ be a convex set on C and $E(x)=x,\ f(E(x))\le\mu,$ when $x\in E(C),$ then f is an $E-$ convex function on C iff the inequality

$f(lE(x)+(1-l)E(y))\ge\alpha\wedge\beta,0\le l\le 1,\alpha,\beta\in(\lambda,\mu]$

holds when $\alpha\le f(E(x))\le\mu,\ \beta\le f(E(y))\le\mu.$

Proof. From Lemma 4.1, C is $E-$ convex.

If f is an $E-$ convex function on C, then, when $\alpha\le f(E(x))\le\mu,\ \beta\le f(E(y))\le\mu,$ $\alpha,\beta\in(\lambda,\mu]$, we have $f(lE(x)+(1-l)E(y))\vee\lambda$

$\ge f(E(x))\wedge f(E(y))\wedge\mu$

$\ge\alpha\wedge\beta\wedge\mu=\alpha\wedge\beta.$

Since $\lambda<\alpha,\lambda<\beta,$ so $\lambda<\alpha\wedge\beta$. Thus,

$f(lE(x)+(1-l)E(y))\ge\alpha\wedge\beta,0\le l\le 1.$

Conversely, assume that f is not an $E-$ convex function on C.

From the theorem 4.1, $E-e(f)$ is not $E-$ convex. Then,

there exist $(x_0,\alpha_0),(y_0,\beta_0)\in E-e(f)$

$,0\le l\le 1,$ such that

$(lE(x_0)+(1-l)E(y_0),\alpha_0\wedge\beta_0)\notin E-e(f).$

On the other hand,

$(x_0,\alpha_0),(y_0,\beta_0)\in E-e(f),$ so

$\alpha_0\le f(E(x_0))\le\mu,\ \ \beta_0\le f(E(y_0))\le\mu.$

Then $f(lE(x_0)+(1-l)E(y_0))\ge\alpha_0\wedge\beta_0,$

That is

$(lE(x_0)+(1-l)E(y_0),\alpha_0\wedge\beta_0)\in E-e(f),$

a contradiction.

Thus, f is a $E-$ convex function on C.

Theorem 4.2 Let $E(C)$ be convex subset of C, and we define $f(E(x))=\sup\{\mu\mid(x,\mu)\in C\times R\}$

When $x\in E(C),E(x)=x,\ f(E(x))\le\mu.$

Then f is a $E-$ convex function on C.

Proof. From Lemma 4.1 C is $E-$ convex. For $x,y\in C,$ there exist $\alpha,\beta\in R,$ which satisfy the following: $\alpha\le f(E(x))\le\mu,\ \beta\le f(E(y))\le\mu,$ based on the definition of $f(E(x)),$

there exist $\xi,\eta\in R,\xi>\alpha,\eta>\beta$

such that $(x,\xi)\in C\times R,(y,\eta)\in C\times R.$

Since $C\times R$ is $E-$ convex,

so $(lE(x)+(1-l)E(y),\xi\wedge\mu)\in C\times R$

Then, $f(lE(x)+(1-l)E(y))$

$=f(E(lE(x)+(1-l)E(y)))\ge\xi\wedge\eta\ge\alpha\wedge\beta$

From the corollary 4.1,

f is a $E-$ convex function on C.

5. REFERENCES

[1] A.K. Katsaras, D.B. Liu, "Fuzzy vector spaces and fuzzy topological vector spaces [J]", Journal of Mathematical Analysis and Applications. , 1977,58:135-146

[2] X. Chen, "Some properties of semi $-E-$ convex functions[J]", Journal of Mathematical Analysis and Applications, 2002, 275:251-262.

[3] H. Gu, Z.H. Liao, " Convex fuzzy sets [J]", Fuzzy System and Mathematics, 2004,18:153-156.

[4] J.G. Brown, "A note on fuzzy sets [J]", Information and Control, 1971,18:32-39.

[5]L.A. Zadeh, "Fuzzy Sets[J]", Information and Control, 1965,8: 338-353

[6] M.D. "Weiss Fixed points, separation and induced topologies for fuzzy sets [J]", Journal of Mathematical Analysis and Applications, 1975,50:142-150

[7] R. Lowen, Convex fuzzy sets [J], Fuzzy Sets and System, 1980,3:291-310

[8] S.K. Bhakat, "On the Definition of a Fuzzy Subgroup [J]", Fuzzy Sets and Systems, 1992,51:235-241.

[9] X.H. Yuan, C. Zhang, et al, "Generalized fuzzy subgroups and many-valued implications [J]", Fuzzy sets and Systems, 2003,138:205-211.

[10] X. M. Yang, "A note on convex fuzzy sets [J]", Fuzzy Sets and Systems, 1993,53:117-118.

[11] X. M. Yang, F. M. Yang, "A property on convex fuzzy sets [J]", Fuzzy Sets and Systems, 2002,126:269~271.

[12] X. M. Yang, "Relationship among three types of convex fuzzy sets [J]", BUSEFAL, 1988,35:53-59.

[13] E.A. Youness, $E-$ Convex Sets, $E-$ Convex Functions, and $E-$ Convex Programming [J]". Journal of

Optimization Theory and Applications, 1999,102:439~450.
[14] Y. M. Liu, “Some properties of convex fuzzy sets [J]”, Journal of Mathematical Analysis and Applications, 1985,111:119-129.

Research on Improvement of B&B Algorithm Performance of Single Machine Total Tardiness Schedule

Xiaochuan Luo
Key Laboratory of Process Industry Automation (Northeastern University)
Shenyang, Liaoning 110004, China
Email: Luoxiaochuan@ise.neu.edu.cn

And

Dancheng LI, Hang Song
Software College, Northeastern University
Shenyang, Liaoning 110004, China
Email: danchengli@126.com , hang_song2003@163.com

ABSTRACT

This paper addresses the schemes to improve B&B algorithm performance for the NP-hard problem of scheduling *N* jobs on a single machine with due dates, sequence-dependent setup times, and no preemption, where the objective is to minimize the total tardiness. Some dominance rules are proposed to reduce the expanded nodes and explored nodes. A comparison with Ragatz's B&B approaches indicates that the algorithm, which adopts our schemes, is competitive and has a certain advantage for larger problems.

Keywords: Dominance Rule, Algorithm Performance, B&B Algorithm, Total Tardiness, Single Machine Scheduling.

1. INTRODUCTION

In this paper, we address the problem of finding an optimal solution of *N* jobs with sequence-dependent setup times to minimize Total Tardiness (TT for short). Sequence-dependent setup times scheduling problems are often encountered in many production systems [1].

One of the earliest works on minimizing tardiness in a sequence dependent environment was from Rinnooy Kan el al. [2]. Recently, Rabadi et al. [3] presented a branch-and-bound algorithm for the single-machine early/tardy scheduling problem with sequence-dependent setup times whose earliness and tardiness are weighted equally and the due date is common and large for all jobs. Ragatz [4] proposed a B&B scheme for minimizing total tardiness on a single processor with sequence dependent setup times.

The objective of this paper is to present some dominance rules and evaluates our branch-and-bound algorithm for TT problem.

2. PROBLEM DESCRIPTION

2.1 Notation

In the remainder of the paper, we make use of the following notation.

1) Indices and sets: N : number of jobs; J : set of all jobs, $J=\{1,2,3,...N\}$; $J_0 = J \cup 0$: extended set of jobs, including a dummy job denoted by 0; i, j, k, l : job indices, $i, j, k, l \in J_0$; S : sequence beginning with the dummy job 0, which is a solution of the problem; K : partial sequence beginning with the dummy job 0, corresponding to a node at level k in the research tree, in which jobs in the first $k+1$ positions have been fixed.

2) Input data: p_i, d_i : processing time and due date of job i ; c_{ij} : cost from job i to job j ; s_{ij} : setup time when job i is scheduled right before job j in a scheduling sequence S .

3) Computed parameters: C_i : finished time of job i ; T_i : tardiness of job i ; T_t : total tardiness of all the jobs.

A job i (without brackets) refers to the job i itself, whereas job $[j]$ (with brackets) refers to the index of the job scheduled in the j^{th} position.

2.2 Assumptions

1) All jobs become available for machine processing simultaneously.

2) All processing times $p_i, i \in J$, are deterministic and known before processing starts.

3) Job preemption and job splitting are not permitted.

4) The machine cannot simultaneously process two or more jobs.

5) At time zero, the machine is processing dummy job 0, with $p_0 = 0$, $d_0 = 0$, $C_0 = 0$. Setup time $s_{0i}, i \in J$, depends on the setup cost from job 0 to job i . Job 0 and job [0] refer the same dummy job.

6) $s_{ij} + s_{jk} \geq s_{ik}$.

2.3 Formulation of Problem

The problem dealt with in this paper can be described as follows. Suppose there are N jobs indexed from 1 to N , and for each job $j, j \in J$, there are such

* Supported by National Nature and Science Foundation of China (70571077)

parameters as processing time p_j, due date d_j, and setup time s_{ij}.

Let $S=\{[0],[1],[2],...,[N]\}$ be a scheduling sequence of all the jobs. The tardiness of the i^{th} job in sequence S is

$$T_{[i]}=\max\{C_{[i]}-d_{[i]},0\} \quad (1)$$

where $C_{[i]}$ is the completion time of job $[i]$,

$$C_{[i]}=\begin{cases}0 & i=0\\ \sum_{k=1}^{i}(s_{[k-1][k]}+p_{[k]}) & i\neq 0\end{cases}$$

We would like to find a sequence that minimizes the total tardiness of the jobs. Total tardiness for a sequence S is

$$T_t=\sum_{i\in\{0,1,2,...,N\}}T_{[i]} \quad (2)$$

Tardiness is known to be a regular performance measure, which means that its value can be increased only by increasing the completion time of a job. To minimize a regular measure, it is sufficient to consider only permutation schedules.

3. DOMINANCE RULES

In this section, we present two dominance rules. Good dominance rules could lead to substantial improvements in algorithmic performance. By changing the job positions in sequence *S*, we get a new sequence *S'* and compare the two sequences costs to decide which sequence is dominant.

It is obvious that inserting idle time in any schedule can not improve the optimal cost. It is sufficient to consider schedules without idle time inserted. Before presenting the dominance rules, we use the following notations in this Section.

$J(K)$: set of jobs in the partial schedule K. $U(K)$: set of unscheduled jobs. $C(K)$: Completion time of the last job in K. $K|u$: New partial schedule obtained by adding job u behind the given partial schedule K. $\sum(K|u)$: Schedule composed of $K|u$, completed by the partial optimal schedule of jobs, which belong to $J-J(K|u)$, starting from the moment $C(K|u)$. $|\cdot|_{[i][j]}$: Number of jobs from position i to position j, including the two jobs at i and j position.

After changing positions of some jobs in sequence S to construct a new sequence S', $[i]$ still refers to the job index of the i^{th} position in S. The tardiness of jobs in S and S' are denoted as $T_{[i]}$ and $T_{[i]}'$ ($i\in J$) respectively. Tt_i or Tt_i' refer to total tardiness of jobs in part i of S or S' for $i=0,1,2$.

