



COMPUTER MODELLING  
AND  
NEW TECHNOLOGIES

**2014**  
**VOLUME 18 NO 6**

ISSN 1407-5806 ISSN 1407-5814 on-line

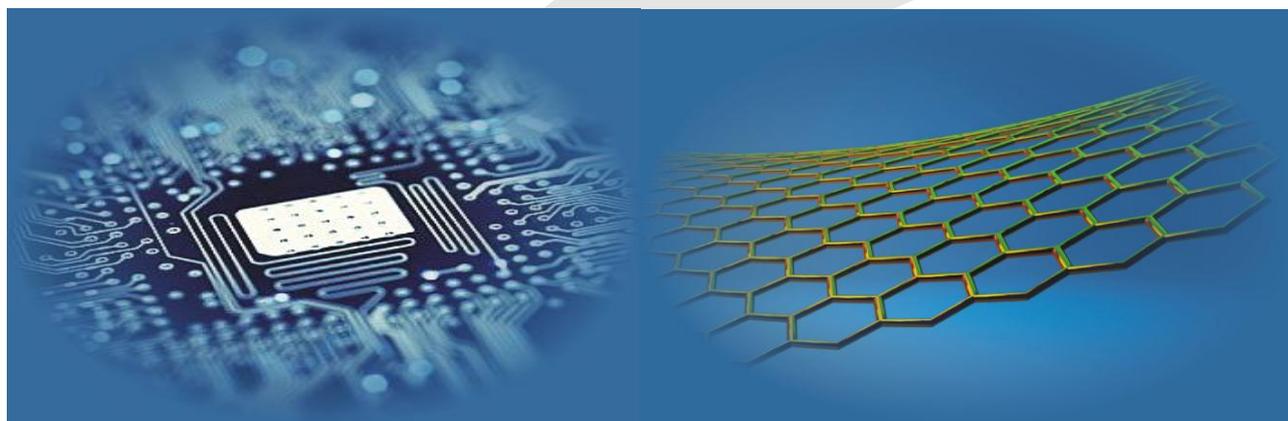
Transport and Telecommunication Institute  
and  
Latvian Transport Development and Education Association

---

# Computer Modelling and New Technologies

**2014 Volume 18 No 6**

ISSN 1407-5806, ISSN 1407-5814 (*On-line: [www.tsi.lv](http://www.tsi.lv)*)



Riga – 2014

## EDITORIAL BOARD

Prof. Igor Kabashkin	<b>Chairman of the Board</b> , <i>Transport &amp; Telecommunication Institute, Latvia</i>
Prof. Yuri Shunin	<b>Editor-in-Chief</b> , <i>Information Systems Management Institute, Latvia</i>
Prof. Adolfas Baublys	<i>Vilnius Gediminas Technical University, Lithuania</i>
Dr. Brent Bowen	<i>Embry-Riddle Aeronautical University, United States of America</i>
Prof. Olgierd Dumbrajs	<i>University of Latvia, Solid State Physics Institute, Latvia</i>
Prof. Sergey Maksimenko	<i>Institute for Nuclear Problem, Belarus State University, Belarus</i>
Prof. Vladimir Litovchenko	<i>V. Lashkaryov Institute of Semiconductor Physics of National Academy of Science of Ukraine, Ukraine</i>
Prof. Pavel D'yachkov	<i>Kurnakov Institute for General and Inorganic Chemistry, Russian Academy of Sciences, Russian Federation</i>
Prof. Stefano Bellucci	<i>Frascati National Laboratories – National Institute of Nuclear Physics, Italy</i>
Prof. Arnold Kiv	<i>Ben-Gurion University of the Negev, Israel</i>
Prof. Alytis Gruodis	<i>Vilnius University, Lithuania</i>
Prof. Michael Schenk	<i>Fraunhofer Institute for Factory Operation and Automation IFF, Germany</i>
Prof. Dietmar Fink	<i>University of Mexico, United Mexican States</i>
Prof. Ravil Muhamedyev	<i>International IT University, Kazakhstan</i>
Prof. Kurt Schwartz	<i>Gesellschaft für Schwerionenforschung mbH, Darmstadt, Germany</i>
Prof. Eva Rysiakiewicz-Pasek	<i>Institute of Physics, Wroclaw University of Technology, Poland</i>
<b>Contributing Editor</b>	Prof. Victor Gopeyenko, <i>Information Systems Management Institute, Latvia</i>
<b>Literary Editor</b>	Prof. Tamara Lobanova-Shunina, <i>Riga Technical University, Latvia</i>
<b>Technical Editor</b> , secretary of Editorial Board	MSc Comp Nataly Burluckaya, <i>Information Systems Management Institute, Latvia</i>

Journal topics:	Host Organization	Supporting Organizations
mathematical and computer modelling computer and information technologies natural and engineering sciences operation research and decision making nanoscience and nanotechnologies innovative education	Transport and Telecommunication Institute	Latvian Transport Development and Education Association  Latvian Academy of Sciences  Latvian Operations Research Society
Articles should be submitted in <b>English</b> . All articles are reviewed.		

<b>EDITORIAL CORRESPONDENCE</b>	<b>COMPUTER MODELLING AND NEW TECHNOLOGIES, 2014, Vol. 18, No.6</b> ISSN 1407-5806, ISSN 1407-5814 (on-line: <a href="http://www.tsi.lv">www.tsi.lv</a> )
<b>Transport and Telecommunication Institute</b>	<b>Scientific and research journal</b> <b>The journal is being published since 1996</b>
1 Lomonosova, <b>Bld 4</b> , LV-1019, Riga, <b>Latvia</b> <b>Phone: (+371) 67100594</b> Fax: (+371) 67100535 E-mail: <a href="mailto:cm&amp;nt@tsi.lv">cm&amp;nt@tsi.lv</a> <a href="http://www.tsi.lv">www.tsi.lv</a>	The papers published in Journal 'Computer Modelling and New Technologies' are included in: <b>INSPEC (since 2010)</b> , <a href="http://www.theiet.org/resources/inspec/">www.theiet.org/resources/inspec/</a> <b>VINITI (since 2011)</b> , <a href="http://www2.viniti.ru/">http://www2.viniti.ru/</a> <b>CAS Database</b> <a href="http://www.cas.org/">http://www.cas.org/</a> <b>El Compindex</b>



# Content

<b>Editors' Remarks</b>	5
-------------------------	---

## Mathematical and Computer Modelling

Jingfei Jiang, Rongdong Hu, Luján Mikel, Yong Dou	Accuracy evaluation of deep belief networks with fixed-point arithmetic	7
Huiyong Li, Yixiang Chen	On denotational semantics of the complex event query language STeCEQL	15
Cuirong Zhao, Honghai Wang	Analysis and research on the simulation and output of discrete event system with fuzzy parameters	24
Weiliang Zhu	Research of simulation techniques based on rough set theory	31
Hong Li, Xiaoping Ma, Zhenghua Xin	A new formal representation of granules based on features	37
Weidong Tang, Jinzhao Wu, Meiling Liu	Interleaving semantics and action refinement in event structures	44
Xiani Yang, Yaqin Lu, Cunzhi Tian	Research on growth opportunity and liquidity monitoring by mathematical optimization	52
I Keshabyan-Ivanova	A computer-assisted analysis of literary texts: a sample study	59

## Information and Computer Technologies

Chunmei Huang, Chunmao Jiang, MingCheng Qu	Design of software error detection system based on SPARC v8 and research on the key technology	65
Honghui Fan, Hongjin Zhu, Qingbang Han	Image reconstruction of concrete based on filtered backprojection method using ultrasonic time of flight data	72
Liang Jia, Nigang Sun	A line segment detection algorithm based on statistical analyses of quantified directions in digital image	79
Dajian Zhang, Minyan Lu, Nan Wu	A model-based assurance case construction approach for system control software	89
Xiao-Yuan Qu	A noise estimation approach by assembling fast edge detection and block based methods	97
Jianjun Chen	Supervised orthogonal tensor neighbourhood preserving embedding for face recognition	101
Jianjun Chen	Constraint-based sparsity preserving projections and its application on face recognition	106
Qinghui Wang, Wangyuan Huang	An effective toa-based scheme for mitigating the influence of the NLOS propagation	113
Kai Song, Caihong Niu	Incomplete character recognition technology in the license plate recognition system	119
Kunliang Liu, Jinming Huang	Research and implementation of 3D reconstruction base on multi-contours	125
You Lu, Xuefeng Xi, Ze Hua, Hongjie Wu, Ni Zhang	An abnormal user behaviour detection method based on partially labelled data	132
Xiaoxiong Wang	Landscape based on three dimensional SketchUp modelling to get visualization applications	142

## Operation Research and Decision Making

Jiansheng Peng	The robot path optimization of improved artificial fish-swarm algorithm	147
Hao-ran Shi, Kejian Liu	Research on lead-time reduction of two-stage supply chain based on stackelberg game	153
Dashan Chen	Effect analysis of speed guidance on traffic demand and driver compliance	159
Xueling Jiang, Shuijie Qin	Simulating a crowd with dynamic emotional transmission based on hidden Markov model	166
Xiangyang Ren, Yu Ren, Qingmei Li	Uncertain demand of farming-enterprise supply chain coordination based on the option contract	173
Fei Shao, BingHua Cheng	Traffic driven epidemic spreading in weighted networks with different routes	179
Guozhang Jiang, Chongwu Lei, Honghai Liu, Gongfa Li	Planning and scheduling model of production process in iron and steel enterprise	186
Gongfa Li, Yikun Zhang, Guozhang Jiang,	Production procedure optimization in iron and steel enterprise	192

Honghai Liu, Jia Liu		
Houxing Tang	Spatial effect of knowledge spillover on regional economic development: an empirical study from China	197
Xing Yu	A fuzzy set approach for a multi-period optimal portfolio selection model	204
Chonghuan Xu	A novel method of user interest drift detection engaging in individual background factors	208
Jing An, Bo Xu	Empirical study of C2C logistics customer satisfaction based on AHP and FCE	213
Bin Wang, Dashe Li, Shue Liu	Research and development of comprehensive communication experiment teaching system	218
Zhenfang He	Artificial neural network model of forecasting relative humidity in different humid and arid areas of China	225
Yulan Zhou	Dynamic pricing model of monopolistic manufacture based on the after-sale service	233
Jian Jin, Jianxiang Wang	Study on local government public expenditure and multi-factor productivity in China based on instrument variable model	240
QiuHong Zheng, Liangrong Song	Study on pre-loan evolutionary stable strategy of bank-enterprise for preventing moral hazard	247
Yuhui Xu, Qiuyue Luo	Study on the city planning for geological disasters defence based on the model of safe city planning	255
V Grekul, L Baydalina	A forecasting model of the economic efficiency of data centre construction project	261
<b>Nature Phenomena and Innovative Engineering</b>		
Hongyan Li	An improved energy-efficient distributed storage system	271
Zhijun Zhang, Hong Liu	Research on ontology-based literature retrieval model	281
Gongfa Li, Jia Liu, Guozhang Jiang, Honghai Liu, Wentao Xiao	Case-based reasoning intelligent prediction model of rotary kiln temperature	290
Chengbing He, Dongchao Chen	Study on sub synchronous resonance problem in series-compensated transmission system	294
Lv Ming, Xu Yan, Xin Nie, Huachen Pan, Haiqiang Liu	Experimental studies of differential heating for artificial upwelling	299
Shengli Liu, Xiang Gao, Pan Xu, Long Liu	Method of multi-feature fusion based on SVM and D-S evidence theory in Trojan detection	304
Yi Zhang, Xiuxia Yang, Weiwei Zhou	UAV trajectory optimization generation based on pythagorean hodograph curve	311
Juan Zhang, Hesong Jiang, Hong Jiang, Chunmei Chen	Reliable UDP over the air transfer in digital radio system	321
Wenqiang Chen	A genetic algorithm for the vehicle routing optimization problem of logistics park distribution	330
Wen-sheng Xiao, Zhong-yan Liu, Jian Liu, Han-chuan Wu	A novel method for identifying system modal parameters using stabilization diagram	335
Yunqi Tao, Dong Liu, Jie Cao, Jiang Xu	Decline in gas pressure influences the deformation and permeability of coal-containing methane	342
Ling Li	Income distribution of the bundled transmission of photovoltaic power plant based on DEA game model	348
Zhiqiang Yin, Lei Wang, Haifeng Ma	Optical measurement method for dynamic mechanical testing based on image grey level distribution difference model	354
Fei Cai, Wenjun Wang	Research on illumination invariance colour index algorithm based on colour ratio	360
Debao Yuan, Xueqian Hong, Shiwei Yu, Liangjian Li, Yanbao Zhao	Study on feasibility of CORS application in surface movement deformation monitoring in mining areas	365
Jianjun Dong	Study on improved elasto-plastic model for unsaturated soils based on Barcelona basic model	372
Wenyu Lv	Water trickling and roof falling of soft argillaceous roadways with its composite supporting and repairing technology	378
<b>Authors' Index</b>		<b>385</b>
<b>Cumulative Index</b>		<b>386</b>

*Editors' Remarks*

\*\*\*\*\*

**Imitation**

*by Edgar Allan Poe*

A dark unfathomed tide  
 Of interminable pride -  
 A mystery, and a dream,  
 Should my early life seem;  
 I say that dream was fraught  
 With a wild and waking thought  
 Of beings that have been,  
 Which my spirit hath not seen,  
 Had I let them pass me by,  
 With a dreaming eye!  
 Let none of earth inherit  
 That vision of my spirit;  
 Those thoughts I would control,  
 As a spell upon his soul:  
 For that bright hope at last  
 And that light time have past,  
 And my worldly rest hath gone  
 With a sigh as it passed on:  
 I care not though it perish  
 With a thought I then did cherish

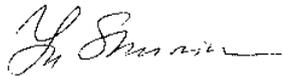
Edgar Allan Poe (1809-1849) \*

\*\*\*\*\*

This 18<sup>th</sup> volume No.6 presents actual papers on main topics of Journal specialization, namely, **Mathematical and Computer Modelling, Computer and Information Technologies, Operation Research and Decision Making and Nature Phenomena and Innovative Engineering.**

Our journal policy is directed on the fundamental and applied sciences researches, which are the basement of a full-scale modelling in practice. This edition is the continuation of our publishing activities. We hope our journal will be interesting for research community, and we are open for collaboration both in research and publishing. We hope that journal's contributors will consider the collaboration with the Editorial Board as useful and constructive.