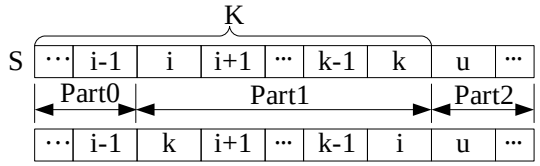


Fig. 1. Interchanging of positions of jobs [*i*] and [*k*]

Theorem 1: If $\exists i,[i]\in J(K),\Delta_1=s_{[i-1][k]}+p_{[k]}+s_{[k][i+1]}-s_{[i-1][i]}-p_{[i]}-s_{[i][i+1]}\le 0,|\cdot|_{[i][k]}=n_1,|\cdot|_{u,end}=n_2,\Delta_2=s_{[i-1][k]}+s_{[k][i+1]}+s_{[k-1][i]}+s_{[i]u}-s_{[i-1][i]}-s_{[i][i+1]}-s_{[k-1][k]}-s_{[k]u}\le 0,\ \Delta_T=(n_1-1)\Delta_1+s_{[i-1][k]}+s_{[k-1][i]}-s_{[i-1][i]}-s_{[k-1][k]}+n_2\Delta_2\le 0$, then $\sum(K|u)$ is dominated.

Proof: Consider the schedule $S=\sum(K|u)$. We construct another schedule S' by exchanging the positions of jobs $[i]$ and $[k]$ (See Fig. 1). The schedule S' has three parts.

In the Part 0: $Tt_0=Tt_0',C_0=C_0'$.

In the Part 1:

The tardiness of job $[l],[i+1]\le l\le[k-1]$, has a change of Δ_1, so the total change from job $[i+1]$ to job $[k-1]$ is $(n_1-2)\Delta_1$. With respect to job $[i]$ and $[k]$, the change of tardiness is

$$T_{[i]}'-T_{[i]}+T_{[k]}'-T_{[k]}=\Delta_1+s_{[i-1][k]}+s_{[k-1][i]}-s_{[i-1][i]}-s_{[k-1][k]}.$$

Therefore, the total change of tardiness in Part 1 is

$$Tt_1'-Tt_1=(n_1-1)\Delta_1+s_{[i-1][k]}+s_{[k-1][i]}-s_{[i-1][i]}-s_{[k-1][k]}.$$

In the Part 2:

Because $\Delta_2\le 0$, $C_u\ge C_u'$. Therefore, any sequence comprised of all the jobs in $U(K)$ can be scheduled earlier in S'. That is $Tt_2'-Tt_2=n_2\Delta_2$.

The total change of tardiness of S' is

$$\Delta_T=Tt_0'+Tt_1'+Tt_2'-Tt_0-Tt_1-Tt_2.$$

Finally, we have

$$\Delta_T=(n_1-1)\Delta_1+s_{[i-1][k]}+s_{[k-1][i]}-s_{[i-1][i]}-s_{[k-1][k]}+n_2\Delta_2$$

Because $\Delta_T\le 0$, $\sum(K|u)$ is dominated.

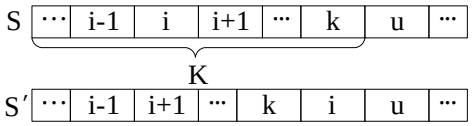


Fig. 2. Inserting job [*i*] after job [*k*]

Theorem 2: if $\exists i,[i] \in J(K), |\cdot|_{[i][k]} = n_1, |\cdot|_{u,end} = n_2, \Delta_1 = s_{[i-1][i+1]} - s_{[i-1][i]} - s_{[i][i+1]} - p_{[i]} \le 0, \Delta_2 = s_{[i-1][i+1]} + s_{[k][i]} + s_{[i]u} - s_{[i-1][i]} - s_{[i][i+1]} - s_{[k]u} \le 0, n_1\Delta_1 + n_2\Delta_2 + C_k - C_{i-1} \le 0,$ then $\sum(K \mid u)$ is dominated.

Proof: Consider the schedule $S = \sum(K \mid u)$. We construct another schedule S' by putting job $[i]$ after job $[k]$ and before job u (See Fig. 2).

From the assumption of the triangular inequality, we have $s_{[i-1][i+1]} \le s_{[i-1][i]} + s_{[i][i+1]} + p_{[i]}$, then we can get $C_{[i+1]} \ge C'_{[i+1]}, \cdots, C_{[k-1]} \ge C'_{[k-1]}, \ C_{[k]} \ge C'_{[k]}$. Also, we have $T_{[i+1]} \ge T'_{[i+1]}, \cdots, T_{[k-1]} \ge T'_{[k-1]}$, and $T_{[k]} \ge T'_{[k]}$. The change value of each job's tardiness is Δ_1.

Consider the jobs in $U(K)$, because $\Delta_2 \le 0$, we have $C_u \ge C'_u$. Consequently, any unscheduled job sequence beginning with job u can be processed earlier in S'. The change value of each job's tardiness is Δ_2. The change of total tardiness is $\Delta_T = (n_1 - 1)\Delta_1 + n_2\Delta_2 + T'_{[i]} - T_{[i]}$.

Because $T'_{[i]} - T_{[i]} = C_k + \Delta_1 - C_{i-1}$, $\Delta_T = n_1\Delta_1 + n_2\Delta_2 + C_k - C_{i-1}$.

According to $n_1\Delta_1 + n_2\Delta_2 + C_k - C_{i-1} \le 0$, we have that $\sum(K \mid u)$ is dominated.

4. COMPUTATION EXPERIMENTS

There are two control parameters to generate instances for the problems with due dates: due date range (DDR) and tardiness factor (TF). In our experiments, the numerical experiments used the set of test problem devised by Ragatz [4] in which five parameters to generate test problems were considered: 1. Number of jobs, N; 2. processing time variance, PTV; 3. Setup times range, SR; 4. tardiness factor, TF; 5. due dates range, DDR.

Table 1. Control parameters of generating random instances

Size	TF	DDR	PTV	SR	number
{10,14,18,22,30}	{0.2,0.4,0.6,0.8}	{0.2,0.9}	{25,625}	{5,19}	5

Total 5×160 instances were randomly generated based on the control parameters sets given in Table 1. We completed our experiments by testing our algorithm on a computer having an Intel Pentium II 600Mhz processor with 256 Meg RAM, running under Windows 2000. The algorithm was coded in Microsoft Visual C++ 6.0.

We show the results in Table 2 when our algorithm was applied to all the random instances. The maximum CPU time of each instance was set at 15 Minutes. Table 2 displays the summary statistics, which are calculated from 160 instances for each job size with different combinations of DDR, TF, PTV and SR. As can be seen, all the instances of 10- and 14-job sizes were solved in an average time of less than 1 second. With the increasing of job size, the solved instance percentage was decreased. This is to be expected since the upper bound, lower bound and dominance rules become less powerful as the number of job sizes increases. The limit of CPU times can explain that the number of nodes of 30- and 26-job instances is less than of 22-job instances. We found that the solved instance percentages were rather good, even for 30-job instances. Maybe, the control parameters generated test instances, which were solved easily for our algorithm.

Table 2. The overall performance of algorithm

Size	Time(*s*)	Expanded Node	Explored Node	Stored Node	Solved(%)
10	0.40596	48.106	5.05	2.6875	100
14	0.98856	204.96	17.788	7.0875	93.75
18	13.176	3571	332.49	98.969	80.10
22	43.998	6555.3	550.75	178.08	68.23
30	100.83	3785.1	327.18	202.84	47.50

We compare our algorithm with Ragatz's algorithm in terms of the average of expanded nodes as Table 3. Because Ragatz's algorithm experiments just only used 10-, 12-, 14- and 16-job instances, we also tested our algorithm using the same size instances. We conclude that our algorithm reduces the expanded node considerable because of good performances of lower bound, upper bound and dominance rules.

Table 3. Comparison with the average of expanded nodes

Size	10	12	14	16
Ours	48.106	166.02	204.96	873.24
Ragatz's	539	1800	2805	33916

5. CONCLUSIONS

The single machine scheduling problem with sequence dependent setup times for total tardiness minimization is NP-hard. We have found that the branch-and-bound algorithm that we have presented is able to efficiently solve problems of a size encountered in industrial situations. The use of the dominance rules developed here improves the performance of the branch and bound algorithm substantially. Significantly better performance over previous published works was also obtained. The computational times to obtain optimal solution by our algorithm are appreciably lower.

6. REFERENCES

[1] C.Y. Liu, S.C. Chang, "Scheduling flexible flow shops with sequence-dependent setup effects", IEEE Transactions on Robotics and Automation, Vol.16, No.4, 2000, pp. 408-419.

[2] A.H.G. Rinnooy, B..J. Lageweg, et al, "Minimizing total costs in one machine scheduling", Operations

Research, Vol.23, No.5, 1975,pp. 908-927.

[3] G. Rabadi, M. Mollaghasemi, et al, "A branch-and-bound algorithm for the early/tardy machine scheduling problem with a common due-date and sequence-dependent setup time", Computers & Operations Research, Vol.31, 2004, pp. 1727-1751.

[4] G.L. Ragatz, "A branch-and-bound method for minimum tardiness sequencing on a single processor with sequence dependent setup times", Proceedings of twenty-fourth annual meeting of the Decision Sciences Institute, San Francisco, 1993, pp.375-1377.

Qos Multicast Routing Algorithms for Mobile Ad Hoc Networks Based On Ga

Fubao[1] Yang, Zeng[2] Yue, Baolin[3] Sun
[1.] School of Information Engineering, Wuhan University of Technology, Wuhan 430063, P.R. China;
[2.] Department of Computer, Jiangxi Vocational College of Finance and Economics, Jiujiang, Jiangxi 332000, P. R. China;
[3.] Department of Mathematics and Physics, Wuhan University of Science and Engineering, Wuhan 430073, P.R. China

ABSTRACT

Mobile Ad hoc Networks (MANET) are termed as mobile distributed multi-hop wireless networks without predetermined topology or central control. Most of the multimedia applications require strict QoS guarantee during the communication between a single source and multiple destinations. This gives rise to the need for an efficient QoS multicast routing strategy. Determination of such QoS-based optimal multicast routes basically leads to a multi-objective optimization problem, which is computationally intractable in polynomial time due to the uncertainty of resources in mobile ad hoc networks. An entropy-based model to support QoS multicast routing algorithm for MANET is presented, which based on Genetic Algorithm (QMRGA). The simulated results show that the proposed approach and parameters provide a correct and effective method to evaluate the route stability in dynamic mobile networks.

Keywords: Quality-of-Service (QoS), Multicast routing, Ad hoc networks, Genetic algorithm, Entropy.

1. INTRODUCTION

Many applications envisioned for mobile ad-hoc networks rely on group communication. Communication during disaster relief, emergency search, battlefields, and data acquisition in remote areas and emergency warnings in vehicular networks are common examples for these applications. As a consequence multicast, routing in mobile ad-hoc networks has received significant attention recent years.

Multicast services have been used by various continuous media applications. For example, the multicast backbone (Mbone) of the Internet has been used to transport real time audio/video for news, entertainment, video conferencing, and distance learning [1…6]. The provision of Quality-of-Service (QoS) guarantees is of utmost importance for the development of the multicast services [1,4…6]. Multicast routing has continued to be a very important research issue in the areas of networks and distributed systems. It attracts the interests of many people.

QoS multicast routing depends on state parameters, which specifying resource availability at network nodes or links, and uses the parameters to determine paths with enough free resources [1…6]. In turn, the successful routing of new flows together with the termination of existing ones, induce constant changes in the amount of resources available. These must then be communicated back to QoS multicast routing. Unfortunately, communicating such changes in a timely fashion is expensive and, at times, not even feasible [1…6]. As a result, changes in resources availability are usually communicated either infrequently or uncertainly. There are two main components to the cost of timely distribution of changes in network state: the number of entities generating such updates and the frequency at which each entity generates updates.

Entropy [7, 8] presents the uncertainty and a measure of the disorder in a system. There are some common characteristics among self-organization, entropy, and the location uncertainty in mobile ad hoc wireless networks. The corresponding methodology, results and observations can be used to select the most stable route between source and destination in an environment using routing protocols, where multiple paths are available. In addition, a convenient performance measure is created to be used for the evaluation of the stability and connectivity in mobile ad hoc networks.

In this paper, an entropy-based model to support QoS multicast routing algorithm for MANET is presented, which based on Genetic Algorithm (QMRGA). The basic motivations of the proposed modeling approach stem from the commonality observed in the location uncertainty in mobile ad hoc wireless networks and the concept of entropy. The purpose of this paper is to develop an algorithm to find out QoS-based multicast routes by simultaneously optimizing end-to-end delay, entropy, and bandwidth provisioning for guaranteed QoS and proper bandwidth utilization without combining them into a single scalar optimization function.

The rest of the paper is organized as follows. Section 2 introduces a QoS-based multicast routing network model. Section 3 describes entropy metric in ad hoc network. Section 4 presents the QoS Multicast Routing algorithms based on Genetic Algorithm (QMRGA). Analysis of convergence and some simulation results are provided in Section 5. The conclusions and future research are in Section 6.

2. ENTROPY METRIC

We associate each node m with a set of variable features denoted by $a_{m,n}$, where node n is a neighbor of node m. In this paper, two nodes are considered neighbors if they can reach each other in one hop. These variable features $a_{m,n}$ represent a measure of the relative speed between two nodes and are defined rigorously later in this section. Any change of the system can be described as a change of variable values $a_{m,n}$ in the course of time t such as $a_{m,n}(t) \to a_{m,n}(t+\Delta_t)$. Let's denote by $v(m,t)$ the velocity vector of node m and by $v(n,t)$ the velocity vector of node n at time t. Pay attention to that velocity

* The research is supported by CSCT (2005BB2212), China.

vectors $v(m,t)$ and $v(n,t)$, which have two parameters, namely speed and direction. The relative velocity $v(m,n,t)$ between nodes m and n at time t is defined as:

$$v(m,n,t) = v(m,t) - v(n,t) \quad (1)$$

Let us also denote by $p(m,t)$ the position vector of node m and by $p(n,t)$ the position vector of node n at time t. The relative position $p(m,n,t)$ between nodes m and n at time t is defined as:

$$p(m,n,t) = p(m,t) - p(n,t) \quad (2)$$

Then, the relative mobility between any pair (m,n) of nodes during some time interval is defined as their absolute relative speed and position averaged over time. Therefore, we have:

$$a_{m,n} = \frac{1}{N}\sum_{i=1}^{N} \frac{|\, p(m,n,t_i) + v(m,n,t_i \times \Delta_{t_i}) \,| - |\, p(m,n,t_{i+1}) \,|}{R}$$

(3) Where N is the number of discrete times t_i that velocity information can be calculated and disseminated to other neighboring nodes within time interval Δ_t. R is radio range of nodes. Based on this, we can define the entropy $H_m(t,\ \Delta_t)$ at mobile during time interval Δt. The entropy can be defined either within the whole neighboring range of node (e.g., within set S_m), or for any subset of neighboring nodes of interest. In general the entropy $H_m(t,\ \Delta_t)$ at mobile is calculated as follows:

$$H_m(t,\ \Delta_t) = \frac{-\sum_{k\in F_m} P_k(t,\Delta_t)\log P_k(t,\Delta_t)}{\log C(F_m)} \quad (4)$$

Where $P_k(t,\Delta_t) = (am, k / \sum_{k\in F_m} a_{m,i})$

In this relation by F_m we denote the set (or any subset) of the neighboring nodes of node m, and by $C(F_m)$ the cardinality (degree) of set F_m. If we want to calculate the local network stability (with reference to node m), then Fm refers to the set that includes all the neighboring nodes of mobile node m (e.g., $F_m = S_m$), while if we are interested in the stability of a part of a specific route then F_m represents the two neighboring nodes of mobile node m over that route. As can be observed from the previous relation the entropy $H_m(t,\Delta_t)$ is normalized so that $0 \le Hm(t,\ \Delta_t) \le 1$. It should be noted that the entropy, as defined here, is small when the change of the variable values in the given region is severe and large when the change of the values is small [7,8]. Let us present the route stability (RS) between two nodes s and $d \in U$ during some interval Δ_t as RS. We also define and evaluate two different measures to estimate and quantify end to end route stability, denoted by $F'(s,\ d)$ and $F(s,\ d)$ and defined as follows respectively:

$$F(s,\ d) = \prod_{i=1}^{N_r} H_i(t,\Delta_t) \quad (6)$$

Where N_r denotes the number of intermediate mobile nodes over a route between the two end nodes $(s,\ d)$.

$$F(s,\ d) = -\ln F'(s,d) = -\sum_{i=1}^{N_r} \ln H_i(t,\Delta_t) \quad (7)$$

3. NETWORK MODEL

A network is usually represented as a weighted digraph $G = (N,\ E)$, where N denotes the set of nodes and E denotes the set of communication links connecting the nodes. $|N|$ and $|E|$ denote the number of nodes and links in the network respectively. In $G(N,\ E)$, considering a QoS constrained multicast routing problem from a source node to multi-destination nodes, namely given a non-empty set $M = \{s,\ u_1,\ u_2,\ \cdots,\ u_m\}$, $M \subseteq N$, s is source node, $U = \{u_1,\ u_2,\ \cdots,\ u_m\}$ be a set of destination nodes. In multicast tree $T = (N_T,\ E_T)$, where $N_T \subseteq N$, $E_T \subseteq E$. If $C(T)$ is the cost of T, $P_T(s,u)$ is the path from source node s to destination $u \in U$ in T, $B_T(s,u)$ is usable bandwidth of $P_T(s,u)$.

Definition 1: The cost of multicast tree T is:

$$C(T_e) = \sum_{e\in E_T} C(e),\ e \in E_T. \quad (8)$$

Definition 2: The bandwidth of multicast tree T is the minimum value of link bandwidth in the path from source node *s* to each destination node $u \in U$. i.e.

$$B_T(s,\ u) = \min(B(e),\ e \in E_T). \quad (9)$$

Definition 3: The delay of multicast tree T is the maximum value of delay in the path from source node *s* to each destination node $d \in U$. i.e.

$$D_T(s,\ d) = \max(\sum_{e\in P_T(n_0,d)} D(e), d \in U) \quad (10)$$

Definition 4: The stability route of multicast tree T is the minimum value of entropy metric in the path from source node *s* to each destination node $d \in U$. i.e.

$$F_T(s,\ u) = \min(F(e),\ e \in E_T). \quad (11)$$

Definition 5: Assume the minimum bandwidth constraint of multicast tree is B, given a multicast demand R, then, the problem of bandwidth constrained multicast routing is to find a multicast tree *T*, satisfying: $B_T(s,\ u) \ge B$, $u \in U$. Suppose $S(R)$ is the set, $S(R)10$ satisfies the conditions above, then, the multicast tree *T* which we find is:

$$C(T) = \min(C(T_S),\ T_S \in S(R)) \quad (12)$$

4. GENETIC OPTIMIZATION ALGORITHM

GAs is based on the mechanics of natural evolution. Throughout their artificial evolution, successive generations each consisting of a population of possible solutions, called individuals (or chromosomes, or vectors of genes), search for beneficial adaptations to solve the given problem [1,3,9]. Applying the Darwinian principles of "reproduction and survival of the fittest" and the genetic operators of crossover and mutation, which derive the new offspring population from the current population, carries out this search. Reproduction involves selecting, in proportion to its fitness level, an individual from the current population and allowing it to survive by copying it to the new population of

individuals. The individual's fitness level is usually based on the cost function given by the problem (e.g., QoS multicast routing) under consideration. Then, crossover and mutation are performed on two randomly chosen individuals of the current population to create two new offspring individuals. Crossover involves swapping two randomly located sub-chromosomes (within the same boundaries) of the two mating chromosomes. Mutation is applied to randomly selected genes, where the values associated with such a gene is randomly changed to another value within an allowed range. The offspring population replaces the parent population, and the process is repeated for many generations. Typically, the best individual that appeared in any generation of the run (i.e. best-so-far individual) is designated as the result produced by the genetic algorithm.

The function QoS based multicast route discovery takes the source node v_s and a specific number of multicast destination nodes, say, v_{d1}, v_{d2},⋯, v_{dn} as input, called the function of path finding to find all possible multicast paths from v_s to each of v_{d1}, v_{d2},⋯, v_{dn}, using the basic depth first search (dfs) algorithm. This gives birth to the initial set of multicast trees. The primary objective of our algorithm is to find the multicast trees from this set, which will satisfy the multiple constrained QoS parameters.

Since the underlying approach is based on Multi-Objective GA (MOGA), our next step is to map the problem in a search space suitable to MOGA. Each of all the generated multicast trees is mapped to a string consisting of the sequence of nodes along the path from the source v_s to each of the destinations v_{d1}, v_{d2},⋯, v_{dn}, To mark the end of a path from a source to a single destination, we use -1 as sentinel. Fig. 1 below gives a clear view of this scenario where a multicast tree is represented by a string. The set of all such strings constitute the initial population. The size of this population depends on how the strings are created, which in turn depends on the network topology and the number of multicast destination nodes.

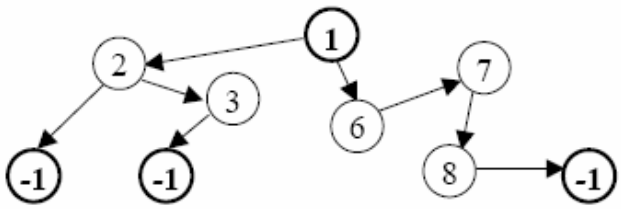


Fig. 1. Representation of the multicast tree and its encoding scheme

5. ANALYSIS OF CONVERGENCE AND SIMULATION EXPERIMENTS

5.1. Analysis of convergence

Theorem 2. The genetic algorithm proposed in this paper converges to the global optimal solution.