**EDITORS**



**Yuri Shunin**



**Igor Kabashkin**

\* **EDGAR ALLEN POE** was born in Boston, January 19, 1809, and after a tempestuous life of forty years, he died in the city of Baltimore, October 7, 1849. His literary work may be summed up as follows: 1838- *The Narrative of Arthur Gordon Pym*; 1839- editor of Burton's *Gentleman's Magazine* Philadelphia, 1840 - editor of *Graham's Magazine*, 1845-*Tales of the Grotesque and Arabesque, The Raven*. The most remarkable of his tales are *The Gold Bug, The Fall of the House of Usher, The Murders of the Rue Morgue, The Purloined Letter, A Descent into Maelstrom* and *The Facts in the Case of M. Valdemar, The Raven* and *The Bells* alone would make the name of Poe immortal. The teachers of Baltimore placed a monument over his grave in 1875.



# Accuracy evaluation of deep belief networks with fixed-point arithmetic

Jingfei Jiang<sup>1\*</sup>, Rongdong Hu<sup>1</sup>, Luján Mikel<sup>2</sup>, Yong Dou<sup>1</sup>

<sup>1</sup>Science and Technology on Parallel and Distributed Processing Laboratory, National University of Defense Technology, ChangSha, Hunan 410073, China

<sup>2</sup>University of Manchester, Manchester, M13 9PL, UK

Received 12 June 2014, www.tsi.lv

---

## Abstract

Deep Belief Networks (DBNs) are state-of-art Machine Learning techniques and one of the most important unsupervised learning algorithms. Training DBNs is computationally intensive which naturally leads to investigate FPGA acceleration. Fixed-point arithmetic can be used when implementing DBNs in FPGAs to reduce execution time, but it is not clear the implications for accuracy. Previous studies have focused only on accelerators using some fixed bit-widths. A contribution of this paper is to demonstrate the bit-width effect on various configurations of DBNs in a comprehensive way by experimental evaluation. Explicit performance changing points are found using various bit-widths. The impact of sigmoid function approximation, required part of DBNs, is evaluated. A solution of mixed bit-widths DBN is proposed, fitting the bit-widths of FPGA primitives and gaining similar performance to the software implementation. Our results provide a guide to inform the design choices on bit-widths when implementing DBNs in FPGAs documenting clearly the trade-off in accuracy.

*Keywords:* deep belief network, fixed-point arithmetic, bit-width, FPGA

---

## 1 Introduction

Deep neural networks have become a “hot topic” in the Machine Learning community with successful results demonstrated with Deep Belief Networks (DBNs) [1], denoising autoencoder [2], sparse coding [3] and etc. DBNs have been shown to be among the best neural networks even for challenging recognition, mining and synthesis tasks. A DBN is built on a subset of neural networks known as Restricted Boltzmann Machine (RBM). Running a DBN is a time-consuming task due to its large scale and processing characteristics. Many experiments have often reported taking weeks, to search the large parameter space (numbers of layers and neurons, learning rate, momentum and all kinds of regulation terms) and calculate millions of parameters (weights and biases). One good example is Quoc et al. [4] who used a cluster in Google of 1,000 machines (16,000 cores) for a week to demonstrate the success of larger scale unsupervised learning from internet images recognition.

Reducing the execution time of the training phase and prediction of a DBN is one critical barrier which has restricted the mass adoption of DBNs. Interest in the acceleration of DBNs has built up in recent years. FPGAs are attractive platforms for accelerating DBNs. For example, a RBM of 256x256 nodes was tested on a platform of four Xilinx Virtex II FPGAs and gained a speedup of 145-fold over an optimized C program

running on a 2.8-GHz Intel processor [5]. Using Altera Stratix III FPGA, Kim et al. [6] also gained significant speedup for a 256x1024 RBM. Multi-FPGA solutions were discussed to determine the extensibility of RBM in [7, 8].

Existing works on FPGA implementations of neural networks often have vast and regular processing units to map neurons partially or wholly at a time. Weights and neuron values are stored in on-chip RAM during processing and are swapped out to off-chip memory after processing. It is too expensive to support a large number of floating-point units on chip and store values using the standard double precision floating-point representations in on-chip RAMs. Many of the previous attempts with FPGAs for neural networks implemented fixed bit-widths (8 bits, 16 bits or 32 bits). Bit-widths with integral multiple of bytes are convenient to align with other components (such as IP cores and user interfaces) and easier to design. Previous works have mainly analysed the impact of bit-widths on accuracy and execution time of old-style neural networks [9-11]. All reported RBM (a building component of DBN) designs on FPGA selected fixed-point arithmetic with a fixed bit-width as well, e.g. 16 bits in [6, 8] or 32 bits in [5] without analyzing in depth the implications for accuracy. Thus, it is not clear whether this kind of fixed bit-width is really the most suitable and area efficient for DBNs.

Using bit-width unequal to the machine word-length on a standard processor or GPU may rarely deliver any

---

\* *Corresponding author* e-mail: jingfeijiang@nudt.edu.cn

speedup. Programs need more instructions to do alignment and splicing which is not a negligible cost. On the other hand, speed and resource usage in FPGAs are more sensitive to the bit-width as many logics are mapped to fine-grain LUTs. As DBNs have grown in size, compared with old-style neural networks, to satisfy the learning demands of contemporary applications, resource saving due to narrower bit-widths has become more attractive to implement larger processing array in FPGAs. However, shrinking the bit-width may harm the convergence and accuracy of DBNs. Antony et al. [12] provided an initial study of the arithmetic effects on RBM for a specific network configuration. This paper reports a comprehensive study where in particular it improves the coverage of the variation of DBN and investigates how mixed bit-widths DBNs can offer a better accuracy and area efficiency. As it is expensive to implement exponential function and division operations directly on FPGA, it is important to understand the implications of approximation on the required sigmoid functions part of DBNs.

## 2 DBNs in a Nutshell

Our work is inspired by the original DBN of [1] and the idea of Stacked Denoising Auto-Encoder (SDAE) [2]. Hinton et al. [1] proposed an algorithm for learning deep networks based on a hierarchical probabilistic graphical model. A DBN is built on a structure of multi-layers RBMs. Each layer of RBM defines an energy function as a goal of minimization, which is represented as the negative log probability of a state between inputs (visible units) and outputs (hidden units):

$$E(v, h) = -\log P(v, h) = \frac{1}{2\sigma^2} \sum_{i \in v} v_i^2 - \frac{1}{\sigma^2} \left( \sum_{i \in v} a_i v_i + \sum_{j \in h} b_j h_j + \sum_{i,j} v_i h_j w_{ij} \right) \quad (1)$$

where,  $w_{ij}$  is the connection weight between visible unit  $v_i$  and hidden unit  $h_j$ ,  $a_i$  and  $b_j$  are biases of  $v_i$  and  $h_j$  respectively.  $\sigma$  is a parameter. In the case of using binary-valued visible units, the first term of Equation (1) will disappear [13]. Training the parameters  $w_{ij}$ ,  $a_i$  and  $b_j$  so as to minimize the energy can take the way of Gibbs Sampling by alternatively sampling each layer's units given the other layer, which uses conditional distributions to approximate the joint distribution. Hinton cut down the process into two steps, which crudely approximate the gradient of the log probability of the training data  $v^0$ :

$$\Delta w_{ij} = \frac{\partial \log(v^0)}{\partial w_{ij}} \approx \varepsilon \left( \langle v_i^0 h_j^0 \rangle - \langle v_i^{rec} h_j^{rec} \rangle \right) \quad (2)$$

The “*rec*” means the second step of Gibbs Sampling.  $\varepsilon$  is the learning rate. The gradient obtained from this simplification is like the gradient of another objective function called Contrastive Divergence (CD). Though it is a kind of approximation, it works well enough to

achieve satisfactory performance in many significant applications. Based on the network model and CD, the overall process of RBM is:

$$\begin{aligned} h^0 &= \text{logistic}(v^0 * W + a) \\ v^{rec} &= \text{logistic}(h^0 * W' + b), \\ h^{rec} &= \text{logistic}(v^{rec} * W + a) \\ \Delta W &= v^0 h^0 - v^{rec} h^{rec} \end{aligned} \quad (3)$$

where *logistic* is the logistic function which is labeled as a sigmoid function  $y = \frac{1}{1+e^{-x}}$ .

SDAE is another type of deep neural network, which is based on a different learning theory but has similar computations to DBN. SDAE denoises its inputs in a corruption level and then gets a distinctive property that even with a high capacity model it can avoid learning the identity mapping. Empirical results showed that SDAE can perform better than non-denoised ones with a suitable corruption level. We try to use the denoising idea on DBN to improve performance. This is done by first corrupting the initial input  $v$  to get a partially transformed version  $\tilde{v}$  by means of a stochastic mapping  $\tilde{v} \sim q_D(\tilde{v} | v)$ . The corrupted input  $\tilde{v}$  is then used to train the RBM using Equation (3).  $q_D$  can use additive Gaussian Noise, random zeroing noise, and salt-pepper noise as well [2]. Random zeroing noise which is most commonly used was selected in our experiments. A fixed percentage of randomly chosen units set their values to 0, while the others are left untouched.

From an information theoretic perspective, converting double precision floating-point arithmetic to fixed-point arithmetic will lose some information of inputs as well as intermediate data. Denoising DBN seems also lose information of inputs, just in a stochastic way. The training process becomes more “coarse” than before in both cases. The advantage of such approximation is that high-dimensional input loose the redundant and useless information during processing and then can learn features easier. The disadvantage is that some critical information may be lost and make the feature more indistinct to be learned. In SDAE, a suitable corruption level can make the advantages of inputs denoising outweigh its disadvantages. For the similar reason, a suitable bit-width may trade-off both-side effects well.

## 3 Experimental Methodology

For our experiments, we modified the floating-point versions of the original DBN (oDN) and the denoising DBN (dDN) into fixed-point versions and we compared them. The dDN version adds a corruption process before the pre-training of each RBM layer. The fixed-point versions take bit-widths as parameters, including bit-widths of neural units, weights, logistic function and random number generator, so it can run in any bit-width

configuration. Fixed-point versions of oDN and dDN were implemented by amplifying each data in RBM by a factor of  $2^{\text{fractional-width}}$  and truncating each operation result by a factor of  $2^{\text{bit-width}}$ , thus simulating the calculation process as a limited fixed-point one. Only pre-training was translated into a fixed-point version. Testing and fine-tuning were still computed using double precision floating-point. All experiments were done in Matlab2010a.

MNIST classification was selected as the objective application because of its popularity in machine learning studies and we use the most common configurations of DBNs in these studies. The dataset is 60,000 training and 10,000 testing samples of 28x28 pixel images of the digits. Three layers DBN with size of 784-400-400-400-10 and one layer DBN with size of 784-400-10 were built, with lower layers of unsupervised pre-training and the top layer of output logistic regression using softmax (multinomial logistic regression). The whole network was then fine-tuned as usual for multi-layers perceptrons, to minimize the output prediction error. The minibatch size for pre-training was 100 and the one for fine-tuning was 1000. Every minibatch in every epoch used different random numbers generated by the same random seed.

50 epochs of pre-training were used in order to find out the whole effect of fixed-point versions. 30 epochs of fine-tuning were used because the network in our configuration can get a rather good result after 30 epochs of fine-tuning and our aim was not to achieve best performance but to compare performance change among different options. We did not use Mean Squared Error (MSE) criterion in our experiments because MSE is sometimes incomparable among different configurations of DBN. Classification error on testing set was directly used as the criterion. For some fixed-point versions with narrower bit-widths, the more of pre-training epochs is not better to gain performance. Our experiments showed that sometimes the MSE between input data and reconstruction data became divergent as the pre-training progressed. Therefore, we used an early-stop strategy when MSE was not convergent during pre-training and recovered the previous update point, which had the smallest MSE as the final pre-training result.

The starting point where the pre-training procedure is initialized has some impact on DBN performance. We ran each version of DBN 40 times under 40 random seeds, which were decided by the "clock" value. Up to 120 random seeds were tried on some bit-width configurations. The result distributions are very similar

with the results running from 40 random seeds. Therefore, we thought 40 random seeds were objective enough to evaluate effect of the initialization point. We used several nodes of the TianHe-1A supercomputer [14] and all the experiments were completed in less than a month. We explored in detail bit-width configurations from 14 bits to 32 bits and found noticeable changing points.

#### 4 Effect of bit-width

When considering a fixed-point representation for real numbers, the integer part of a number mainly influences the representation scope while the fractional part mainly decides the precision. Overflow may affect the DBN performance heavily.

We ordered the results by the bit-width of the integer part. Figure 1 shows the test classification error distributions of three layers DBNs, obtained with 5 bits, 6 bits and 7 bits integers, as the fractional part increase (The X axis shows as "integer-decimal" pair denoting integer width and fractional part). In Figure 1(left), the noticeable thing is that 5 bits integer is not wide enough for scope representation. Therefore, the errors are very large and unstable. When the integer width increases one or two bits, the representation scope is mostly satisfied and the representation precision effects dominated. Configurations with 6 bits or 7 bits integer seem much more robust with respect to the random initialization seed and achieve better performance. The wider the fractional width is the better performance they can achieve. For different configurations with the same bit-widths (7-8 and 6-9, 7-10 and 6-11, 7-12 and 6-13), 6 bits integer perform almost the same as (or just a little better than) 7 bits integer (It can be explained that precision limitation harms the performance more severely when integer width is enough). We also evaluated 8 bits integer and got the same trend. It indicates that the changing points where most reasonable performance (below 1.5%, the best is about 1.2%) can be achieved for our three layers oDN is 6~7 bits integer and 10~13 bits fractional part. A bit-width of 19 bits is wide enough to achieve the best performance. In a FPGA implementation of such DBN, there is really a precision lost if 16 bits are used and there is a big waste if 32 bits are used, as some previous implementations do. We also evaluated wider bit-widths from 21 bits to 32 bits and the distributions are indistinguishable from the floating-point version, which confirms the conclusion.

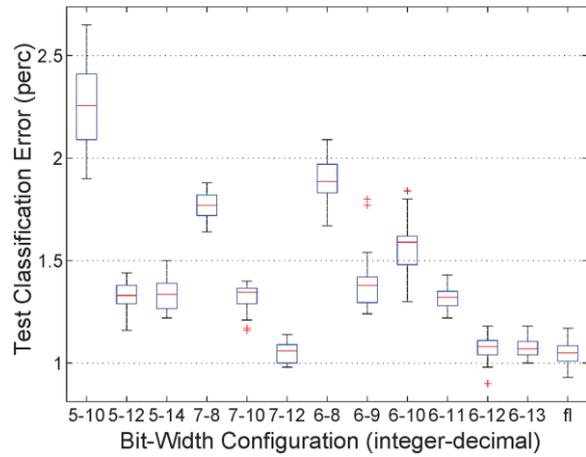
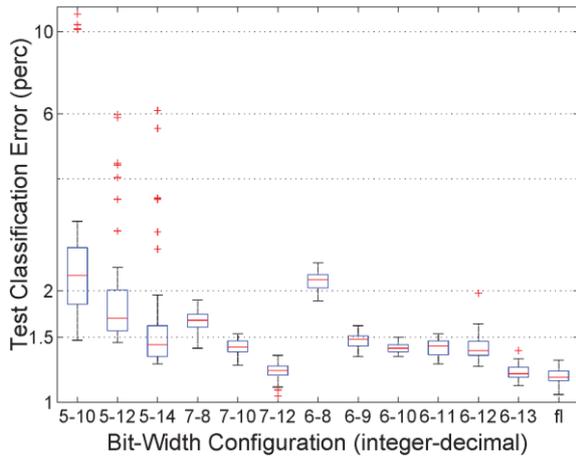


FIGURE 1 Performance of three layers oDN (left, Y axis is in log scale to show clearly) and dDN (right). fl: Floating-point version

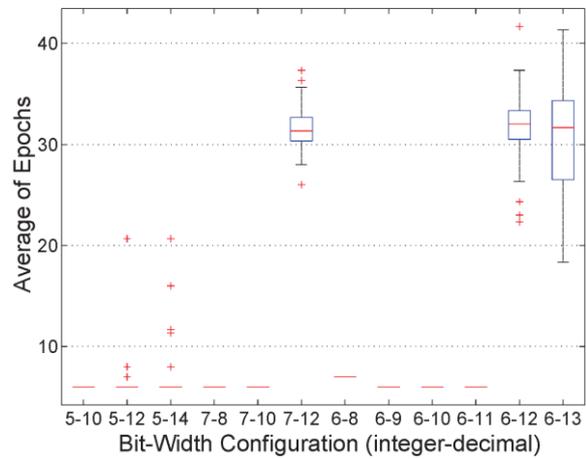
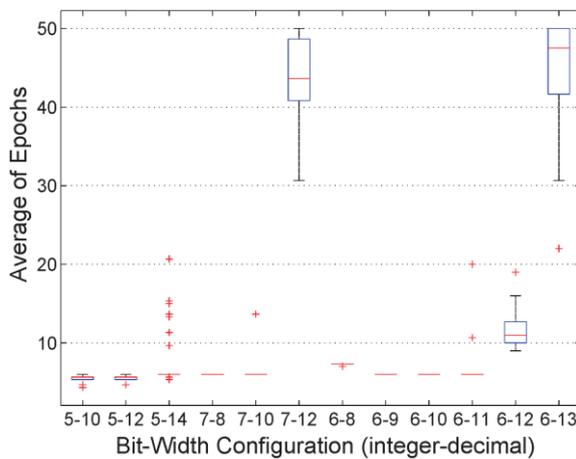


FIGURE 2 Epochs taken by configurations of Figure 1 in oDN (left) and dDN (right)

Figure 2 (left) shows the number of epochs (average of three layers) each configuration in Figure 1 (left) takes. It indicates the direct reason why narrower bit-widths get worse performance. DBNs with narrower bit-widths take less epochs up to converge during pre-training (under the control of early stop). There is a large difference among configurations: only the configurations achieving best performance (6-13 and 7-12) take as many epochs as the floating-point version (up to 50), others take rather a few epochs (below 20). It seems more difficult for narrow bit-width to converge efficiently. All comparisons were based on the same DBN parameters except the bit-width. If other parameters are adjusted distinctively, performance for narrow bit-width may be better. But we only wanted to compare the effects in the same situation. We also tried to adjust weight costs but got no better results. The good news is that using narrow bit-width not only reduces the executing time of DBN kernel assuming that narrower multiplication and addition in Equation (3) are calculated faster, but also reduces the number of epochs, thus reduces the whole executing time greatly.

Figure 1 (right) and Figure 2 (right) show the similar situations on dDN. The corruption level of 5% was used

for dDN, which is the suitable corruption level to gain better performance. dDN performs better than oDN, especially on stability. 18 bits width (6-12) can achieve best performance requiring epochs below 44. Better performance and stability of fixed-point dDN is attributed to input corruptions, which not only highlight the discard of redundant information but also neutralize some effects of bit-width shrink.

Figure 3 shows the performance comparisons between DBNs with three layers and DBNs with one layer. The overall performance of one layer DBN is off cause worse than three layers DBN. The performance variations of three layers DBNs are a little larger than one layer DBNs because of the better sensitivity of deeper DBN. For example, variations of one layer oDN and dDN from 1L6-9 to 1L6-13 are about 0.2%. Variations of three layers are about 0.3%. Comparing Figure 1 (left) with Figure 3 (left) and Figure 1 (right) with Figure 3 (right), the trend of variations are almost the same. 19 bits oDN and 18 bits dDN of one layer can gain best performances just as their three layers contrasts. These results indicate that bit-width influences on DBNs with different numbers of layers are relatively consistent.

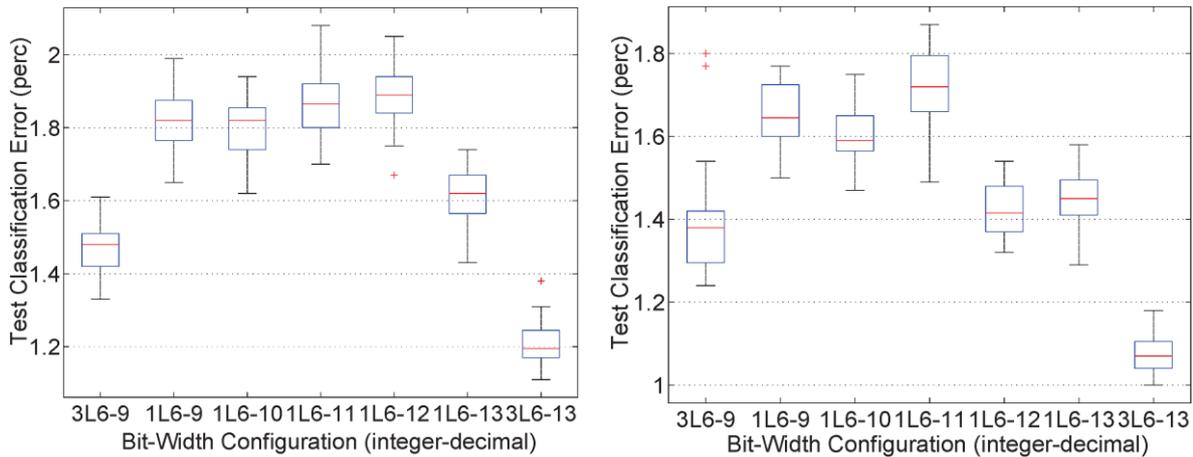


FIGURE 3 Bit-width impact on different layers of oDN (left) and dDN (right). 3L: three layers DBNs, 1L: one layer DBNs

**5 Effect of sigmoid function approximation**

In our experiments aforementioned, a software version of exponential function and division was used to calculate the logistic function. Only input and output were constrained by bit-width. As it is very expensive to implement exponential function and division directly on a FPGA, approximations applicable to FPGA implementation can be considered. Therefore, the results above may have been more optimistic than real implementation. The sigmoid function approximation impact should be evaluated.

The methods of sigmoid function approximation can be divided into two groups from the perspective of implementation complexity. One group includes algorithms based on higher order Taylor Series Expansion [15], Least Square Approximation [16] and Lookup Table combined with interpolation [7]. These algorithms need nonlinear functions implementation or large volume of memory. The latency and resources usage are relatively high. They are more suitable for the design choices of small number of units and high precision. The other group includes many Piecewise Linear Approximation of nonlinearity algorithms (PLAs) [10, 17, 18], which use linear functions and can be implemented on FPGA easily, but may be not so accurate. They are suitable for the design choices of vastly replicated units. FPGA implementations of DBNs need as many units as possible to execute in parallel. Accordingly, the same number of sigmoid function modules is needed. PLA is preferred in this situation. Whether the precision of PLAs will harm the DBN performance needs to be considered. Two PLAs with different precisions were used in our next experiment. One (PLA1) is from [18], which is a typical algorithm and used in an implementation of RBM [6]. The other (PLA2) was used in MLP-BP neural networks by Antony et al. [10]. Two algorithms are shown in Table 1.

PLAs have uniform structures like Table 1. PLA1 has 4 pieces and PLA2 has 3 pieces. The linear functions in

each piece, the numbers of pieces and the bit-width determine the actual precision of PLA implementation. A PLA module was built in Verilog according to Table 1, parameterizing bit-width, number of piece, input scope in each piece, bias number (such as the addends of 2, 5, 27 in PLA1 and 2, 56 in PLA2) and shift number (such as the divisors of 4, 8, 32 in PLA1 and 4, 64 in PLA2). PLA1 and PLA2 configured with 6 bits integer and different fractional parts were simulated in Modelsim. Software versions of sigmoid function with corresponding bit-widths were also run in Matlab.

TABLE 1 Piecewise Linear Approximation algorithms

PLA1		PLA2	
x	y	x	y
$0 \leq  x  < 1$	$y = ( x +2)/4$	$0 \leq  x  < 8/5$	$y = ( x +2)/4$
$1 \leq  x  < 19/8$	$y = ( x +5)/8$	$8/5 \leq  x  < 8$	$y = ( x +56)/64$
$19/8 \leq  x  < 5$	$y = ( x +27)/32$		
$ x  \geq 5$	$y = 1$	$ x  \geq 8$	$y = 1$
$x < 0$	$y = 1 - y$	$x < 0$	$y = 1 - y$

Figure 4 shows function curves and absolute errors (compared with its corresponding software versions) of PLA1 and PLA2, with 6 bits integer and 13 bits fractional part. The maximum precision difference between PLA1 and PLA2 is about 5% (6.8% of PLA2 vs 1.9% of PLA1). Figure 5 shows the maximum and mean absolute errors of PLAs in different fractional parts. It shows that the precisions are stable when fractional width is up to 9~10 bits for both PLAs. We selected these two PLAs as representatives to show how PLAs do with different precision lost affect DBN performance.

The sigmoid functions in oDN and dDN of one layer were replaced by PLA1 and PLA2 respectively. Figure 6 shows the results using PLA1. The software version and PLA1 configured with same bit-width get almost the same distributions on both oDN and dDN, which means PLA1 configured with the same bit-width as other operations has enough precision to benefit the overall performance.

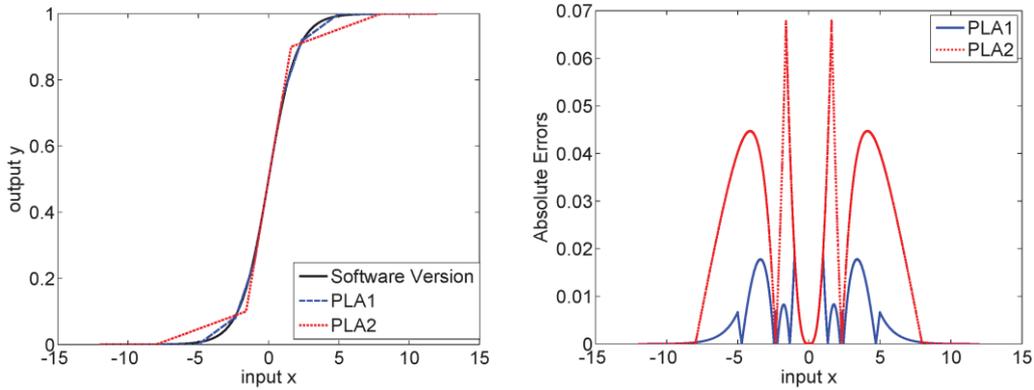


FIGURE 4 left: Sigmoid function curve of software version, PLA1 and PLA2. right: Absolute errors of PLA1 and PLA2 with 6 bits integer and 13 bits fractional part

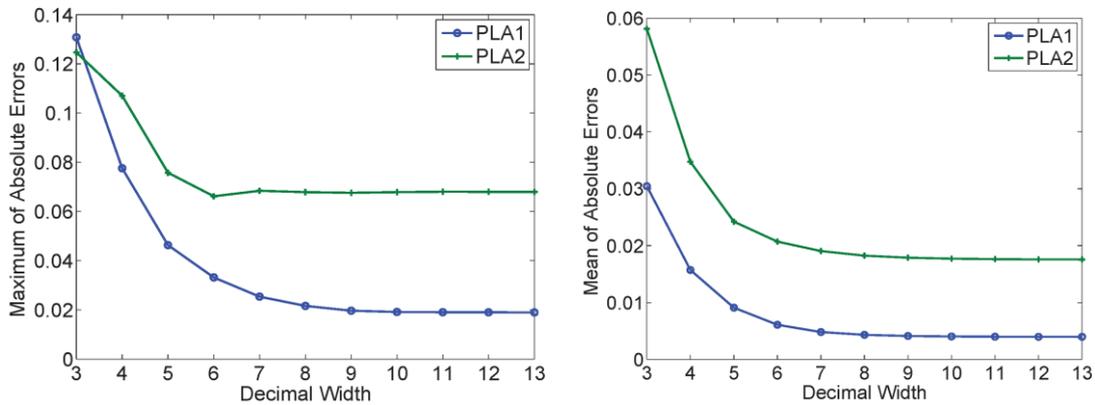


FIGURE 5 Maximum (left) and Mean (right) absolute errors of PLA1 and PLA2

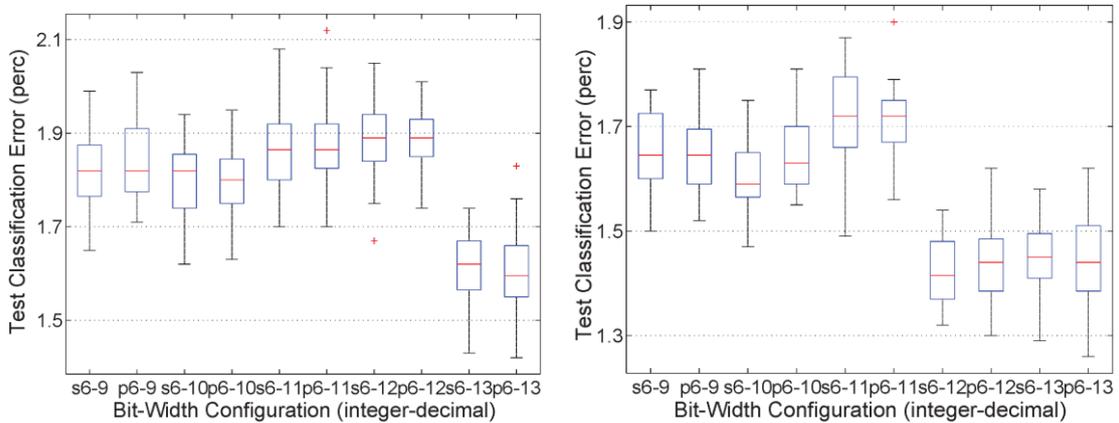


FIGURE 6 Performance of oDN (left) and dDN (right) using software version sigmoid function (s) and PLA1 (p)

Figure 7 shows the performance comparison between PLA1 and PLA2. The performance difference becomes a little larger as the bit-width increases. It means that PLA2 with lower precision may not satisfy the precision requirement of the whole DBN. Harm appears more clearly when more epochs are taken (when wider bit-width as 6-13 is used).