Proof. The genetic algorithm has following merits: (1) The method, which uses the candidate routing set from source node to each destination node, makes the searching speed faster, and the whole process could be done in shorter time; (2) selecting by proportion and pertaining the optimal individual before selection; (3) Changeable length chromosome encoding method which based on routing expression is used; (4) Crossover probability between [0,1]; (5) Mutation probability between [0,1], by theorem 2.7 in Ref. [1,3,9]: have crossover probability between [0,1], mutation probability between (0,1). At the same time, the genetic algorithm obtained according to the method proposed above, this could converge to the global optimal solution. Hence, the genetic algorithm proposed in this paper can converge to the global optimal solution.

5.2. Experiments Environment

We conduct simulation experiments to evaluate QMRGA. QMRGA is implemented by using the network simulator and its performance is compared with MAODV [6]. The network topology is randomly generated using the method proposed in Ref. [10]. For each pair of nodes, an edge is placed connecting the two nodes with probability:

$$P_e(u,v) = \beta \exp(-\frac{d(u,v)}{\alpha L}) \quad (13)$$

Where $d(u,v)$ is geometric distance from node u to node v, L is maximum distance between two nodes. The parameters α and β are in the range (0, 1) and can be used to obtain certain desirable characteristics in the topology. Parameter α can be used to control short edge and long edge of the random graph, and parameter β can be used to control the value of average degree of the random graph.

This paper constructs a simulation of a QoS-based multicast system based on one to many multicast application. Our simulation modeled a network of mobile nodes placed randomly within 1000m × 1000 m area. Radio propagation range for each node was 250 meters and channel capacity of 10 Mbps is chosen. For a multicast system of node degree of 4 with 20 to 100 nodes randomly distributed. In order to simulate the real situations, multicast group size are always made less than 30% of the total nodes, in that multicast applications running in a wide area network usually involve only a small number of nodes in the network, such as video conference systems, distance learning, co-operative editing systems, etc. Constant Bit Rate (CBR) traffic was used. At each simulation point, the simulation runs 30 times. Each time the nodes in the group *G* are randomly picked out from the network graph.

5.3. Simulation Results

Fig. 2 depicts the comparison of data transmission rate under nodes' changing movement speed for these three protocols: the faster the node's movement speed, the smaller the protocol's data transmission rate, which due to the fact that when the movement speed increase for the nodes, the network's topology structure changes faster. From the Fig. 2 we can see that when the node's movement speed increases, QMRGA data transmission rate is higher than that of MAODV. When the node movement speed is controlled with a range, the network's topology structure will not change fast, the link's break rate of the multicast tree is low, make QMRGA QoS constraints assured within most of user's movement speed range, so QMRGA has a good performance within the network node's constrained movement speed scope.

Fig. 3 and Fig. 4 depict a comparison among success rate to find the path through MAODV. With the relaxation of

delay and bandwidth constraints, the success rate becomes larger for MAODV, Meanwhile QMRGA's success rate is still higher than that of MAODV, which means QMRGA is more suitable for the routing choosing under timely data transmission application and dynamic network structure.

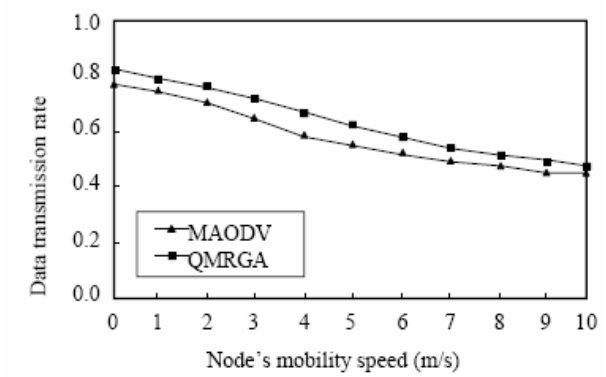


Fig. 2. Comparison of data transmission rate

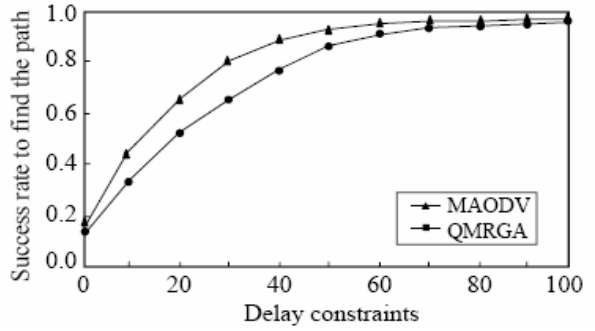


Fig. 3. Comparison of success rate to find the path

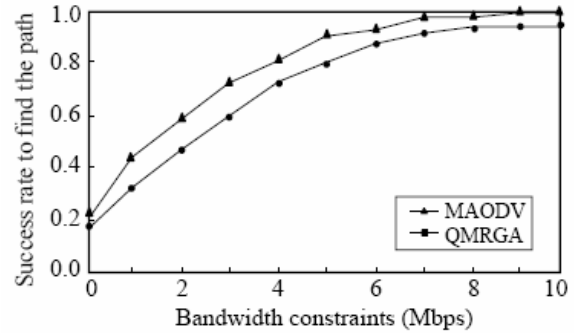


Fig. 4. Comparison of success rate to find the path

6. CONCLUSIONS AND FUTURE WORK

Multicast applications involving real-time audio and/or video transmissions require strict QoS constraints (end-to-end delay, entropy and bandwidth availability) to be met by the network. To guarantee real-time delivery of multimedia packets, a multicast channel needs to be established in advance by using a path selection policy that takes into account the QoS constraints. Among numerous advances in high-performance networking technology, the multicast routing with QoS constraints has continued to be a very important research area.

This paper discusses the multicast routing problem with multiple QoS constraints, which may deal with the delay, bandwidth and cost metrics, and describes a network model for researching the Ad Hoc network QoS multicast routing problem. It presents an entropy-based model to support QoS multicast routing algorithm for MANET based on Genetic Algorithm (QMRGA). The key idea of QMRGA protocol is to construct the new metric-entropy and select the stability path with the help of entropy metric to reduce the number of route reconstruction so as to provide QoS guarantee in the ad hoc network. The simulated results demonstrate that the proposed approach and parameters provide a correct and effective method to evaluate the route stability in dynamic mobile networks.

In a word, the deep research of QoS constraint multicast routing will increase the technology of high performance network routing system, and it will be widely applied in video, multimedia broadcasting and distance education fields, etc.

7. ACKNOWLEDGMENT

Supported by National Natural Science Foundation of China (90304018), Natural Science Foundation of Hubei Province (2005ABA231).

8. REFERENCES

[1] B.L. Sun, L.Y. Li, "A QoS Multicast Routing Optimization Algorithms Based on Genetic Algorithm," Journal of Communications and Networks, 2006, 8(1): 1-7.

[2] B.L. Sun, L.Y. Li, et al, "An Entropy-Based Stability QoS Multicast Routing Protocol in Ad Hoc Network," The First International Conference on Grid and Pervasive Computing (GPC 2006), Taichung, Taiwan, 2006, 3(5): 217-226.

[3] B.L. Sun, Q. Yang, et al, "Fuzzy QoS Controllers in Diff-Serv Scheduler using Genetic Algorithms," 2005 International Conference on Computational Intelligence and Security (CIS2005), Xian, China, 2005, 15(19): 101-106.

[4] J. J. Wu, R.H. Hwang, "Multicast routing with multiple constraints," Information Sciences, 2000, 12(4): 29-57.

[5] B. Wang, J.C. Hou, "Multicast routing and its QoS extension: Problems, algorithms, and protocols," IEEE Network, 2000, 14(1): 22-36.

[6] E.M. Royer, C.E. Perkins, "Multicast Operation of the Ad Hoc On-Demand Distance Vector Routing Protocol," ACM MOBICOM, 1999, 207-218.

[7] B. An, S. Papavassiliou, "An Entropy-Based Model for Supporting and Evaluating Route Stability in Mobile Ad hoc Wireless Networks", IEEE Communications Letters, 2002, 6(8): 328-330.

[8] A. Shiozaki, "Edge extraction using entropy operator", Comp. Vis., Graphics, Image Processing, 1986, 36: 1-9, 66~72.

[9] G.L. Chen, "Genetic Algorithm with Applications," Beijing: Posts & Telecom Press, 1996, 53-70.

[10] P.V. Mieghem, "Paths in the simple Random Graph and the Waxman Graph," Probability in the Engineering and Informational Sciences (PEIS), 2001, 15(3): 535-555.

Optimization of PCA-RBF Based Soft-Sensing Model*

Zhang Hong
School of Mechanical Engineering, Southern Yangtze University
Wuxi 214122,China
Email: jndxzh@sohu.com
Xu Wenbo
School of information Engineering, Southern Yangtze University
Wuxi 214122,China
Email: xwb@sytu.edu.cn
Liu Fei
Institute of Automation, Southern Yangtze University
Wuxi 214122, China

ABSTRACT

Principal Component Analysis (PCA) can effectively extract the characteristic information of data and eliminate co-linearity in variables. Radial Basis Function Neural-Network has the properties of fast learning and Global Convergence. This article puts forward a soft-sensing modeling method combining PCA and RBFNN, which selects the RBF centers by means of Data Clustering based on Quanta Particle Swarm Optimization. The transform function in Hidden Layer Unit is Gauss Transform. The weight from hidden to output layer is also optimized by QSPO. The obtained soft-sensing is verified in measuring the oxygen content in CCR heating furnace.

Keywords: soft-sensing , PCA, Neural-Network, QSPO, RBF

1. INTRODUCTION

Based on a certain optimization principal, using measurable information consisting of assistant variables, and by means of computing calculation, soft-sensing technology can achieve the measurement of main variables. The core of soft sensor technique is characterizing the soft-sensing model of mathematical relationship between assistant variables and main variables. Therefore, in fact factoring soft sensor technique itself is how to establish soft-sensing model, a matter of model building. Accordingly, the process of soft-sensing establishing is the process of soft sensor technique factoring.

Soft-sensing, based on Artificial Neural Network (ANN), is a kind of technology that has been put into research the most, developed fast, and applied widely in recent years. Because ANN possesses the characterization of self-learning, associational memory, and self-adaptation, nonlinear approximating, etc. soft-sensing can build a model directly by using of input and output data, without the condition of pre-known knowledge (using assistant variables as the input to ANN, while using main variables as the output to ANN. Fulfilling the soft-sensing by means of ANN'S self-learning ability). Such a model is strongly capable of linear-tuning, and it also fits the high non-linear and serious uncertain system, so it provides a highly effective method for cracking the problem of processing parameters for soft-sensing in complex system. During past few years, the methods of combining Neural Network with PCA, WAVELET, GA, or PSO, etc [2,3,4,5] have been put forward to improve the precision of Robustness and on-line prediction of soft-sensing model, which has achieved great developement in either theory or practice.