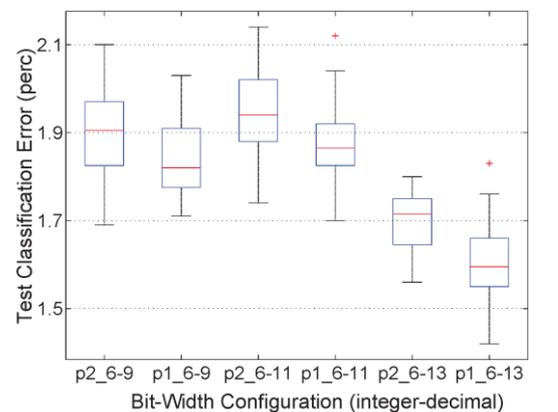


FIGURE 7 Performance of oDN using PLA1 (p1) and PLA2 (p2)

## 6 Mixed Bit-widths

Matrix operations for DBN include multiplications, accumulations and additions. Modern FPGAs supply built-in primitives to support such operations. For example, the primitive of DSP48E slice from Xilinx Inc. supports many functions such as multiply, multiply accumulate (MACC) and multiply add. The primitives have their favourite bit-widths. For example, One DSP48E contains one 25x18 two's complement multiplier, an adder, and an accumulator. The output of DSP48E is commonly wider than sum of inputs bit-widths for accumulation (up to 48 bits). Block RAM is built by primitives, which also have fixed bit-width (RAMB18 and RAMB36).

It is somehow unfortunate that our result of 19 bits is the narrowest bit-width to achieve best performance on oDN. Table 2 show the resource usage of a MACC unit generated from Xilinx Core Generator. It shows that resource cost (mainly the DSP48s) will double or triple if the bit-width DBN required exceeds the basic bit-width of primitive. Meanwhile, multiple primitives will cascade to form a processing unit, which will increase the pipeline stages of the unit. The latency and control complexity will increase finally. Multiplier resources like DSP48 are relatively precious in a FPGA (several tens or hundreds). Therefore, using one suitable bit-width for DBN may not be area efficient.

TABLE 2 Resources usage of a MACC

Description	Values					
multiplier width	18	18	19	19	19	25
multiplcand width	18	25	19	25	30	25
DSP48s	1	1	3	3	2	2
LUTs	0	0	0	0	50	27
Flip-Flops	0	0	0	0	99	53

A mixed bit-widths solution is proposed to accommodate the requirement of hardware primitives. In DBNs, multiply operations are multiplication of a neural unit value and a weight or multiplication of two neural unit values. Weights record the feature values, which need higher precision, while neural units can be corrupted to some extent according to our experiments. Therefore, a narrower bit-width can be used for neural units and a wider one can be used for weights. A mixed bit-widths version of one layer DBN was modified and evaluated. 8 bits integer and 17 bits fractional part were used for weights (fitting in one DSP48). 6 bits and several narrow decimals were used for neural units. PLA1 was used for sigmoid function with the same bit-width of neural units. Figure 8 shows the results. It is clear that the mixed bit-

## References

- [1] Hinton G, Osindero S, Teh Y 2006 A fast learning algorithm for deep belief nets *Neural computation* 18(7) 1527-54
- [2] Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol P 2010 Stacked denoising autoencoders: Learning useful representations in a deep network with a local de-noising criterion *Journal of Machine Learning Research* 11 3371-408

width configurations can gain best performances like 6-13. It further indicates that weight precision dominates the overall accuracy.

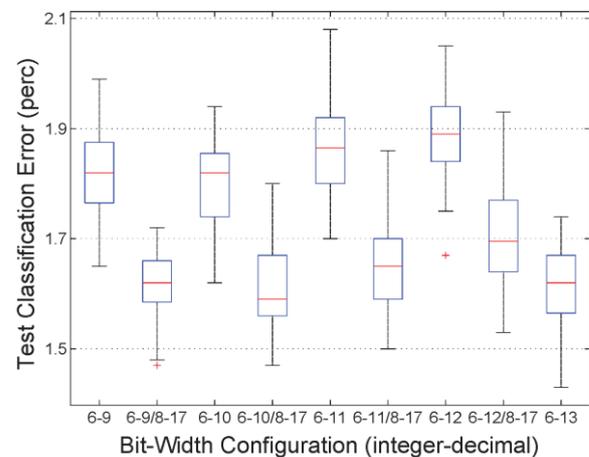


FIGURE 8 Performance of oDN using mixed bit-widths

## 7 Conclusions

Our work gives a comprehensive evaluation for implementing DBNs on FPGAs by studying a wide range of bit-width achieving best performance and area efficiency. Bit-width impacts show similar trend on different layers of DBNs, but are a little different between oDN and dDN. The PLA with higher precision can satisfy the overall DBN, but the other with lower precision does become the precision bottleneck of DBN. From these results, a mixed bit-widths solution is proposed. Assigning different bit-widths to neural units and weights can fit hardware primitives better and gain better performance. The control complexity implementing irregular bit-width (not integral multiple of bytes) seems a little high. But our design experience on a memory sub-system of DBN accelerators supporting various bit-widths has shown that it is not as difficult as it may sound. The cost is only little in hardware and does not affect the critical path.

## Acknowledgments

This work is funded by National Science Foundation of China (number 61303070) in cooperation with Dr. Lujan who is supported by a Royal Society University Research Fellowship. Dr. Jingfei Jiang is an academic visitor at University of Manchester. We acknowledge Antony W. Savich for his feedback and TianHe-1A supercomputing system service.

- [3] Lee H, Ekanadham C, Ng A 2008 Sparse deep belief net model for visual area v2 *Advances in neural information processing systems* 20 873-80
- [4] Le Q, Monga R, Devin M, Corrado G, Chen K, Ranzato M, Dean J, Ng A 2011 Building high-level features using large scale unsupervised learning. preprint arXiv:1112.6209
- [5] Ly D, Chow P 2009 A multi-fpga architecture for stochastic restricted Boltzmann machines: *International Conference on Field*

- Programmable Logic and Applications, (Czech Republic, August 31-September 2, 2009)* pp 168-73
- [6] Kim S, McMahon P, Olukotun K 2010 A large-scale architecture for restricted boltzmann machines *Proc. of the 18th IEEE Annual International Symposium on Field-Programmable Custom Computing Machines (Charlotte, North Carolina, May 2-4, 2010)* pp201-208
- [7] Le Ly D, Chow P 2010 *IEEE Transactions on Neural Networks* 21(11) 1780-92
- [8] Lo C, Chow P 2011 Building a multi-fpga virtualized restricted boltzmann machine architecture using embedded mpi *Proc. of the 19th ACM/SIGDA international symposium on Field programmable gate arrays (Monterey, California, February 27-1 March 1, 2011)* pp 189-98
- [9] Gomperts A, Ukil A, Zuruh F 2011 *IEEE Transactions on Industrial Informatics* 7(1) 78-89
- [10] Savich A, Moussa M, Areibi S 2007 *IEEE Transactions on Neural Networks* 18(1) 240-52
- [11] Draghici S 2002 On the capabilities of neural networks using limited precision weights *Neural Networks* 15(3) 395-414
- [12] Savich A, Moussa M 2011 Resource efficient arithmetic effects on rbm neural network solution quality using mnist *International Conference on Reconfigurable Computing and FPGAs (Cancun, Mexico, Nov 30- Dec 2, 2011)* pp 35-40
- [13] Hinton G, Salakhutdinov R 2006 Reducing the dimensionality of data with neural networks *Science* 313(5786) 504-7
- [14] National supercomputer centre in tianjin <http://www.nssc-tj.gov.cn/> 16 Aug 2013.
- [15] Arroyo Leon M, Ruiz Castro A, Leal Ascencio R 1999 An artificial neural network n a field programmable gate array as a virtual sensor *Proc. of the third International Workshop on Design of Mixed-Mode Integrated Circuits and Applications (Cat. No.99EX303)* pp 114-7
- [16] Al-Nsour M, Abdel Aty Zohdy H 1998 Implementation of programmable digital sigmoid function circuit for neuro-computing *Proc. of the Midwest Symposium on Circuits and Systems* pp 571-4
- [17] Alippi C, Storti Gajani G 1991 Simple approximation of sigmoidal functions: realistic design of digital neural networks capable of learning *Proc. of the IEEE International Symposium on Circuits and Systems* pp 1505-8
- [18] Amin H, Curtis K, Hayes Gill B 1997 Piecewise linear approximation applied to nonlinear function of a neural network *IEE Proc. of Circuits, Devices and Systems* 144 pp 313-7

Authors	
	<p><b>Jingfei Jiang, born in 1974, Inner Mongolia, China</b></p> <p><b>Current positions, grades:</b> She was a lecturer in School of Computer from 2004 to 2006 and is an associate professor at the same school from 2007 until now, founded by Science and Technology on Parallel and Distributed Processing Laboratory.</p> <p><b>University studies:</b> She was awarded BS. (1997), MS. (2000) and Ph.D. (2004) degrees in Computer Science by the University of School of Computer, National University of Defense Technology.</p> <p><b>Scientific interests:</b> She works in high-performance embedded system design and reconfigurable computing. More recently she has been involved in acceleration methods of machine learning algorithms such as Deep Belief Network and Stacked Auto-encoders.</p>
	<p><b>Rongdong Hu, born in 1986, Chongqing, China</b></p> <p><b>Current positions, grades:</b> He is a PhD student under the supervision of Prof. Guangming Liu, who is the vice chief designer of the well-known TH-1A supercomputer.</p> <p><b>University studies:</b> Master Degree at the same school in 2010 for work in Hierarchical Parallel Mass Storage System.</p> <p><b>Scientific interests:</b> cloud computing and statistical learning. He aim to use the intelligent techniques to improve the efficiency of cloud resource management - maximizing resource utilization, reducing energy consumption and cost of services, while ensuring the quality of cloud services.</p>
	<p><b>Mikel Luján, born in 1975, San Sebastian, Spain</b></p> <p><b>Current positions, grades:</b> After graduation he worked as a postdoctoral researcher in the Centre for Novel Computing at the University of Manchester. In 2005 Mikel started working for Sun Microsystem Research Laboratories in California. In late 2006 Mikel returned back to Manchester as a Career Development Fellow (cf. Research Assistant Professor), but now as a member of the Advanced Processor Technologies Group. In October 2009 he started his Royal Society University Research Fellowship on how to co-design future many-core architectures and managed virtual execution environments.</p> <p><b>University studies:</b> M.Phil. (1999) and Ph.D. (2002) degrees in Computer Science by the University of Manchester.</p>
	<p><b>Yong Dou, born in 1966, Jilin, China</b></p> <p><b>Current positions, grades:</b> He is the director of Science and Technology on Parallel and Distributed Processing Laboratory.</p> <p><b>Scientific interests:</b> design and implementation of the high performance accelerators. Now he is the project leader building large-scale parallel computers for deep learning applications in National University of Defense Technology.</p>

# On denotational semantics of the complex event query language STeCEQL

Huiyong Li\*, Yixiang Chen

*Software Engineering Institute, East China Normal University, N. Zhongshan Rd. 3663, Shanghai, China*

*Received 12 June 2014, www.tsi.lv*

---

## Abstract

With the complex event processing technology has been widely used in processing the information of the internet of things, many scholars have proposed a lot of event query languages(EQL) for different scenarios. Early scholars generally study the operational semantics of EQL. Recently, many researchers begin to pay attention to the correctness of the operational semantics of the EQL. Some researchers have shown the correctness of the operational semantics by proven the equivalence between the denotational semantics and the operational semantics of EQL. The internet of vehicles is an important research branch of internet of things and it has a very wide range of applications. STeCEQL is a spatial and temporal constraint EQL for the internet of vehicles. In this paper, we focus on the correctness of the operational semantics of STeCEQL. We mainly establish the denotational semantics of STeCEQL. Finally, we prove the equivalence between the two semantics of STeCEQL. Therefore, the operational semantics of STeCEQL are correct.

*Keywords:* Complex Event Query Language, Internet of things, Mobile System, Denotational Semantics, Operational Semantics

---

## 1 Introduction

In recent years, many researchers have concerned the internet of things and they has achieved a great deal of results [1]. Internet of vehicles is an important kind of the internet of things and it has very broad applications. Unlike the other internet of things, there are a lot of non-moving agents in internet of vehicles and many fast moving agents in it. All kinds of sensors of agents in the internet of vehicles produce great amount of temporal, spatial and other data. Meanwhile, the internet of vehicles is a performance critical system, which requires real-time processing the data in the system [2, 3]. However, the database technology cannot solve the daunting task.

In order to real-time processing these data of the internet of vehicles, some researchers have introduced the complex event processing technology into it. The complex event processing technology is filtering the amounts of data flow into the events by the EQL. When there are some events occurs, the system will real-time or near real-time to make the appropriate treatment, which based on the predefined rules base. Moody has proposed an EQL SpaTec and it has been applied to monitoring the bus system of London [4, 5]. Jin has proposed an EQL CPSL and it can describe the relationship between the properties of the internet of vehicles [6]. We have proposed STeCEQL and given its syntax and the operational semantics, which can effectively describe the internet of vehicles.

The operational semantics is an important means to describe the computer language. In the early studies, the

researchers only give the operational semantics of EQL. Zhu has proposed an EQL SEL and given its operational semantics [7]. Seiriö has proposed an EQL ruleCore and given its operational semantics [8]. Wu has proposed an EQL SASE and given its operational semantics [9]. Demers has proposed an EQL Cayuga and given its operational semantics [10].