Neural Network varies a lot, and Multilayer BP and RBF are more used in soft-sensing technology development than others. BP Network and RBF Network are both accomplished by neural elements forward connection and there is no connection in the same layers, but the connection between adjoining neural elements with the signal propagating from lower level neural to higher level neural. Traditional BP Network is based on gradient method, which is simple in algorithm and definite in sense, but slow in converge and inevitable in local minimum. RBF Network (Local Apperceiving Field Network) is composed of three layers, and divides the input space naturally, which makes the input vectors expending to high dimensional hidden layer unit space, so it is more possible being linearly divided and reduce the possibility of local minimum largely.

2. QUESETION RAISED

CCR furnace is a fire heating equipment to provide heat resource for CCR plant and plays a crucial role. The furnace heats up hydro-cracked heavy naphtha gradually before it is fed into the reactor. The combustion of the furnace directly affects the reaction depth and its downstream product quality. Oxygen content is considered as a significant indicator for evaluating CCR heater performance，as a matter of fact, it gives a lot of impacts to the heater, and direct influences the efficiency of the furnace. Lower O_2 content may cause incomplete combustion, increase thermal losses, and consequently the heat efficiency is reduced. Whereas, if O_2 content is too high the excessive air will carry away substantial amount of heat into the atmosphere at flue gas temperature, which may increase flue gas heat losses and in turn discount heat efficiency. Moreover, the O_2 content offers a direct influence to flue gas resistance. It is worth to address that the higher O_2 content in the flue gas can accelerate oxidation of the internals of the heater. Excessive O_2 content in the flue gas will promote SO_2 to convert to SO_3, and worsen low temperature corrosion at dew point of the flue gas. By analyzing operation mechanism of furnace performance it is observed the physical quantity such as combustion chamber temperature, convection temperature, exhaust flue gas temperature, down stream temperature, combustion chamber pressure are all

Author: Zhang Hong (1966－), from Wuxi, Jiangsu, PhD. in automatic control theory and Associate Professor.

closely correlated with O_2 content [12]. At present, O_2 content is measured by using a special but imported O_2 analyzer, based on zirconia measuring theory. Due to severe condition and environment around sampling locations, the function of this O_2 analyzer is quite inconsistently and prone to erosion damage. However, soft measurement technique is a good methodology for the job, simply by measuring O_2 content related auxiliary variables and setting up its mathematical models, it is able to give us an optimum and on line estimated O_2 figure. Nevertheless, we have to see that because of the intricate internal relations between O_2 and auxiliary variables, and its poor non-linearity, correlativity and uncertainty, the soft measurement model built with traditional concept is difficult to truly depict the Process Characteristics. Artificial Neural Network (ANN) on the other hand is rather competent for achieving the on line measurement function, without necessary acquirement of prior knowledge of steady states and dynamic process.

3. BUILDING OF SOFT-SENSING MODEL BASED ON RBF NEURAL NETWORK

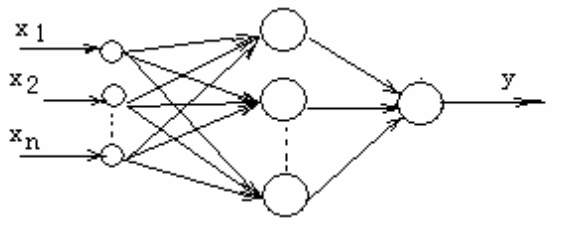
Establishment of soft measurement model based on Ra- dial Basis Functions Neural Network（RBFNN）Radial Basis Functions Neural Network（RBFNN）is able to acquire natural division on input space, which enables input quantity extending to high dimensional hidden unit, and consequently reducing the local minima greatly. If a right center of RBFNN can be located, very few hidden layers would be required to get approximation together with quick learning speed and optimum approximation features. With appropriate optimization algorithm RBF can even ensure global conver- gence. In addition, RBF modeling accuracy is also excellent [13,14,15].

3.1 STRUCTURE OF RBF NEURAL NETWORK

The structure of RBF Neural Network [6,7,8,9] is similar as that of the Multilayer Feed Forward Network. It is a three-layer feed forward network. The input layer is composed of signal source nodes. The second layer is the hidden layer, with the number of units depending on the requirement of the described question. The third layer is the output layer, which responds to the function of the input mode. The transform from the input space to the hidden space is non-linear, whereas the transform from the hidden space to the input space is linear. The transform function of the hidden unit is RBF, which is a locally distributed non-negative, non-linear function that attenuates radially symmetrically with respect to the center point. In the case of RBF network, the learning algorithm does not have the problem of local optimum and, as the parameter adjustment is linear, higher convergence speed is expected. Moreover, it is of a global approximation nature and has optimum approximation performance. The structure of RBF Neural Network is as shown in the Fig. The most commonly used transform function is Gaussian function:

$$\Phi_k(x) = \exp\left[-\frac{1}{(\sigma_k)^2}\left\|x-\underline{\mu}_k\right\|^2\right], (k = 1,2,...,h)$$

where x represents n-dimensional input vector; μ_k is the center of the k_{th} basis function and is the vector with the same dimensions as x; σ_k is the i_{th} perceptron variable (it can be selected freely), which determines the width of this basis function. H is the number of perceptrons, $\left\|x-\underline{\mu}_k\right\|$ denotes the distance between x and μ_k, Φ_k（x）reaches a exclusive maximum at the point of μ_k. As $\left\|x-\underline{\mu}_k\right\|$ is increasing, Φ_k（x）attenuate to zero rapdilly. Given input $x \in R^n$, only a small part of centers near to x are activated.



Input layer Hidden layer Output layer

Fig. 1. RBF Neural Network Strucure Diagram

The major factors that may affect Robustness and accuracy of RBF Network soft-sensing are as follows: selection of assistant variables, preprocessing of sample data, selection of proper RBF and the best approximation algorithm of RBF Network.

3.2 SELECTION OF ASSISTANT VARIABLE AND PROCESSING OF DATA

In the neural network soft sensing technique, the assistant variables as the input layer are usually selected from the measurable parameters using a method combining mechanism analysis and empirical deduction, and the optimum number of assistant variables selected is related to the degree of freedom of the process, the measurement noise and the uncertainty of model. Before the network learns, the input data should be processed. This has an essential effect on the network, affecting not only the learning speed and complexity of the network, but also the accuracy of the network. The data processing is normally conducted in three steps: collection of relevant variables (data), evaluation and extraction of valid variables and data transform data conversion and processing. Abnormal data may appear among data from time to time, which will affect the network learning. Before the network learning, the data must be checked for any abnormal data included. If any, check if those data are induced by deviation and, if yes, delete them.

Principal Components Analysis (PCA) is capable of mapping the high-dimensional data with multiple correlated variables to the mutually independent low-dimensional spaces, so as to significantly mitigate the difficulty in direct analysis of multi-dimensional complex process variables. PCA can be used to realize data simplification, data denoising, data compression, modeling, and singular value detection and variable selection. PCA can adequately dig out the information hidden in data, which is a kind of comparatively successful data mining technology.

After normalization, the sampling data is saved as a $m \times n$ data matrix X, in which each column corresponds to a variable and each line corresponds to a sampling time. Matrix X can be decomposed into n sums of outer products of vector:

$$\boldsymbol{X} = \boldsymbol{t}_1 \boldsymbol{p}_1^T + \boldsymbol{t}_2 \boldsymbol{p}_2^T + \cdots + \boldsymbol{t}_n \boldsymbol{p}_n^T$$

Where, $\boldsymbol{p}_i \in \boldsymbol{R}^n$ is load vector, and

$$\boldsymbol{p}_i^T \boldsymbol{p}_j = \boldsymbol{0} \qquad i \neq j$$

$$\boldsymbol{p}_i^T \boldsymbol{p}_j = \boldsymbol{1} \qquad i = j$$

i.e. the load vectors are orthogonal to each other and the length of each load vector is 1.

$\boldsymbol{t}_i \in \boldsymbol{R}^n$ is the principal component and the projection of data matrix X in the direction of the load vector corresponding to this scoring vector, i.e.:

$$\boldsymbol{t}_i = \boldsymbol{X}\boldsymbol{p}_i$$

If the scoring vectors are arranged by their lengths as $\|\boldsymbol{t}_1\| > \|\boldsymbol{t}_2\| > \cdots > \|\boldsymbol{t}_m\|$, the load vector $\boldsymbol{p}_1$ will represent the direction of maximum variation of datum X, $\boldsymbol{p}_2$ will be perpendicular to $\boldsymbol{p}_1$ and represent the direction of second maximum variation of datum X, and $\boldsymbol{p}_m$ will represent the direction of minimum variation of the datum.

When there exists a certain degree of linear dependence among the variables in matrix X, the variation of data is mainly reflected in the directions of the foremost few load vectors and the projection of data matrix X on the last few load vectors will be very small, which are caused mainly by the sensing noises. Thus, after principal component decomposition of matrix X, the following formula can be written:

$$\boldsymbol{X} = \boldsymbol{t}_1 \boldsymbol{p}_1^T + \boldsymbol{t}_2 \boldsymbol{p}_2^T + \cdots + \boldsymbol{t}_k \boldsymbol{p}_k^T + \boldsymbol{E}$$

where E is the deviation matrix, representing the variation of $\boldsymbol{X}$ in the directions of $\boldsymbol{p}_{k+1}$ through $\boldsymbol{p}_m$ load vectors. In many actual applications, k is often much less than m, which can remarkably compress the number of dimensions of the data. Since the deviation matrix mainly results from sensing noises, neglect of the same often has the effect of removing the sensing noises, without causing any noticeable loss of usable information in the data. Therefore, the original datum $\boldsymbol{X}$ may be approximately expressed as

$$\boldsymbol{X} \approx \boldsymbol{t}_1 \boldsymbol{p}_1^T + \boldsymbol{t}_2 \boldsymbol{p}_2^T + \cdots + \boldsymbol{t}_k \boldsymbol{p}_k^T$$

In this way, the original enormous data can be successfully compressed into a low-dimensional data matrix.

3.3 DETERMINATION OF PARAMETERS FOR RBF NEURAL NEUTRAL NETWORK

In RBF network, as the tasks fulfilled by the output layer and the hidden layer are different, their learning strategies are also different. The output layer adjusts the linear weight and uses the linear optimization strategy, so its learning speed is faster. The hidden layer adjusts the parameter of the action function and uses the non-linear optimization strategy, so its learning speed is slower. This reveals that the "time marks" of the learning process of the two layers are also different. Therefore, learning is normally carried out separately at two layers.