In recent years, some researchers begin to concern the correctness of the EQL's operational semantics. Michael has proposed an EQL XChange and given its operational semantics and the denotational semantics [11]. Finally, he has demonstrated the equivalence of two semantics. Darko has proposed an EQL ETALIS and demonstrated the equivalence of its two semantics [12]. The denotational semantics is more abstract than the operational semantics. The equivalence of two semantics is often used to verify the correctness of the operational semantics.

Therefore, we establish the denotational semantics of STeCEQL and proved the equivalence between the two semantics of STeCEQL in this paper. The remainder of this paper is organized as follows: Section 2 restates the syntax of STeCEQL. Section 3 defines the denotational semantics of STeCEQL. Section 4 proves the equivalence of two semantics of STeCEQL by structural inductive method. The last Section concludes this paper.

## 2 Syntax and operational semantics of STeCEQL

The STeCEQL can express the base events of the internet of vehicles and the complex events composed by the base

---

\* *Corresponding author* e-mail: lihuiyongchina@126.com

events in a specific relationship. The syntax of the STeCEQL is as follows:

**ABexp:**

$attribute ::= true \mid false \mid x_a = a \mid x_a \neq a$

$\mid x_a > a \mid x_a \geq a \mid x_a < a \mid x_a \leq a$

$\mid attribute_0 \wedge attribute_1 \mid attribute_0 \vee attribute_1$

**TBexp:**

$time ::= true \mid false \mid x_t \text{ BEFORE } t \mid x_t \text{ AFTER } t$

$\mid x_t \text{ EQUAL } t \mid x_t \text{ OVERLAP } t \mid x_t \text{ DURING } t$

$\mid time_0 \vee time_1 \mid time_0 \wedge time_1$

**LBexp:**

$location ::= true \mid false \mid x_l \text{ EQ } l \mid x_l \text{ OP } l \mid x_l \text{ IN } l$

$\mid x_l \text{ NORTH } l \mid x_l \text{ SOUTH } l \mid x_l \text{ EAST } l \mid x_l \text{ WEST } l$

$\mid x_l \text{ NORTHWEST } l \mid x_l \text{ NORTHEAST } l$

$\mid x_l \text{ SOUTHWEST } l \mid x_l \text{ SOUTHEAST } l$

$\mid location_0 \vee location_1 \mid location_0 \wedge location_1$

**DBexp:**

$direction ::= true \mid false \mid x_d = d \mid x_d \neq d$

**EBexp:**

$e ::= agent^{time}(attribute1; attribute2; attribute3 \dots)$

$\mid agent^{location}(attribute1; attribute2; attribute3 \dots)$

$\mid agent^{time}_{location}(attribute1; attribute2; attribute3 \dots)$

$\mid agent^{time}_{(location, direction)}(attribute1; attribute2; \dots)$

**CEexp:**

$ce ::= e1 \wedge e2 \mid e1 \vee e2$

The operational semantics of the STeCEQL is as follows:

**ABexp:**

$\langle true, \sigma \rangle \rightarrow true$

$\langle false, \sigma \rangle \rightarrow false$

$\langle x_a = a, \sigma \rangle \rightarrow true, \text{ if } \sigma(x_a) = a$

$\langle x_a = a, \sigma \rangle \rightarrow false, \text{ if } \sigma(x_a) \neq a$

$\langle x_a \neq a, \sigma \rangle \rightarrow true, \text{ if } \sigma(x_a) \neq a$

$\langle x_a \neq a, \sigma \rangle \rightarrow false, \text{ if } \sigma(x_a) = a$

$\langle x_a > a, \sigma \rangle \rightarrow true, \text{ if } \sigma(x_a) > a$

$\langle x_a > a, \sigma \rangle \rightarrow false, \text{ if } \sigma(x_a) \leq a$

$\langle x_a \geq a, \sigma \rangle \rightarrow true, \text{ if } \sigma(x_a) \geq a$

$\langle x_a \geq a, \sigma \rangle \rightarrow false, \text{ if } \sigma(x_a) < a$

$\langle x_a < a, \sigma \rangle \rightarrow true, \text{ if } \sigma(x_a) < a$

$\langle x_a < a, \sigma \rangle \rightarrow false, \text{ if } \sigma(x_a) \geq a$

$\langle x_a \leq a, \sigma \rangle \rightarrow true, \text{ if } \sigma(x_a) \leq a$

$\langle x_a \leq a, \sigma \rangle \rightarrow false, \text{ if } \sigma(x_a) > a$

$\frac{\langle attribute_0, \sigma \rangle \rightarrow b_0 \quad \langle attribute_1, \sigma \rangle \rightarrow b_1}{\langle attribute_0 \wedge attribute_1, \sigma \rangle \rightarrow b},$

if  $b_0 = true$  and  $b_1 = true, b = true$ ; else  $b = false$

$\frac{\langle attribute_0, \sigma \rangle \rightarrow b_0 \quad \langle attribute_1, \sigma \rangle \rightarrow b_1}{\langle attribute_0 \vee attribute_1, \sigma \rangle \rightarrow b},$

if  $b_0 = true$  or  $b_1 = true, b = true$ ; else  $b = false$

**TBexp:**

$\langle true, \sigma \rangle \rightarrow true$

$\langle false, \sigma \rangle \rightarrow false$

$\langle x_t \text{ BEFORE } t, \sigma \rangle \rightarrow true, \text{ if } \sigma(x_t).endn < t.start1$

$\langle x_t \text{ BEFORE } t, \sigma \rangle \rightarrow false, \text{ if } \sigma(x_t).endn \geq t.start1$

$\langle x_t \text{ AFTER } t, \sigma \rangle \rightarrow true, \text{ if } \sigma(x_t).start1 > t.endn$

$\langle x_t \text{ AFTER } t, \sigma \rangle \rightarrow false, \text{ if } \sigma(x_t).start1 \leq t.endn$

$\langle x_t \text{ EQUAL } t, \sigma \rangle \rightarrow true,$

if  $(\forall i \in N. \sigma(x_t).starti = t.starti \text{ and } \sigma(x_t).endi = t.endi)$

$\langle x_t \text{ EQUAL } t, \sigma \rangle \rightarrow false,$

if  $(\exists i \in N. \sigma(x_t).starti \neq t.starti \text{ and } \sigma(x_t).endi \neq t.endi)$

$\langle x_t \text{ OVERLAP } t, \sigma \rangle \rightarrow true,$

if  $(\sigma(x_t).endn \geq t.start1 \text{ and } \sigma(x_t).endn \leq t.endn)$

or  $(\sigma(x_t).start1 \geq t.start1 \text{ and } \sigma(x_t).start1 \leq t.endn)$

$\langle x_t \text{ OVERLAP } t, \sigma \rangle \rightarrow false,$

if  $\sigma(x_t).endn < t.start1$  or  $\sigma(x_t).start1 > t.endn$

$\langle x_t \text{ DURING } t, \sigma \rangle \rightarrow true,$

if  $\sigma(x_t).start1 \geq t.start1$  and  $\sigma(x_t).end1 \leq t.endn$

$\langle x_t \text{ DURING } t, \sigma \rangle \rightarrow false,$

if  $\sigma(x_t).start1 < t.start1$  or  $\sigma(x_t).end1 > t.endn$

$\frac{\langle time, \sigma \rangle \rightarrow true}{\langle \neg time, \sigma \rangle \rightarrow false}$

$$\frac{\langle time, \sigma \rangle \rightarrow false}{\langle \neg time, \sigma \rangle \rightarrow true}$$

$$\frac{\langle time_0, \sigma \rangle \rightarrow b_0 \quad \langle time_1, \sigma \rangle \rightarrow b_1}{\langle time_0 \wedge time_1, \sigma \rangle \rightarrow b},$$

if  $b_0 = true$  and  $b_1 = true, b = true; else b = false$

$$\frac{\langle time_0, \sigma \rangle \rightarrow b_0 \quad \langle time_1, \sigma \rangle \rightarrow b_1}{\langle time_0 \vee time_1, \sigma \rangle \rightarrow b},$$

if  $b_0 = true$  or  $b_1 = true, b = true; else b = false$

**Lexp:**

$$\langle true, \sigma \rangle \rightarrow true$$

$$\langle false, \sigma \rangle \rightarrow false$$

$$\langle x_l EQ l, \sigma \rangle \rightarrow true,$$

if  $(\forall i \in N. \sigma(x_l).row_i = l.row_i \text{ and } \sigma(x_l).column_i = l.column_i)$

$$\langle x_l EQ l, \sigma \rangle \rightarrow false,$$

if  $(\exists i \in N. \sigma(x_l).row_i \neq l.row_i \text{ or } \sigma(x_l).column_i \neq l.column_i)$

$$\langle x_l OPl, \sigma \rangle \rightarrow true,$$

if  $(\exists i, j \in N. \sigma(x_l).row_i = l.row_j \text{ and } \sigma(x_l).column_i = l.column_j)$

$$\langle x_l OPl, \sigma \rangle \rightarrow false,$$

if  $(\forall i, j \in N. \sigma(x_l).row_i = l.row_j \text{ and } \forall \sigma(x_l).column_i = l.column_j)$

$$\langle x_l IN l, \sigma \rangle \rightarrow true, \text{ if } \sigma(x_l) \subset l$$

$$\langle x_l IN l, \sigma \rangle \rightarrow false, \text{ if } \sigma(x_l) \not\subset l$$

$$\langle x_l NORTH l, \sigma \rangle \rightarrow true,$$

if  $(\forall i, j \in N. \sigma(x_l).row_i < l.row_j \text{ and } \sigma(x_l).column_j = l.column_j)$

$$\langle x_l NORTH l, \sigma \rangle \rightarrow false,$$

if  $(\exists i, j \in N. \sigma(x_l).row_i \geq l.row_j \text{ or } \sigma(x_l).column_j \neq l.column_j)$

$$\frac{\langle location, \sigma \rangle \rightarrow true}{\langle \neg location, \sigma \rangle \rightarrow false}$$

$$\frac{\langle location, \sigma \rangle \rightarrow false}{\langle \neg location, \sigma \rangle \rightarrow true}$$

$$\frac{\langle location_0, \sigma \rangle \rightarrow b_0 \quad \langle location_1, \sigma \rangle \rightarrow b_1}{\langle location_0 \wedge location_1, \sigma \rangle \rightarrow b},$$

if  $b_0 = true$  and  $b_1 = true, b = true; else b = false$

$$\frac{\langle location_0, \sigma \rangle \rightarrow b_0 \quad \langle location_1, \sigma \rangle \rightarrow b_1}{\langle location_0 \vee location_1, \sigma \rangle \rightarrow b},$$

if  $b_0 = true$  or  $b_1 = true, b = true; else b = false$

**DBexp:**

$$\langle true, \sigma \rangle \rightarrow true$$

$$\langle false, \sigma \rangle \rightarrow false$$

$$\langle x_d = d1, \sigma \rangle \rightarrow true, \text{ if } \sigma(x_d) = d1$$

$$\langle x_d = d1, \sigma \rangle \rightarrow false, \text{ if } \sigma(x_d) \neq d1$$

$$\langle x_d \neq d1, \sigma \rangle \rightarrow true, \text{ if } \sigma(x_d) \neq d1$$

$$\langle x_d \neq d1, \sigma \rangle \rightarrow false, \text{ if } \sigma(x_d) = d1$$

**EBexp:**

$$\frac{\langle time, \sigma \rangle \rightarrow b1 \quad \langle a1, \sigma \rangle \rightarrow b2 \quad \langle a2, \sigma \rangle \rightarrow b3 \quad \dots}{\langle agent^{time}(attribute1; attribute2; \dots), \sigma \rangle \rightarrow true},$$

if  $\forall b \in (b1, b2, b3, \dots), b = true$

$$\frac{\langle time, \sigma \rangle \rightarrow b1 \quad \langle a1, \sigma \rangle \rightarrow b2 \quad \langle a2, \sigma \rangle \rightarrow b3 \quad \dots}{\langle agent^{time}(attribute1; attribute2; \dots), \sigma \rangle \rightarrow false},$$

if  $\exists b \in (b1, b2, b3, \dots), b = false$

$$\frac{\langle t, \sigma \rangle \rightarrow b1 \quad \langle l, \sigma \rangle \rightarrow b2 \quad \langle a1, \sigma \rangle \rightarrow b3 \quad \langle a2, \sigma \rangle \rightarrow b4 \quad \dots}{\langle agent_{location}^{time}(attribute1; attribute2; \dots), \sigma \rangle \rightarrow true},$$

if  $\forall b \in (b1, b2, b3, b4, \dots), b = true$

$$\frac{\langle t, \sigma \rangle \rightarrow b1 \quad \langle l, \sigma \rangle \rightarrow b2 \quad \langle a1, \sigma \rangle \rightarrow b3 \quad \langle a2, \sigma \rangle \rightarrow b4 \quad \dots}{\langle agent_{location}^{time}(attribute1; attribute2; \dots), \sigma \rangle \rightarrow false},$$

if  $\exists b \in (b1, b2, b3, b4, \dots), b = false$

$$\frac{\langle t, \sigma \rangle \rightarrow b1 \quad \langle l, \sigma \rangle \rightarrow b2 \quad \langle d, \sigma \rangle \rightarrow b3 \quad \langle a1, \sigma \rangle \rightarrow b4 \quad \langle a2, \sigma \rangle \rightarrow b5 \quad \dots}{\langle agent_{(l,d)}^{time}(attribute1; attribute2; \dots), \sigma \rangle \rightarrow true},$$

if  $\forall b \in (b1, b2, b3, b4, b5, \dots), b = true$

$$\frac{\langle t, \sigma \rangle \rightarrow b1 \quad \langle l, \sigma \rangle \rightarrow b2 \quad \langle d, \sigma \rangle \rightarrow b3 \quad \langle a1, \sigma \rangle \rightarrow b4 \quad \langle a2, \sigma \rangle \rightarrow b5 \quad \dots}{\langle agent_{(location,direction)}^{time}(attribute1; attribute2; \dots), \sigma \rangle \rightarrow false},$$

if  $\exists b \in (b1, b2, b3, b4, b5, \dots), b = false$

**CEBexp:**

$$\frac{\langle e1, \sigma \rangle \rightarrow b1 \quad \langle e2, \sigma \rangle \rightarrow b2}{\langle e1 \wedge e2, \sigma \rangle \rightarrow true}, \text{ if } \forall b \in (b1, b2), b \equiv true$$

$$\frac{\langle e1, \sigma \rangle \rightarrow b1 \quad \langle e2, \sigma \rangle \rightarrow b2}{\langle e1 \wedge e2, \sigma \rangle \rightarrow false}, \text{ if } \exists b \in (b1, b2), b \equiv false$$

$$\frac{\langle e1, \sigma \rangle \rightarrow b1 \quad \langle e2, \sigma \rangle \rightarrow b2}{\langle e1 \vee e2, \sigma \rangle \rightarrow true}, \text{ if } \exists s \in (s1, s2), b \equiv true$$

$$\frac{\langle e1, \sigma \rangle \rightarrow b1 \quad \langle e2, \sigma \rangle \rightarrow b2}{\langle e1 \vee e2, \sigma \rangle \rightarrow false}, \text{ if } \forall b \in (b1, b2), b \equiv false$$

### 3 Denotational semantics of STeCEQL

Let the states set  $\Sigma$  is composed by the function  $\sigma$  that from the storage set to different attribute values set. And then,  $\sigma(X)$  is the value of the storage unit X under the state  $\sigma$ . The ordered pair  $\langle \text{attribute}, \sigma \rangle \rightarrow \text{true}$  means that the value of the expression **attribute** is true under the state  $\sigma$ . The value of the complex event expressions is Boolean. Let Boolean set is  $\mathbf{B}=\{\text{true}, \text{false}\}$  and the element of the set express by **b**. Therefore, in the STeCEQL, the denotational functions of the all kinds of Boolean expressions are the mappings from the states set  $\Sigma$  to the Boolean set  $\mathbf{B}$ .