The common learning methods of RBF network include: random selection of RBF center (direct calculation method) and selection of RBF center by self-organized learning. In the first method, the center of hidden unit RBF is randomly selected among the sample input data and the center is fixed. When the center of RBF is determined, the output of hidden unit is known. Then the connection weight of the network can be determined through resolving of a set of linear equations. For a given question, if the distribution of sample data is representative, this method could be regarded as a simple and practical method. In the second method, the RBF center is movable and its position is determined through self-organized learning. The linear weight of the output layer is calculated through supervised learning. This shows that it is a mixed learning method. The self-organized learning part allocates the resources of the network in a sense and the purpose of learning is to locate the RBF center in an important zone of the input space. The k-mean clustering algorithm may be used to select the RBF center c_i. The output weight can be calculated with the standard linear least squares regression method, which is a supervised learning process. This method is the more commonly used RBF training algorithm at present time.

The data cluster method based on Quantum-Behaved Particle Swarm Optimization (QPSO) proposed in this thesis for selection of RBF center is a new cluster method with low convergence error of cluster, greater distance between clusters formed and smaller distance inside a cluster. The output weight can also be optimized with QPSO.

Quantum-behaved Particle Swarm Optimization is a new Particle Swarm Optimization (PSO) model proposed by Sun Jun in 2004 from the perspective of quantum mechanics. He believed that particles had the behavior of quanta and proposed the Quantum-behaved Particle Swarm Optimization on the basis of such model [10, 11].

In 1995, the American social psychologist James Kennedy and electric engineer Russell Eberhart jointly proposed the Particle Swarm Optimization (PSO). The basic idea was inspired by the result of their research on group behavior of birds in their early years and the life-form population model of biologist Frank Heppner was also utilized [6]. In a d-dimension space, the position of particle i is expressed as vector $Xi=(x_1, x_2, \cdots, x_d)$. $P_i = (p_{i1}, p_{i2}, ..., p_{id})$ is the best position experienced by particle i, namely the position with the best fitness value that particle i has experienced, called the personal best position (pbest). $P_g = (p_{g1}, p_{g2}, ..., p_{gd})$ is the best position that all particles in the group have experienced (gbest). Nevertheless, none of the developed PSO algorithm can assure global convergence of the algorithm, because the evolution equation of PSO causes all particles to search in a limited sample space.

The QPSO method proposed by Sun Jun not only reduces the number of parameters, but also has better searching capability than any developed PSO algorithms. Optimization of the neural network parameters brings about increased convergence speed and accuracy, less undetermined parameters and higher operability, rendering the method superior to the Speed Gradient Descent Method, Genetic Algorithm and other optimization algorithms. It can assure global convergence of the algorithm and is superior to all developed PSO algorithms in the searching capability. In QPSO, the particle swarm moves according to the following three formulas:

$$mbest = \frac{1}{M}\sum_{i=1}^{M} P_i = \left(\frac{1}{M}\sum_{i=1}^{M} P_{i1}, ..., \frac{1}{M} P_{id} \right) \qquad (1)$$

$$x_{id} = p_{id} \pm \alpha * \left| mbest_d - x_{id} \right| * \ln(\frac{1}{u}) \qquad (3)$$

$$p_{id} = \varphi * P_{id} + (1-\varphi) * P_{gd} \qquad \varphi = rand \qquad (2)$$

$$u = rand$$

Where $mbest$ is the middle position of $pbest$ and p_{id} is the random point between P_{id} and P_{gd}. α is the contraction and expansion coefficient of QPSO, an important parameter for QPSO convergence. It is normally as

α =(1.0-0.5)*(MAXITER-T)/MAXITER+0.5;

with a quite good effect to be achieved, where MAXITER is the maximum times of iteration.

The procedure of selecting RBF center with the data cluster method based on QPSO algorithm is as follows:

Input: number of clusters, data vector, number of particles

Output: the optimum fitness value J_e of particle after each iteration.

① Initialization.

For T=1: MAXITER

② Conduct clustering based on Euclidean distance.

③ Calculate $mbest$ with formula (1).

④ Update local optimum pbest.

⑤ Update global optimum gbest.

⑥ Calculate the random point with formula (2).

⑦ Update the center vector of particle with formula (3).

End

Repeat the calculation in ② through ⑦ steps, till the times of iteration is met.

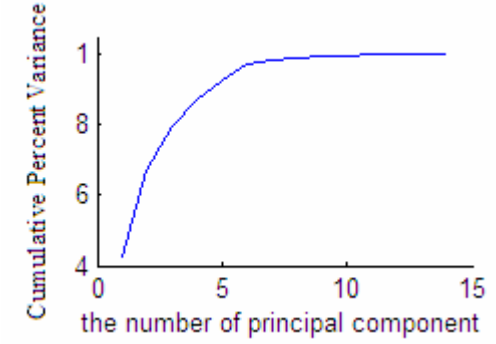
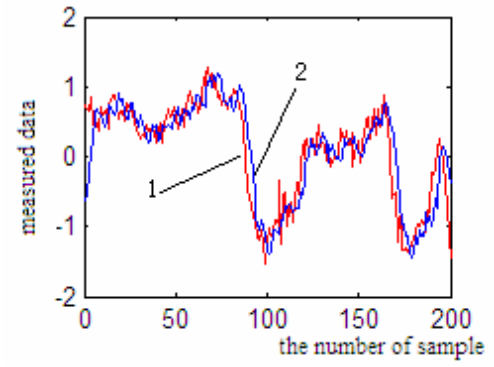
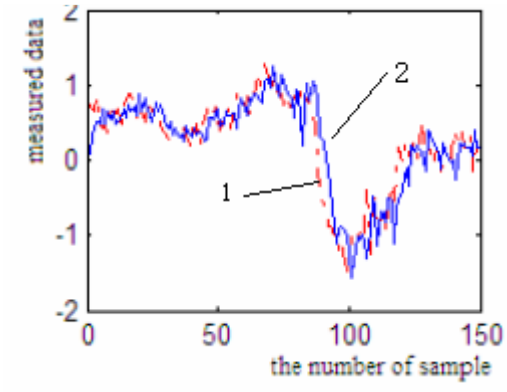


Fig. 2. The degree of principal components



1— Actuallu measured value,
2— Soft-sensing value

Fig 3 Training curve of oxygen content soft-sensing model



1— Actually measured value,
2— Soft-sensing value

Fig. 4. Training simulation curve for oxygen content soft-sensing

The output weight W can also be trained with QPSO algorithm. First of all, encode weight W into an individual expressed by a real number code string. Assuming the network incorporates M weights that need to be optimized, each individual will be expressed by a m-dimensional vector composed of M parameters. Based on the scale of the particle swarm and the above individual structure, a certain number of (N) individuals (particles) will be generated randomly to form a population and, at the same time, pbest and gbest will be initialized.

Table 1 Radial Basis Function Center

1	-8.2531	1.7283	-9.0262	-4.4429	-9.7021
.2	-6.9643	3.1417	3.0502	-2.7278	8.8139
3	-7.3922	1.0789	3.9846	-0.3503	3.6476
4	-0.4422	-1.3844	-9.6121	-0.6945	0.1883
5	4.5309	-9.7974	-0.0631	-1.3791	8.171
6	-6.8373	4.2467	-3.0584	1.0921	1.635
7	7.1295	1.5385	5.9679	-0.496	6.9734
8	3.6972	2.844	0.4478	6.2852	-2.4228
9	-3.2583	-7.3182	-1.9685	0.2822	4.6181

10	-8.7413	3.3148	-5.5805	2.5289	9.0964

Table 2. Width of Radial Basis Function

1	2	3	4	5
9.7824	0.9024	-10.888	-11.521	-11.803
6	7	8	9	10
17.45	14.694	-14.939	-15.065	12.481

Table 3. Weights between Hidden Layer and Output Layer

1	2	3	4	5
5.8817	26.902	3.0051	22.581	-31.876
6	7	8	9	10
-31.57	11.75	13.584	-7.3354	17.886

Then map each individual in the particle swarm into a network parameter to form a RBF neural network. Input the training sample into the neural network corresponding to each individual and conduct training. And calculate the mean square error produced by each network on the training set:

$$E(X_p) = \frac{1}{2n}\sum_{p=1}^{n}\sum_{k=0}^{c}[y_{k,p}(X_p) - d_{k,p}]^2$$

where $y_{k,p}$ is the actual network output of training sample p at output k, and $d_{k,p}$ is the corresponding given output. It is taken as the objective function. Evaluate all individuals in the particle swarm and find the best individual (i.e. the individual with the least mean square error in the particle swarm), which will be used to determine if it is necessary to update the pbest and gbest of the particles. Then update the position vector of each individual as per the QPSO model so as to produce new individual particles. Continue to map the newly produced particles into the parameters of network and input training sample to train the network. Repeat the steps again and again till the termination condition for the algorithm is reached.

4. SOFTING RESULT

After data preprocessing of the pre-selected 14 secondary variables correlated with oxygen content using the PCA method, vector matrixes P orthogonal to each other, with characteristic value ordering, can be obtained, as shown in Fig. 2. The cumulative percentage of interpretation of the first 5 independent components is 90%. Taking the first 5 independent components for data mapping, as input to the RBF neural network, using 10 neurons of the hid- den layer of RBF neural network and 200 groups of training samples, with the transform function of the hidden unit being the above mentioned Gauss function, conduct clustering of the RBF center with QSPO method. The center of each hidden unit is as shown in the table below.

By optimizing the mean square error produced by the training set with the weights between the hidden layer and the output layer,the train ingcurve of oxygen Content soft-sensingmodel as shown in Fig. 3 can be obtained. Simulate the trained model with another 150 groups of samples and the training simulation curve as shown in Fig. 4 can be obtained. The speed of convergence at optimization will differ with different number of particles.

When 50 or more particles are taken, the convergence effect will be remarkable. Fig. 5 is the convergence chart, where the interative convergence curve for optimization with 100 particles is shown.

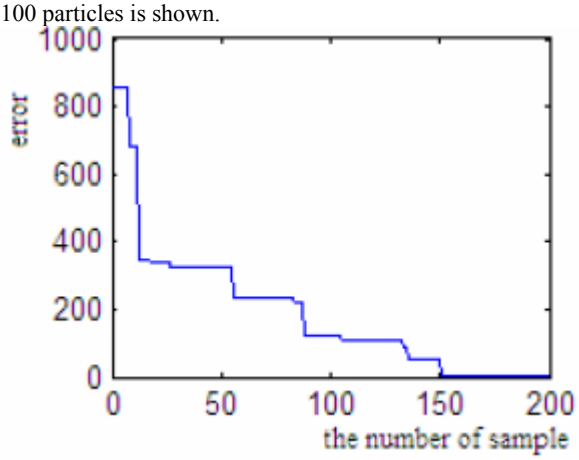


Fig. 5 Training convergence curve

5.CONCLUDING REMARKS

This thesis introduces the advanced QPSO algorithm into PCA_RBF soft-sensing model to optimize the basis function center value and the weights between hidden layer and output layer, with which better sensing accuracy and higher convergence speed are achieved during the oxygen content sensing of a continuous reforming heating furnace. However, actual engineering application requires further researching and improving the main elements that compose the soft sensor, such as the quantity and quality of network learning and training sample, the learning algorithm and the topology structure and type of network, and further improving the robustness and accuracy of the soft-sensing model.

6. REFERENCES

[1] H.Q. Li, Soft-Sensing technique principle and application, Beijing: Chemical Industry Press, 2000.
[2] B. Widrow, “A daptive noise canceling: principles and applications”, Proc. IEEE, 1975, 63(12): 1 692-1 716.
[3] S.L. Lai, F. Ming, “A Novel Local PCA- Based Method for Detecting Activation Signals in FMRI. Magnetic Resonance Imaging”, 1999, 17(6) 827–836.
[4] Y.G. CEN, Y.Q. Qing, D.B. Sun, et al, “Application of PSO Algorithm in Wavelet Neural Network”, Joural of System Simulation, 2004, 16(12): 2783-2785.
[5] J.R. Koza, J.P. Rice, “Genetic generation of both the weights and architure for a neural network”, International Joint Conference on Neuralworks, 1991, pp.397-404.
[6] J. Moody, C. Darken, Learning with Localized Rece- ptive Fields, In Proceedings of the1988 Connectionist Models Summer School, Pittsburgh, PA, 1988.
[7] J. Moody, C. Darken, “Fast Learning in Networks of Locally-Tuned Processing Units”, Neura l Computation, 1989, 1(2):281-294.
[8] T. Poggio, F. Girosi, Networks for Approximation and

Learning, Proceedings of IEEE, 1990, 78(9): 1481-1497.
[9] B. Ben, G.C. Christophe, "GA-RBF: A Self-Optimising RBF Network", http://www.cs.bris.ac.uk/Publications/Papers/1000219.pdf.
[10]J. Sun, "Particle Swarm Optimization with Particles Having Quantum Behavior", IEEE proc. Of Congress on Evolutionary computation, 2004.
[11] J. Sun ,W.B. Xu ,et al, "Aglobal Seach Strategy of Quantum-Behaved Particle Swarm Optimization", Proceedings of the 204 IEEE Conference on Cybernetics and Intelligent Systems,2004.
[12] J.L. Qian, Tubular furnace, Beijing:Sinopec Press, 2003.
[13] T.Y. Chai, X.B. Wang, RBF Neural-Network application in accelerated cooling control ststem, Chinese Journal of Automation 2000, 26(2): 219225.
[14] L.D. Pan, X.F. Huang, et al, "RBF Neural Networks and Its Application in Establishing Nonlinear System Self-Tuning Model", IFAC, ADCHEM'97, Banf, Canada, 1997.
[15] Y.C. Huang, S.H. Sun, "A Novel learning algorithm for RBF neural networks to identify non-linear system", Journal of Harbin Institute of Technology, 1995, 27(4): 74-81.

Research on the Implementation Techniques of BEA Weblogic Server Clusters Based on J2EE

Hongmei Nie, C. J. Duanmu, Jianmin Zhao
Dept. Of Information Science and Engineering
Zhejiang Normal University, Zip Code: 321004
Jinhua, Zhejiang Province, P. R. China
Email: nhm@zjnu.cn, duanmu@zjnu.cn , zjm@zjnu.cn

ABSTRACT

A WebLogic server which is based on the principles of J2EE is a currently advanced platform of electronic business. This paper discusses three techniques to implement a WebLogic cluster. They are the JSP/Servlet cluster, the object cluster, and the cluster based on java message services (JMS). It also illustrates how to realize the functionalities of workload equalization, and breakdown switching in a JSP/Servlet cluster.

Keywords: J2EE, WebLogic Server, clusters, EJB, JMS

1. INTRODUCTION

In the implementations of electronic business, thousands of hundreds of customs access the application software through the internet. In order to let all the clients successfully access the application software at any time without declination, the J2EE server must have high availability, scalability, and error-resilience. A WebLogic server is just such ideal J2EE server. Several WebLogic servers can simultaneously serve the core application software for the same task and thus form a cluster. A WebLogic cluster is a seamless combination of many WebLogic servers, which act as a single and powerful server. This paper will discuss the techniques for implementing a WebLogic server cluster by using the author's practical experience.

2. TECHNIQUES FOR A JSP/SERVLET CLUSTER

In this technique, the WebLogic server for application software is implemented by replicating the status of the transaction sessions of the customs accessing the JSP/Servlet cluster.

2.1 Saving of the Information of a Transaction Session

There are four choices for processing the information of a transaction session. They are (1) no protects of the information of transaction sessions (2) protects of the information of transaction sessions by replicating memory contents (3) protects of the information of transaction sessions by using files (4) protects of the information of transaction sessions by using databases.

(1) no protects of the information of transaction sessions

In this mode, whenever there is a breakdown of a WebLogic server, all the information of the transaction sessions will be lost. The advantage of this mode is that this mode has the least influence on the performance of a server.

(2) Protects of the information of transaction sessions by replicating memory contents

In general, a transaction session is directed to the same server in a cluster. This server is called the master server. In this mode, a cluster automatically chooses another server as the slave server to replicate the information of the transaction sessions of the clients on the master server. Whenever an object of a Http session is modified, this information is replicated to the slave server by using IP sockets. That is, whenever there is a call of the function of SetAttribute() or removeAttribute(), the replication of the Http session on the slave server will be automatically and simultaneously updated by the master server of the WebLogic cluster. In this way, whenever there is a breakdown of a WebLogic server, all the following requests from the clients to this server will be automatically switched to the slave server and there is no loss of information of the transaction sessions. When adapting this mode, it must be considered whether the cost of replicating memory contents is acceptable for the cluster or not.

(3) Protects of the information of transaction sessions by using files

In this mode, all of the information of a transaction session is saved in files. Whenever there is a modification request from a transaction session of a client, there is a writing operation of a file. This mode has some influence on the performance of a server.

(4) Protects of the information of transaction sessions by using databases

In this mode, all of the information of a transaction session is saved in a database. As the mode of (3), whenever there is a modification request of a transaction session from a client, there is a transaction session of a database. As the mode of (3), this mode has some influence on the performance of a server.

2.2 Implementations of Workload Eqalization

The process of workload equalization is to equally allocate the workload of the cluster to the servers according to some principles. Through workload equalization, every server can process the transaction sessions of the clients and this process can increase the availability and scalability of the cluster. The workload equalization is transparent to the clients. There are several modes of workload equalization. They are (1) workload equalization using a WebLogic proxy, which can be integrated with other servers (2) workload equalization using hardware (3) workload equalization using firewall (4) workload equalization using software (5) workload equalization using the DNS round robin algorithm.

2.2.1 Worklaod equalization using WebLogic proxy

WebLogic itself has the functionality of workload equalization. It can not only act as a proxy for workload equalization, but also serve as an insertion component

integrated with other servers. The JSP/Servlet workload equalization in WebLogic must satisfy the following two requirements. (1) It can interpret the strings in the cookies sent from the clients to the servers or from the servers to the clients. (2) It can encrypt and decrypt the SSL transaction sessions. Only when it satisfies the above two requirements can it recognizes and supports the master servers for the transaction sessions between the clients.

As a mechanism for workload equalization, WebLogic can set the parameters for routing, which can not be realized by using other kinds of hardware. A WebLogic server provides the means to control the workload equalization in a micro-granularity way. Any cluster can designate a call router, which is a target for objects in the cluster.

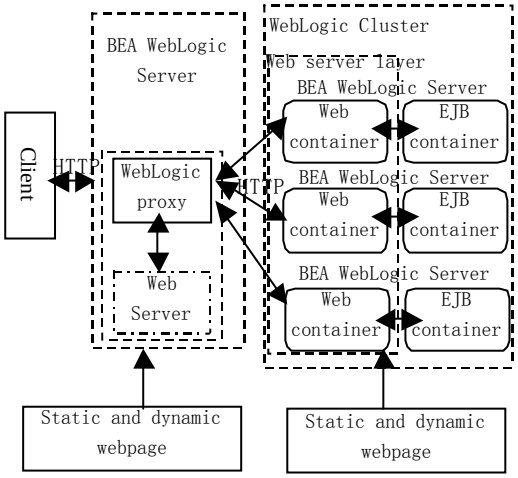


Fig. 1 Settings of a cluster using BEA WebLogic as workload equalizer

(1) Use a BEA WebLogic server as a proxy for workload equalization

As a proxy, a WebLogic server an automatically redirect most of the requests from the clients to the cluster, which can process these requests. Fig. 1 demonstrates the settings of a cluster, where a BEA WebLogic acts as a workload equalizer. The HttpClusterServlet in a WebLogic server allocates the requests of the clients to different servers.

(2) Integration with the web servers from a third party

A BEA WebLogic server can be integrated with the web servers from a third party. It supports the web servers currently in use, such as iPlanet (Netscape) enterprise server, Microsoft IIS, and Apache.

This mechanism enables the clients to use the web servers from a third party and thus offers the smooth transitions to the environment based on WebLogic servers. This mechanism also provides a close integration of a WebLogic server with servers from a third party and thus supports workload equalization and breakdown switching inherent in a WebLogic server. Fig. 2 illustrates the integration of a BEA WebLogic server with a WebLogic cluster, where the BEA WebLogic server acts as a proxy component.

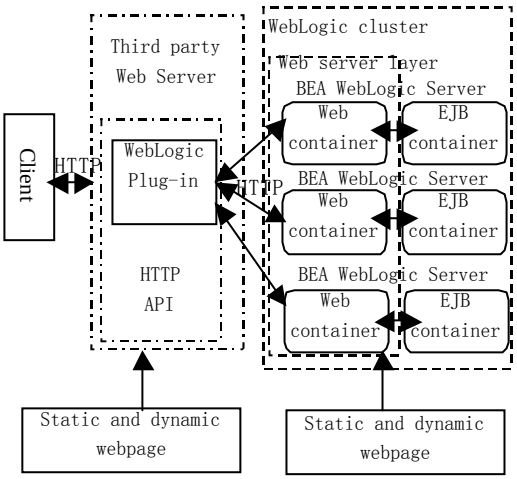


Fig. 2 Connection with WebLogic cluster using BEA WebLogicas a proxy of Plug-in

2.2.2 Workload equalization by using hardware

This mechanism of workload equalization by using hardware is suitable for a very larger web. It performs better than the workload equalization by using the software of a proxy. This mechanism can track the availability of every server and thus avoid sending the requests to abnormal servers. In this process of workload equalization, the information of the workload can be utilized by the hardware. The following are some mainly advantages of the workload equalization using hardware.

(1) More broad choices of the algorithms for the workload equalization. Various algorithms for workload equalization can be selected in this mechanism.

(2) Skipping of the breakdown server. In this mechanism, the workload equalizer can skip the breakdown servers in the cluster.

(3) Acceleration of the transaction sessions in the Secure Socket Layer (SSL). The workload equalizer can usually switch the processing of a SSL transaction session from a server of application software to some specific SSL accelerators. In this way, this mechanism can speedup the transaction sessions of SSL.

(4) Less hops in the network. A hardware workload equalizer is placed directly before a cluster of BEA WebLogic servers and thus there is only one hop in the network. In this way, this system performs better than the system, where the workload equalizer is placed before the web proxy of a third party and thus has two hops.