Numeric Boolean expressions **attribute**  $\in$  **ABexp**, denotational function  $\mathcal{A}[\text{attribute}]: \Sigma \rightarrow B$ .

Temporal Boolean expressions **time**  $\in$  **TBexp**, denotational function  $\mathcal{T}[\text{time}]: \Sigma \rightarrow B$ .

Spatial Boolean expressions **location**  $\in$  **LBexp**, denotational function  $\mathcal{L}[\text{location}]: \Sigma \rightarrow B$ .

Directional Boolean expressions **direction**  $\in$  **DBexp**, denotational function  $\mathcal{D}[\text{direction}]: \Sigma \rightarrow B$ .

Event Boolean expressions **e**  $\in$  **EBexp**, denotational function  $\mathcal{E}[e]: \Sigma \rightarrow B$ .

We define the denotational semantic function by the structural induction as below:

$$\mathcal{A}: \text{ABexp} \rightarrow (\Sigma \rightarrow B)$$

$$\mathcal{T}: \text{TBexp} \rightarrow (\Sigma \rightarrow B)$$

$$\mathcal{L}: \text{LBexp} \rightarrow (\Sigma \rightarrow B)$$

$$\mathcal{D}: \text{DBexp} \rightarrow (\Sigma \rightarrow B)$$

$$\mathcal{E}: \text{EBexp or CEBexp} \rightarrow (\Sigma \rightarrow B)$$

**ABexp:**

$$\mathcal{A}[\text{true}] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma\}$$

$$\mathcal{A}[\text{false}] = \{(\sigma, \text{false}) \mid \sigma \in \Sigma\}$$

$$\mathcal{A}[x_a = a] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) = a\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) \neq a\}$$

$$\mathcal{A}[x_a \neq a] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) \neq a\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) = a\}$$

$$\mathcal{A}[x_a > a] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) > a\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) \leq a\}$$

$$\mathcal{A}[x_a \geq a] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) \geq a\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) < a\}$$

$$\mathcal{A}[x_a < a] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) < a\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) \geq a\}$$

$$\mathcal{A}[x_a \leq a] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) \leq a\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) > a\}$$

$$\mathcal{A}[\text{attribute}_0 \wedge \text{attribute}_1] = \{(\sigma, b_0 \wedge_T b_1) \mid \sigma \in \Sigma$$

$$\text{and } (\sigma, b_0) \in \mathcal{A}[\text{attribute}_0] \text{ and } (\sigma, b_1) \in \mathcal{A}[\text{attribute}_1]\}$$

$$\mathcal{A}[\text{attribute}_0 \vee \text{attribute}_1] = \{(\sigma, b_0 \vee_T b_1) \mid \sigma \in \Sigma$$

**TBexp:**

$$\mathcal{T}[\text{true}] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma\}$$

$$\mathcal{T}[\text{false}] = \{(\sigma, \text{false}) \mid \sigma \in \Sigma\}$$

$$\mathcal{T}[x_t \text{ BEFORE } t] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma \text{ and } \sigma(x_t).n < t.1\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma \text{ and } \sigma(x_t).endn \geq t.start1\}$$

$$\mathcal{T}[x_t \text{ AFTER } t] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma \text{ and } \sigma(x_t).1 > t.n\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma \text{ and } \sigma(x_t).start1 \leq t.endn\}$$

$$\mathcal{T}[x_t \text{ EQUAL } t] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma$$

$$\text{and } (\forall i \in N. \sigma(x_t).si = t.si \text{ and } \sigma(x_t).ei = t.ei)\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma$$

$$\text{and } (\exists i \in N. \sigma(x_t).si \neq t.si \text{ and } \sigma(x_t).ei \neq t.ei)\}$$

$$\mathcal{T}[x_t \text{ OVERLAP } t] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma$$

$$\text{and } (\sigma(x_t).endn \geq t.start1 \text{ and } \sigma(x_t).endn \leq t.start1)$$

$$\text{or } (\sigma(x_t).start1 \geq t.start1 \text{ and } \sigma(x_t).start1 \leq t.endn)\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma \text{ and } \sigma(x_t).n < t.1 \text{ or } \sigma(x_t).1 > t.n\}$$

$$\mathcal{T}[x_t \text{ DURING } t] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma$$

$$\text{and } \sigma(x_t).s1 \geq t.s1 \text{ and } \sigma(x_t).e1 \geq t.en\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma$$

$$\text{and } \sigma(x_t).start1 < t.start1 \text{ and } \sigma(x_t).end1 > t.endn\}$$

$$\mathcal{T}[\text{time}_0 \wedge \text{time}_1] = \{(\sigma, b_0 \wedge_T b_1) \mid \sigma \in \Sigma$$

$$\text{and } (\sigma, b_0) \in \mathcal{A}[\text{time}_0] \text{ and } (\sigma, b_1) \in \mathcal{A}[\text{time}_1]\}$$

$$\mathcal{T}[\text{time}_0 \vee \text{time}_1] = \{(\sigma, b_0 \vee_T b_1) \mid \sigma \in \Sigma$$

$$\text{and } (\sigma, b_0) \in \mathcal{A}[\text{time}_0] \text{ and } (\sigma, b_1) \in \mathcal{A}[\text{time}_1]\}$$

**LBexp:**

$$\mathcal{L}[\text{true}] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma\}$$

$$\mathcal{L}[\text{false}] = \{(\sigma, \text{false}) \mid \sigma \in \Sigma\}$$

$$\mathcal{L}[x_l \text{ EQL}] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma$$

$$\text{and } \forall i \in N. \sigma(x_l).rowi = l.rowi \text{ and } \sigma(x_l).columni = l.columni\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma$$

$$\text{and } \exists i \in N. \sigma(x_l).rowi \neq l.rowi \text{ or } \sigma(x_l).columni \neq l.columni\}$$

$$\mathcal{L}[x_l \text{ OPL}] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma$$

$$\text{and } \exists i, j \in N. \sigma(x_l).rowi = l.rowj \text{ and } \sigma(x_l).columni = l.columnj\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma$$

$$\text{and } \forall i, j \in N. \sigma(x_l).rowi \neq l.rowj \text{ and } \sigma(x_l).columni \neq l.columnj\}$$

$$\mathcal{L}[x_l \text{ IN } l] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma \text{ and } \sigma(x_l) \subset l\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma \text{ and } \sigma(x_l) \not\subset l\}$$

$$\mathcal{L}[x_l \text{ NORTH } l] = \{(\sigma, \text{true}) \mid \sigma \in \Sigma$$

and  $\forall i, j \in N. \sigma(x_i).ri < l.rj$  and  $\sigma(x_j).cj = l.cj$   
 $\cup \{(\sigma, false) \mid \sigma \in \Sigma \text{ and } \exists i, j \in N. \sigma(x_i).ri \geq l.rj\}$   
 $\mathcal{L}[\![location_0 \wedge location_1]\!] = \{(\sigma, b_0 \wedge_T b_1) \mid \sigma \in \Sigma$   
 and  $(\sigma, b_0) \in \mathcal{A}[\![location_0]\!]\}$  and  $(\sigma, b_1) \in \mathcal{A}[\![location_1]\!]\}$   
 $\mathcal{L}[\![location_0 \vee location_1]\!] = \{(\sigma, b_0 \vee_T b_1) \mid \sigma \in \Sigma$   
 and  $(\sigma, b_0) \in \mathcal{A}[\![location_0]\!]\}$  and  $(\sigma, b_1) \in \mathcal{A}[\![location_1]\!]\}$

**DBexp:**

$\mathcal{D}[\![true]\!] = \{(\sigma, true) \mid \sigma \in \Sigma\}$   
 $\mathcal{D}[\![false]\!] = \{(\sigma, false) \mid \sigma \in \Sigma\}$   
 $\mathcal{D}[\![x_d = d]\!] = \{(\sigma, true) \mid \sigma \in \Sigma \text{ and } \sigma(x_d) = d\}$   
 $\cup \{(\sigma, false) \mid \sigma \in \Sigma \text{ and } \sigma(x_d) \neq d\}$   
 $\mathcal{D}[\![x_d \neq d]\!] = \{(\sigma, true) \mid \sigma \in \Sigma \text{ and } \sigma(x_d) \neq d\}$   
 $\cup \{(\sigma, false) \mid \sigma \in \Sigma \text{ and } \sigma(x_d) = d\}$

**EBexp:**

$\mathcal{E}[\![agent^{time}(attribute1; attribute2; attribute3; \dots)]\!]$   
 $= \{(\sigma, b1 \wedge_T b2 \wedge_T b3 \wedge_T \dots) \mid \sigma \in \Sigma$   
 and  $(\sigma, b1) \in \mathcal{T}[\![t]\!]\}$  and  $(\sigma, b2) \in \mathcal{A}[\![a1]\!]$   
 and  $(\sigma, b3) \in \mathcal{A}[\![a2]\!]\}$  and  $(\sigma, b4) \in \mathcal{A}[\![a3]\!]\}$   
 $\mathcal{E}[\![agent_{location}(attribute1; attribute2; attribute3; \dots)]\!]$   
 $= \{(\sigma, b1 \wedge_T b2 \wedge_T b3 \wedge_T \dots) \mid \sigma \in \Sigma$   
 and  $(\sigma, b1) \in \mathcal{L}[\![location]\!]\}$  and  $(\sigma, b2) \in \mathcal{A}[\![a1]\!]$   
 and  $(\sigma, b3) \in \mathcal{A}[\![a2]\!]\}$  and  $(\sigma, b4) \in \mathcal{A}[\![a3]\!]\}$   
 $\mathcal{E}[\![agent^{time}_{location}(attribute1; attribute2; attribute3; \dots)]\!]$   
 $= \{(\sigma, b1 \wedge_T b2 \wedge_T b3 \wedge_T \dots) \mid \sigma \in \Sigma$   
 and  $(\sigma, b1) \in \mathcal{T}[\![t]\!]\}$  and  $(\sigma, b2) \in \mathcal{L}[\![location]\!]\}$   
 and  $(\sigma, b3) \in \mathcal{A}[\![a1]\!]\}$  and  $(\sigma, b4) \in \mathcal{A}[\![a2]\!]\}$  and  $(\sigma, b5) \in \mathcal{A}[\![a3]\!]\}$   
 $\mathcal{E}[\![agent^{time}_{location, direction}(attribute1; attribute2; \dots)]\!]$   
 $= \{(\sigma, b1 \wedge_T b2 \wedge_T b3 \wedge_T \dots) \mid \sigma \in \Sigma$   
 and  $(\sigma, b1) \in \mathcal{T}[\![t]\!]\}$  and  $(\sigma, b2) \in \mathcal{L}[\![l]\!]\}$  and  $(\sigma, b3) \in \mathcal{D}[\![d]\!]\}$   
 and  $(\sigma, b4) \in \mathcal{A}[\![a1]\!]\}$  and  $(\sigma, b5) \in \mathcal{A}[\![a2]\!]\}$  and  $(\sigma, b6) \in \mathcal{A}[\![a3]\!]\}$

**CEBexp:**

$\mathcal{E}[\![e1 \wedge e2]\!] = \{(\sigma, b1 \wedge_T b2) \mid \sigma \in \Sigma$   
 and  $(\sigma, b1) \in \mathcal{E}[\![e1]\!]\}$  and  $(\sigma, b2) \in \mathcal{E}[\![e2]\!]\}$   
 $\mathcal{E}[\![e1 \vee e2]\!] = \{(\sigma, b1 \vee_T b2) \mid \sigma \in \Sigma$   
 and  $(\sigma, b1) \in \mathcal{E}[\![e1]\!]\}$  and  $(\sigma, b2) \in \mathcal{E}[\![e2]\!]\}$

**4 Equivalence between operational semantics and denotational semantics**

The operational semantics of STeCEQL describes the behavioural characteristics of each step. The denotational semantics is more abstract than the operational semantics.

The denotational semantics describes the relationships between the state sets. To illustrate the correctness of the operation semantics, we prove the equivalence between the operational semantics and the denotational semantics of STeCEQL.

**Theorem 1:** For every expression  $attribute \in \mathbf{ABexp}$ , we have  $\mathcal{A}[\![attribute]\!] = \{(\sigma, b) \mid \langle attribute, \sigma \rangle \rightarrow b\}$ .

**Proof.** We prove the theorem by structural induction. We have that

$$P(attribute) \Leftrightarrow_{def}$$

$$\mathcal{A}[\![attribute]\!] = \{(\sigma, b) \mid \langle attribute, \sigma \rangle \rightarrow b\}$$

**The case: attribute=true.**

Let  $(\sigma, b) \in \mathcal{A}[\![true]\!] \Leftrightarrow \sigma \in \Sigma$  and  $b \equiv true$ .

Obviously, if  $(\sigma, b) \in \mathcal{A}[\![true]\!]$ , then  $b \equiv true$  and  $\langle true, \sigma \rangle \rightarrow true$ .

Conversely, if  $\langle true, \sigma \rangle \rightarrow true$ , then the only possible derive is  $b \equiv true$ , thus  $(\sigma, b) \in \mathcal{A}[\![true]\!]$ .