However, the price of the hardware workload equalizer is relatively more expensive than that of BEA Weblogic. Moreover, the hardware workload equalizer does not have the particular intelligence of a BEA WebLogic, which can apply its intelligence to offer better performance of breakdown switching and less network chaos. In addition to this, a hardware workload equalizer can not use Http cookies to carry some information of the location of the master servers and slave servers that is particular to a WebLogic and is related to the objects in the cluster. Fig.3 illustrates the mechanism where the hardware workload equalizer is connected with a BEA WebLogic server.

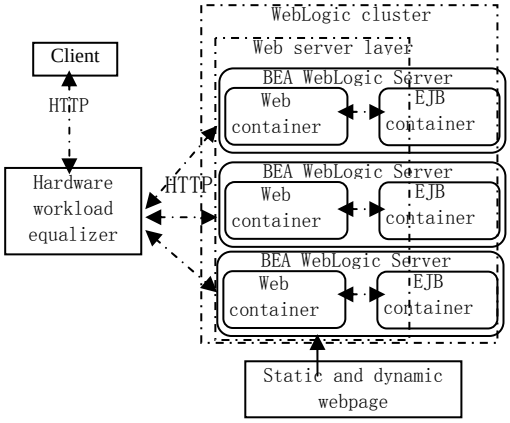


Fig. 3 Connection of hardware workload equalizer with BEA WebLogic servers

2.2.3 Workload equalization by using firewall

Some companies implements the workload equalization by using a firewall, such as the software provided by the Check Point Company. In general, this mechanism is only suitable for some very small web settings.

2.2.4 Workload equalization using software

This mechanism use only software to offer the same functionalities as those of the workload equalization by using hardware. Currently, the company, which has a leading role in this technique, is the Resonate. This company offers the product of Local Dispatch.

2.2.5 Workload equalization by using the algorithm of DNS round robin

This is the oldest mechanism in the workload equalization. A DNS server has a list of IP addresses of the web servers. Each IP address corresponds to a name of a DNS domain. Whenever a client accesses the cluster, the DNS server provides an IP address to the client in a round robin way and thus the workload equalization is realized. Since there are a lot of restrictions for the algorithm of round robin, this mechanism is generally not utilized in an important web site of electronic business. The main problem of this mechanism is that there is only one algorithm of round robin for the workload equalization. This leads to the difficult management of the DNS. Furthermore, this mechanism can not achieve breakdown switches.

2.3 Breakdown Switches

Whenever a cluster is required to offer continuous service with no stop and high availability, which guarantees the availability of its application software at any time, it must have the functionality of breakdown switching. There are two modes for breakdown switching.

(1) JSP/Servlet breakdown switching. In this mode, the breakdown switching is carried out by an inserted proxy component or a hardware workload equalizer and this process is transparent to the clients.

(2) Breakdown switching based on objects. In this mode, the breakdown switching is handled by the stub which has the functionality of replication.

The information of the JNDI update and the synchronization of a whole cluster is transmitted by the broadcasting methods between the WebLogic servers. The broadcasting methods are not robust for message transmission. However, they have a high efficiency for processing the messages.

The information of the transaction sessions of HTTP and EJB can be replicated by the utilization of IP sockets in WebLogic servers. The communications between IP sockets are robust.

The replication of java naming and directory interface (JNDI) is very important in the network. It maintains and updates the lists of the objects and services which have been accessed by an instance of a WebLogic server. A server in a cluster can utilize the JNDI tress of the whole cluster. The JNDI tree of the whole cluster is like a JNDI tree in a single server, which contains the information of all the available services. However, the JNDI tree of the whole cluster also saves the virtual position software, which points to all the available objects in the servers of the cluster, in addition to the names of local servers. Each BEA WebLogic server creates and updates the local copy of the JNDI tree of the whole cluster.

Fig. 4 illustrates the mechanism which the client uses a BEA WebLogic proxy or hardware workload equalizer to access the WebLogic servers of the cluster.

(1) Break down switching which use the WebLogic proxy as an insertion component.

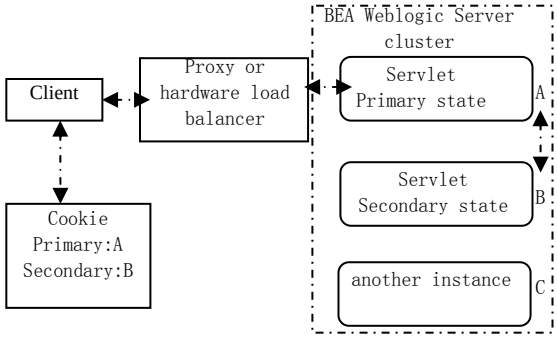


Fig. 4 Client access the cluster of WebLogic servers

During the transaction sessions of a client, if there is a breakdown of server A, the following connection requests from the client to the server A will fail. Under this situation, the proxy will determine (from cookie or URL) the slave server (server B in this example), which will be accessed by the objects of JSP or Servlet. After the request is rerouted to the slave server B by the proxy, server B will designate itself as a master server and create a slave server on any of the available server of the cluster, as shown in Fig. 5.

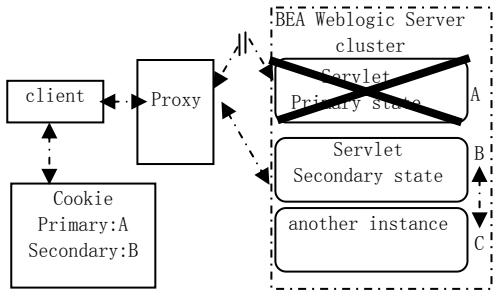


Fig. 5 A slaver server which is rerouted to the cluster

(2) Breakdown switching which uses the hardware workload equalizer.

In order to respond to the connection fail, the hardware workload equalizer will send the request to an available BEA WebLogic server of the cluster according to its workload allocation principle. In the above example, let us assume that after the breakdown of server A, the workload equalizer will reroute the requests from the client to serve C. After the client connects to the server C, server C will connect to the server B by using the information of the cookies of the client or the request of the HTTP, and then server C will obtain and copy the status of the transaction session. Under this situation, server C will become the master server of the transaction session and the server B will replicate the status of the transaction session as a slave server. In the cookies of the client, the information of the master and slave servers will be updated by the utilization of the rewriting function of URL, as shown in Fig. 6.

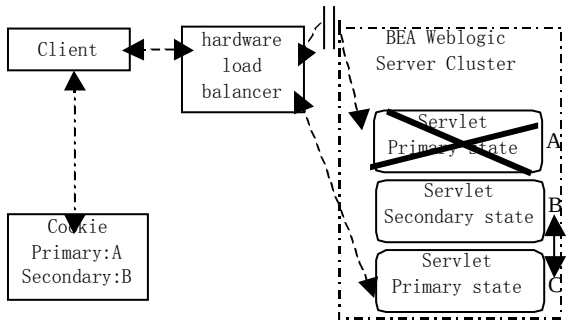


Fig. 6 Reroute to server C

3. CLUSTER OBJECTS

A cluster object is an object which has been accessed by the client through the remote function call (RFC). The common cluster object is EJB. The cluster objects on a server of a cluster are replicated to all the servers of the cluster.

When a client access a cluster object, the stub which supports replication will be automatically sent to the client. This stub contains the information of the algorithms of workload equalization, such as the round-robin algorithm, the weight-base algorithm, or random algorithm. When a client sends the request for an available server of a cluster, it will choose one of the stubs. Whenever there is a breakdown, the stub will intercept the exception and obtain another stub according to the algorithm of the workload equalization. In this way, the breakdown switching is carried out automatically.

EJB is the most common cluster object. A WebLogic cluster handles various kinds of EJB according to different methods.

(1) Stateless EJB

For an idem-potent stateless EJB, the caller function has the same return value, irrespective of the number of times which this function is called. For an idem-potent stateless EJB, its method can automatically make a switching. If all the methods in a stateless EJB are idem-potent, the stateless-methods-are-idem-potent is set as true in weblogic-ejb-jar.xml. Otherwise it is set as false. For an idem-potent staeless EJB, the stub can automatically make breakdown switching.

However, for a non-idem-potent stateless EJB, it cannot always automatically make breakdown switching, since the stub cannot know whether or not the method called by the custom under the breakdown situation will leads to exceptions due to switching.

(2) State EJB

As the method to replicate the status of the transaction session of Http, when a client creates the position software for EJB objects, the initial instance chosen by the workload equalizer will automatically select the slave server which already replicates the status of EJB.

The client receives the virtual position software which lists the positions of the master and slave servers. Under the default situation, the client only communicates with the master server. If there is a breakdown, the client will send the request to the slave server. Then, the slave server will become the main serve and it will automatically choose another server in the cluster as its slave server.

(3) Bean (including CMP and BMP)

When an EJBHome finds or creates a bean, it returns to the position software of the server. Before the breakdown of this server, all of the requests related to the position software are sent to the same server. If there is a breakdown, the BEA WebLogic server will be switched to another server. Under this situation, there is no replication. The BEA WebLogic server uses a database to save the switching information. Since the EJB is positioned inside the cluster, it only needs another available server to reload its status from the database. Then, the position software updates the position of the EJB server by using the position of newly created server.

4. JAVA MESSAGE SERVICE SERVER CLUSTER

A BEA WebLogic Server (JMS) supports a cluster of many JMS servers by using the target addresses of JMS distributed all over the cluster and their connecting factories. A target address is the instance to receive the message. A client uses the connecting factory of JMS to make a connection with a target.

The workload of application software is allocated to many JMS servers by using the connecting factories. In this way, it equalizes the workload among these servers, reduces the workload on a single JMS server, and centralizes the transaction sessions by designating a specific server to record the connection paths.

In this mode, the target addresses which are across many JMS servers are labeled as virtual targets. In this way, it reduces the workload of the physical addresses on a designated server and increases significantly the robustness and capacity of message transmission.

5. CONCLUSION

This paper discusses the cluster of BEA WebLogic server. It describes how the BEA WebLogic architecture achieves high availabbnility and scalability, reqired by enterprise clients, by using cluster techniques.

Solving the problem of a single breakdown in the cluster is one of the objectives of a BEA WebLogic cluster. This is achieved by using redundancies. Another objective is to have a high scalable architecture for electronic business of enterprises, which can make the ability to increase the capacity limitlessly under the current changing business environment.

The technique of the cluster of a BEA WebLogic is the most advanced technique. It offers the most scalability for the clients and the robustness which is mostly needed by the application software.

A BEA WebLogic cluster having high scalability and other high performances becomes the most robust and the highest scalable Java servers in the market.

6. REFERENCES

[1] BEA systems, Achiving scalablity and high availability for E-Bussiness, 2004.

[2] R. Michael,C. Wollen, et al, Application of J2EE and BEA WebLogic server, 2002.

[3] A. Ali, S. Keyur, Management of BEA WebLogic Server, 2003.