**The case: attribute= $(x_a=a)$ ,  $x_a$  is the storage unit.**

By the definition:

$$\mathcal{A}[\![x_a = a]\!] = \{(\sigma, true) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) = a\}$$

$$\cup \{(\sigma, false) \mid \sigma \in \Sigma \text{ and } \sigma(x_a) \neq a\}.$$

Then  $(\sigma, true) \in \mathcal{A}[\![x_a = a]\!] \Leftrightarrow \sigma \in \Sigma$  and  $\sigma(x_a) = a$ .

If  $(\sigma, true) \in \mathcal{A}[\![x_a = a]\!]$ , then  $\sigma(x_a) = a$ .

By the operational semantics of the expression, we get  $\langle x_a=a, \sigma \rangle \rightarrow true$ .

Conversely, suppose  $\langle x_a=a, \sigma \rangle \rightarrow true$ , then there must be a derivation as below:

$$\frac{\sigma(x_a) = a}{\langle x_a=a, \sigma \rangle \rightarrow true}$$

Thus,  $(\sigma, true) \in \mathcal{A}[\![true]\!]$ .

Hence,  $(\sigma, true) \in \mathcal{A}[\![true]\!] \Leftrightarrow \langle x_a=a, \sigma \rangle \rightarrow true$

Similarly,

$$(\sigma, false) \in \mathcal{A}[\![true]\!] \Leftrightarrow \langle x_a=a, \sigma \rangle \rightarrow false.$$

Thus, we can get:

$$\mathcal{A}[\![x_a=a]\!] \Leftrightarrow \{(\sigma, b) \mid \langle x_a=a, \sigma \rangle \rightarrow b\}$$

**The case: attribute= $(attribute_0 \wedge attribute_1)$ , let  $attribute_0$  and  $attribute_1$  are  $\mathbf{ABexp}$ .**

Suppose  $P(attribute_0)$  and  $P(attribute_1)$  are true.

By the definition:

$$(\sigma, b) \in \mathcal{A}[\![attribute_0 \wedge attribute_1]\!]$$

$$\Leftrightarrow \sigma \in \Sigma \text{ and } \exists b_0, b_1. b = b_0 \wedge_T b_1$$

$$\text{and } (\sigma, b_0) \in \mathcal{A}[\![attribute_0]\!]\} \text{ and } (\sigma, b_1) \in \mathcal{A}[\![attribute_1]\!]\}.$$

Thus, suppose  $(\sigma, b) \in \mathcal{A}[\![attribute_0 \wedge attribute_1]\!]$ , then  $\exists b_0, b_1$   $(\sigma, b_0) \in \mathcal{A}[\![a_0]\!]\}$  and  $(\sigma, b_1) \in \mathcal{A}[\![a_1]\!]\}$

By the suppose, the  $P(attribute_0)$  and  $P(attribute_1)$  are true, then

$\langle \text{attribute}_0, \sigma \rangle \rightarrow b_0$  and  $\langle \text{attribute}_1, \sigma \rangle \rightarrow b_1$   
 Hence, we can derive  
 $\langle \text{attribute}_0 \wedge \text{attribute}_1, \sigma \rangle \rightarrow b, b = b_0 \wedge_T b_1$ .  
 Conversely, every derivation of  
 $\langle \text{attribute}_0 \wedge \text{attribute}_1, \sigma \rangle \rightarrow b$  must have the follows:  

$$\frac{\begin{array}{c} \vdots \\ \langle \text{attribute}_0, \sigma \rangle \rightarrow b_0 \end{array} \quad \frac{\begin{array}{c} \vdots \\ \langle \text{attribute}_1, \sigma \rangle \rightarrow b_1 \end{array}}{\langle \text{attribute}_1, \sigma \rangle \rightarrow b_1}}{\langle \text{attribute}_0 \wedge \text{attribute}_1, \sigma \rangle \rightarrow b}$$

For a  $b_0$  and  $b_1$ , we can derive  $b = b_0 \wedge_T b_1$ .

Because the  $P(\text{attribute}_0)$  and  $P(\text{attribute}_1)$  are true,  
 $(\sigma, b_0) \in \mathcal{A}[\text{attribute}_0]$  and  $(\sigma, b_1) \in \mathcal{A}[\text{attribute}_1]$ .

Hence,  $(\sigma, b) \in \mathcal{A}[\text{attribute}]$ .

The proofs of other cases are completely analogous.  
 We finish the proof of this theorem.

**Theorem 2:** For every expression  $\text{time} \in \mathbf{TBexp}$ , we have

$$\mathcal{T}[\text{time}] = \{(\sigma, b) \mid \langle \text{time}, \sigma \rangle \rightarrow b\}$$

**Proof.** We prove the theorem by structural induction.  
 We have that

$$P(\text{time}) \Leftrightarrow_{\text{def}} \mathcal{T}[\text{time}] = \{(\sigma, b) \mid \langle \text{time}, \sigma \rangle \rightarrow b\}$$

**The case:  $\text{time} \equiv \text{true}$**

Let  $(\sigma, b) \in \mathcal{T}[\text{true}] \Leftrightarrow \sigma \in \Sigma$  and  $b \equiv \text{true}$ .

Obviously, if  $(\sigma, b) \in \mathcal{T}[\text{true}]$  then  
 $b \equiv \text{true}$  and  $\langle \text{true}, \sigma \rangle \rightarrow \text{true}$ .

Conversely, if  $\langle \text{true}, \sigma \rangle \rightarrow \text{true}$ , then the only possible derive is  $b \equiv \text{true}$ , thus  $(\sigma, b) \in \mathcal{T}[\text{true}]$ .

**The case:  $\text{time} \equiv (x_t \text{ BEFORE } t)$ ,  $x_t$  is the storage unit.**

By the definition:

$$\mathcal{T}[x_t \text{ BEFORE } t] = \{(\sigma, \text{true})$$

$$\mid \sigma \in \Sigma \text{ and } \sigma(x_t).endn < t.start1\}$$

$$\cup \{(\sigma, \text{false}) \mid \sigma \in \Sigma \text{ and } \sigma(x_t).endn \geq t.start1\}$$

Then

$$(\sigma, \text{true}) \in \mathcal{T}[x_t \text{ BEFORE } t]$$

$$\Leftrightarrow \sigma \in \Sigma \text{ and } \sigma(x_t).endn < t.start1.$$

If  $(\sigma, \text{true}) \in \mathcal{T}[x_t \text{ BEFORE } t]$ , then  
 $\sigma(x_t).endn < t.start1$ .

By the operational semantics of the expression, we get  
 $\langle x_t \text{ BEFORE } t, \sigma \rangle \rightarrow \text{true}$ .

Conversely, suppose  $\langle x_t \text{ BEFORE } t, \sigma \rangle \rightarrow \text{true}$ ,  
 then there must be a derivation as below:

$$\frac{\sigma(x_t).endn < t.start1}{\langle x_t \text{ BEFORE } t, \sigma \rangle \rightarrow \text{true}}$$

Thus,  $(\sigma, \text{true}) \in \mathcal{T}[x_t \text{ BEFORE } t]$ .

Hence,  $(\sigma, \text{true}) \in \mathcal{T}[x_t \text{ BEFORE } t]$

$$\Leftrightarrow \langle x_t \text{ BEFORE } t, \sigma \rangle \rightarrow \text{true}.$$

Similarly,

$$(\sigma, \text{false}) \in \mathcal{T}[x_t \text{ BEFORE } t] \Leftrightarrow \langle x_t \text{ BEFORE } t, \sigma \rangle \rightarrow \text{false}.$$

Thus we can get:

$$\mathcal{T}[x_t \text{ BEFORE } t] \Leftrightarrow \{(\sigma, b) \mid \langle x_t \text{ BEFORE } t, \sigma \rangle \rightarrow b\}$$

**The case:  $\text{time} \equiv (\text{time}_0 \wedge \text{time}_1)$ , let  $\text{time}_0$  and  $\text{time}_1$  are  $\mathbf{TBexp}$ .**

Suppose  $P(\text{time}_0)$  and  $P(\text{time}_1)$  are true.

By the definition:

$$(\sigma, b) \in \mathcal{T}[\text{time}_0 \wedge \text{time}_1] \Leftrightarrow \sigma \in \Sigma$$

$$\text{and } \exists b_0, b_1. b = b_0 \wedge_T b_1$$

$$\text{and } (\sigma, b_0) \in \mathcal{T}[\text{time}_0] \text{ and } (\sigma, b_1) \in \mathcal{T}[\text{time}_1]$$

Thus, suppose  $(\sigma, b) \in \mathcal{T}[\text{time}_0 \wedge \text{time}_1]$ , then

$$(\sigma, b_0) \in \mathcal{T}[\text{time}_0] \text{ and } (\sigma, b_1) \in \mathcal{T}[\text{time}_1].$$

By the suppose, the  $P(\text{time}_0)$  and  $P(\text{time}_1)$  are true,  
 then

$$\langle \text{time}_0, \sigma \rangle \rightarrow b_0 \text{ and } \langle \text{time}_1, \sigma \rangle \rightarrow b_1$$

Hence, we can

derive  $\langle \text{time}_0 \wedge \text{time}_1, \sigma \rangle \rightarrow b, b = b_0 \wedge_T b_1$ .

Conversely, every derivation of  
 $\langle \text{time}_0 \wedge \text{time}_1, \sigma \rangle \rightarrow b$  must have the follows:

$$\frac{\begin{array}{c} \vdots \\ \langle \text{time}_0, \sigma \rangle \rightarrow b_0 \end{array} \quad \frac{\begin{array}{c} \vdots \\ \langle \text{time}_1, \sigma \rangle \rightarrow b_1 \end{array}}{\langle \text{time}_1, \sigma \rangle \rightarrow b_1}}{\langle \text{time}_0 \wedge \text{time}_1, \sigma \rangle \rightarrow b}$$

For a  $b_0$  and  $b_1$ , we can derive  $b = b_0 \wedge_T b_1$

Because the  $P(\text{time}_0)$  and  $P(\text{time}_1)$  are true,  
 $(\sigma, b_0) \in \mathcal{T}[\text{time}_0]$  and  $(\sigma, b_1) \in \mathcal{T}[\text{time}_1]$ .

Hence,  $(\sigma, b) \in \mathcal{T}[\text{time}]$ .

The proofs of other cases are completely analogous.  
 We finish the proof of this theorem.

**Theorem 3:** For every expression  $\text{location} \in \mathbf{LBexp}$ , we have  $\mathcal{L}[\text{location}] = \{(\sigma, b) \mid \langle \text{location}, \sigma \rangle \rightarrow b\}$

**Proof.** We prove the theorem by structural induction.  
 We have that

$$P(\text{location}) \Leftrightarrow_{\text{def}}$$

$$\mathcal{L}[\text{location}] = \{(\sigma, b) \mid \langle \text{location}, \sigma \rangle \rightarrow b\}$$

**The case:  $\text{location} \equiv \text{true}$ .**

Let  $(\sigma, b) \in \mathcal{L}[\text{true}] \Leftrightarrow \sigma \in \Sigma$  and  $b \equiv \text{true}$

Obviously, if  $(\sigma, b) \in \mathcal{L}[\text{true}]$ , then  
 $b \equiv \text{true}$  and  $\langle \text{true}, \sigma \rangle \rightarrow \text{true}$ .

Conversely, if  $\langle \text{true}, \sigma \rangle \rightarrow \text{true}$ , then the only possible derive is  $b \equiv \text{true}$ , thus  $(\sigma, b) \in \mathcal{L}[\text{true}]$ .

**The case:  $\text{location} \equiv (x_t \text{ EQ } l)$ ,  $x_t$  is the storage unit.**

By the definition:

$$\mathcal{L}[\llbracket x_l EQl \rrbracket] = \{(\sigma, true) \mid \sigma \in \Sigma\}$$

and  $\forall i \in N. \sigma(x_l).row_i = l.row_i$  and  $\sigma(x_l).column_i = l.end_i$

$$\cup \{(\sigma, false) \mid \sigma \in \Sigma\}$$

and  $(\exists i \in N. \sigma(x_l).ri \neq l.ri$  or  $\sigma(x_l).ci \neq l.ci)$

$$\text{Then } (\sigma, true) \in \mathcal{L}[\llbracket x_l EQl \rrbracket] \Leftrightarrow$$

$\sigma \in \Sigma$  and  $\forall i \in N. \sigma(x_l).row_i = l.row_i$  and  $\sigma(x_l).column_i = l.end_i$ .

If  $(\sigma, true) \in \mathcal{L}[\llbracket x_l EQl \rrbracket]$ , then

$\forall i \in N. \sigma(x_l).row_i = l.row_i$  and  $\sigma(x_l).column_i = l.end_i$ .

By the operational semantics of the expression, we get  $\langle x_l EQl, \sigma \rangle \rightarrow true$ .

Conversely, suppose  $\langle x_l EQl, \sigma \rangle \rightarrow true$ , then there must be a derivation as below:

$$\frac{\forall i \in N. \sigma(x_l).row_i = l.row_i \text{ and } \sigma(x_l).column_i = l.end_i}{\langle x_l EQl, \sigma \rangle \rightarrow true}$$

Thus,  $(\sigma, true) \in \mathcal{L}[\llbracket x_l EQl \rrbracket]$ .

Hence,  $(\sigma, true) \in \mathcal{L}[\llbracket x_l EQl \rrbracket] \Leftrightarrow \langle x_l EQl, \sigma \rangle \rightarrow true$ .

Similarly,

$(\sigma, false) \in \mathcal{L}[\llbracket x_l EQl \rrbracket] \Leftrightarrow \langle x_l EQl, \sigma \rangle \rightarrow false$ .

Thus we can get:

$$\mathcal{L}[\llbracket x_l EQl \rrbracket] \Leftrightarrow \{(\sigma, b) \mid \langle x_l EQl, \sigma \rangle \rightarrow b\}$$

The case:  $location \equiv (location_0 \wedge location_1)$ , let  $location_0$  and  $location_1$  are LBexp.

Suppose  $P(location_0)$  and  $P(location_1)$  are true.

By the definition:

$$(\sigma, b) \in \mathcal{L}[\llbracket location_0 \wedge location_1 \rrbracket] \Leftrightarrow \sigma \in \Sigma$$

and  $\exists b_0, b_1. b = b_0 \wedge_T b_1$

and  $(\sigma, b_0) \in \mathcal{L}[\llbracket location_0 \rrbracket]$  and  $(\sigma, b_1) \in \mathcal{L}[\llbracket location_1 \rrbracket]$

By suppose, the  $P(location_0)$  and  $P(location_1)$  are true, then

$(\sigma, b_0) \in \mathcal{L}[\llbracket location_0 \rrbracket]$  and  $(\sigma, b_1) \in \mathcal{L}[\llbracket location_1 \rrbracket]$

Hence, we can derive  $\langle location_0 \wedge location_1, \sigma \rangle \rightarrow b$ ,  $b = b_0 \wedge_T b_1$ .

Conversely, every derivation of  $\langle location_0 \wedge location_1, \sigma \rangle \rightarrow b$  must have the follows:

$$\frac{\begin{array}{c} \vdots \\ \langle location_0, \sigma \rangle \rightarrow b_0 \end{array} \quad \frac{\begin{array}{c} \vdots \\ \langle location_1, \sigma \rangle \rightarrow b_1 \end{array}}{\langle location_0 \wedge location_1, \sigma \rangle \rightarrow b}}$$

For a  $b_0$  and  $b_1$ , we can derive  $b = b_0 \wedge_T b_1$ .

Because the  $P(location_0)$  and  $P(location_1)$  are true,  $(\sigma, b_0) \in \mathcal{L}[\llbracket location_0 \rrbracket]$  and  $(\sigma, b_1) \in \mathcal{L}[\llbracket location_1 \rrbracket]$ .

Hence,  $(\sigma, b) \in \mathcal{L}[\llbracket location \rrbracket]$ .

The proofs of other cases are completely analogous.

We finish the proof of this theorem.

**Theorem 4:** For every expression  $direction \in \mathbf{LBexp}$ , we have  $\mathcal{D}[\llbracket direction \rrbracket] = \{(\sigma, b) \mid \langle direction, \sigma \rangle \rightarrow b\}$

**Proof.** We prove the theorem by structural induction.

We have that  $P(direction) \Leftrightarrow_{def}$

$$\mathcal{D}[\llbracket direction \rrbracket] = \{(\sigma, b) \mid \langle direction, \sigma \rangle \rightarrow b\}$$

**The case:  $direction \equiv true$ .**

Let  $(\sigma, b) \in \mathcal{D}[\llbracket true \rrbracket] \Leftrightarrow \sigma \in \Sigma$  and  $b \equiv true$ .

Obviously, if  $(\sigma, b) \in \mathcal{D}[\llbracket true \rrbracket]$ , then  $b \equiv true$  and  $\langle true, \sigma \rangle \rightarrow true$ .

Conversely, if  $\langle true, \sigma \rangle \rightarrow true$ , then the only possible derive is  $b \equiv true$ , thus  $(\sigma, b) \in \mathcal{D}[\llbracket true \rrbracket]$ .

**The case:  $direction \equiv (x_d = d)$ ,  $x_d$  is the storage unit.**

By the definition:

$$\mathcal{D}[\llbracket x_d = d \rrbracket] = \{(\sigma, true) \mid \sigma \in \Sigma \text{ and } \sigma(x_d) = d\} \cup \{(\sigma, false) \mid \sigma \in \Sigma \text{ and } \sigma(x_d) \neq d\}$$

Then

$$(\sigma, true) \in \mathcal{D}[\llbracket x_d = d \rrbracket] \Leftrightarrow \sigma \in \Sigma \text{ and } \sigma(x_d) = d.$$

If  $(\sigma, true) \in \mathcal{D}[\llbracket x_d = d \rrbracket]$ , then  $\sigma(x_d) = d$ .

By the operational semantics of the expression, we get  $\langle x_d = d, \sigma \rangle \rightarrow true$ .

Conversely, suppose  $\langle x_d = d, \sigma \rangle \rightarrow true$ , then there must be a derivation as below:

$$\frac{\sigma(x_d) = d}{\langle x_d = d, \sigma \rangle \rightarrow true}$$

Thus,  $(\sigma, true) \in \mathcal{D}[\llbracket true \rrbracket]$ .

Hence,  $(\sigma, true) \in \mathcal{D}[\llbracket true \rrbracket] \Leftrightarrow \langle x_d = d, \sigma \rangle \rightarrow true$ .

Similarly,

$(\sigma, false) \in \mathcal{D}[\llbracket true \rrbracket] \Leftrightarrow \langle x_d = d, \sigma \rangle \rightarrow false$ .

Thus, we can get:

$$\mathcal{D}[\llbracket x_d = d \rrbracket] \Leftrightarrow \{(\sigma, b) \mid \langle x_d = d, \sigma \rangle \rightarrow b\}$$

The proofs of other cases are completely analogous.

We finish the proof of this theorem.

**Theorem 5:** For every expression  $e \in \mathbf{EBexp}$ , we have  $\mathcal{E}[\llbracket e \rrbracket] = \{(\sigma, b) \mid \langle e, \sigma \rangle \rightarrow b\}$

**Proof.** We prove the theorem by structural induction.

We have that  $P(e) \Leftrightarrow_{def} \mathcal{E}[\llbracket e \rrbracket] = \{(\sigma, b) \mid \langle e, \sigma \rangle \rightarrow b\}$ .

**The**

**case:**

$e \equiv agent^{time}(attribute_1; attribute_2; attribute_3)$ , let **time** is **EBexp**, **attributes** are **ABexp**,

Suppose  $P(\mathbf{time})$  and  $P(\mathbf{attribute})$ s are true.

By the definition:

$$(\sigma, b) \in \mathcal{E}[\llbracket agent^{time}(attribute_1; attribute_2; attribute_3) \rrbracket]$$

$$\Leftrightarrow \sigma \in \Sigma \text{ and } \exists b_1, b_2, b_3, b_4. b = b_1 \wedge_T b_2 \wedge_T b_3 \wedge_T b_4$$

and  $(\sigma, b_1) \in \mathcal{T}[\![time]\!]$  and  $(\sigma, b_2) \in \mathcal{A}[\![attribute_1]\!]$

and  $(\sigma, b_3) \in \mathcal{A}[\![attribute_2]\!]$  and  $(\sigma, b_4) \in \mathcal{A}[\![attribute_3]\!]$

Thus, suppose  $(\sigma, b) \in \mathcal{E}[\![agent^{time}(attribute_1; attribute_2; attribute_3)]\!]$ , then  $\exists b_1, b_2, b_3, b_4, (\sigma, b_1) \in \mathcal{T}[\![time]\!]$  and  $(\sigma, b_2) \in \mathcal{A}[\![a1]\!]$  and  $(\sigma, b_3) \in \mathcal{A}[\![a2]\!]$  and  $(\sigma, b_4) \in \mathcal{A}[\![a3]\!]$ .

By the suppose  $P(\mathbf{time})$  and  $P(\mathbf{attribute})$ s are true, then  $\langle time, \sigma \rangle \rightarrow b_1$  and  $\langle attribute_1, \sigma \rangle \rightarrow b_2$  and  $\langle attribute_2, \sigma \rangle \rightarrow b_3$  and  $\langle attribute_3, \sigma \rangle \rightarrow b_4$ .

Hence, we can derive  $\langle agent^{time}(attribute_1; attribute_2; attribute_3), \sigma \rangle \rightarrow b$ ,  $b = b_1 \wedge_T b_2 \wedge_T b_3 \wedge_T b_4$ .

Conversely, every derivation of  $\langle agent^{time}(attribute_1; attribute_2; attribute_3), \sigma \rangle \rightarrow b$  must have the follows:

$$\frac{\begin{array}{c} \vdots \\ \langle t, \sigma \rangle \rightarrow b_1 \end{array} \quad \frac{\begin{array}{c} \vdots \\ \langle a_1, \sigma \rangle \rightarrow b_2 \end{array} \quad \frac{\begin{array}{c} \vdots \\ \langle a_2, \sigma \rangle \rightarrow b_3 \end{array}}{\langle agent_t(a_1, a_2, a_3), \sigma \rangle \rightarrow b}}$$

For a **time** and **attributes**, we can derive  $b = b_1 \wedge_T b_2 \wedge_T b_3 \wedge_T b_4$ .

Because the  $P(\mathbf{time})$  and  $P(\mathbf{attribute})$  are true,  $(\sigma, b_1) \in \mathcal{T}[\![time]\!]$  and  $(\sigma, b_2) \in \mathcal{A}[\![attribute_1]\!]$  and  $(\sigma, b_3) \in \mathcal{A}[\![attribute_2]\!]$  and  $(\sigma, b_4) \in \mathcal{A}[\![attribute_3]\!]$ .

Hence,  $(\sigma, b) \in \mathcal{E}[\![agent^{time}(attribute_1; attribute_2; attribute_3)]\!]$ .

The proofs of other cases are completely analogous.

We finish the proof of this theorem.

**Theorem 6:** For every expression  $ce \in \mathbf{CEBexp}$ , we have  $\mathcal{E}[\![ce]\!] = \{(\sigma, b) \mid \langle ce, \sigma \rangle \rightarrow b\}$

**Proof.** We prove the theorem by structural induction. We have that

$$P(e) \Leftrightarrow_{def} \mathcal{E}[\![ce]\!] = \{(\sigma, b) \mid \langle ce, \sigma \rangle \rightarrow b\}.$$

**The case:**  $ce \equiv (e_1 \wedge e_2)$ , let  $e_1$  and  $e_2$  are **EBexp**,

Suppose  $P(e_1)$  and  $P(e_2)$  are true.

By the definition:  $(\sigma, b) \in \mathcal{E}[\![e_1 \wedge e_2]\!] \Leftrightarrow \sigma \in \Sigma$  and  $\exists b_0, b_1. b = b_0 \wedge_T b_1$  and  $(\sigma, b_0) \in \mathcal{E}[\![e_1]\!]$  and  $(\sigma, b_1) \in \mathcal{E}[\![e_2]\!]$ .

Thus, suppose  $(\sigma, b) \in \mathcal{E}[\![e_1 \wedge e_2]\!]$ , then  $\exists b_0, b_1, (\sigma, b_0) \in \mathcal{E}[\![e_1]\!]$  and  $(\sigma, b_1) \in \mathcal{E}[\![e_2]\!]$ .

By the suppose  $P(e_1)$  and  $P(e_2)$  are true, then  $\langle e_1, \sigma \rangle \rightarrow b_0$  and  $\langle e_2, \sigma \rangle \rightarrow b_1$ .

Hence, we can derive  $\langle e_1 \wedge e_2, \sigma \rangle \rightarrow b, b = b_0 \wedge_T b_1$ .

Conversely, every derivation of  $\langle e_1 \wedge e_2, \sigma \rangle \rightarrow b$  must have the follows:

$$\frac{\begin{array}{c} \vdots \\ \langle e_1, \sigma \rangle \rightarrow b_0 \end{array} \quad \frac{\begin{array}{c} \vdots \\ \langle e_2, \sigma \rangle \rightarrow b_1 \end{array}}{\langle e_1 \wedge e_2, \sigma \rangle \rightarrow b}$$

For a  $e_1$  and  $e_2$ , we can derive  $b = b_0 \wedge_T b_1$ .

Because the  $P(e_1)$  and  $P(e_2)$  are true,  $(\sigma, b_0) \in \mathcal{E}[\![e_1]\!]$  and  $(\sigma, b_1) \in \mathcal{E}[\![e_2]\!]$

Hence,  $(\sigma, b) \in \mathcal{E}[\![ce]\!]$ .

The proofs of other cases are completely analogous.

We finish the proof of this theorem.

Up to date, we have finished the proof of equivalence between the operational semantics and the denotational semantics of STeCEQL.

## 5 Conclusion and outlook

In this paper, focusing on the correctness of the operational semantics of the EQL STeCEQL, we give the denotational semantics of it and prove the equivalence of two semantics of STeCEQL by structural inductive method. From the view of formal semantics of computer language, the equivalence of the operational semantics and the denotational semantics show the correctness of its operational semantics.

Since the internet of vehicles is a typical real-time distributed mobile networked system, we will study the processing algorithm of the STeCEQL in next steps

## Acknowledgments

This work is supported by the National Basic Research Program of China (No. 2011CB302802), the National Natural Science Foundation of China (No.61370100 and No.61021004) and Shanghai Knowledge Service Platform Project (No. ZF1213).

## References

- [1] Atzori L, Iera A, Morabito G 2010 The internet of things: A survey *Computer networks* 54(15) 2787-805
- [2] Chen Y 2012 Stec: A location-triggered specification language for real-time systems *Object/Component/Service-Oriented Real-Time Distributed Computing Workshops (ISORCW)*, 15th IEEE International Symposium on, IEEE pp 1-6
- [3] Wu H, Chen Y, Zhang M 2013 On Denotational Semantics of Spatial-Temporal Consistency Language-SteC *Theoretical Aspects of Software Engineering (TASE)*, 2013 International Symposium on, IEEE 113-20
- [4] Schwiderski-Grosche S, Moody K 2009 The SpaTeC composite event language for spatio-temporal reasoning in mobile systems *Proceedings of the Third ACM International Conference on Distributed Event-Based Systems*, ACM p 11
- [5] Moody K, Bacon J, Evans D 2010 *Implementing a practical spatio-temporal composite event language, From active data management to event-based systems and more* Springer Berlin Heidelberg, pp 108-23

[6] Jin B, Zhuo W, Hu J 2013 Specifying and detecting spatio-temporal events in the internet of things *Decision Support Systems* 55(1) 256-69

[7] Zhu D, Sethi A S 2001 SEL, a new event pattern specification language for event correlation *Computer Communications and Networks, Proceedings. Tenth International Conference on IEEE* p 586-9

[8] Seiriö M, Berndtsson M 2005 *Design and implementation of an ECA rule markup language, Rules and rule markup languages for the semantic web* Springer Berlin Heidelberg, p 98-112

[9] Wu E, Diao Y, Rizvi S 2006 High-performance complex event processing over streams *Proceedings of the 2006 ACM SIGMOD international conference on Management of data. ACM* p 407-18

[10] Demers A J, Gehrke J, Panda B 2007 Cayuga: A General Purpose Event Monitoring System *CIDR* p 412-22

[11] Eckert M 2008 *Complex event processing with XChange EQ: language design, formal semantics, and incremental evaluation for querying events* LMU München: Faculty of Mathematics, München

[12] Anicic D, Rudolph S, Fodor P 2012 Stream reasoning and complex event processing in etalis *Semantic Web* 3(4) 397-407

Authors



**Huiyong Li, born on February 2, 1980, Taiyuan, China**

**Current position, grades:** PhD student of software engineering institute, East China Normal University.  
**University studies:** M.Sc. in Computer Sciences (2011) from Taiyuan University of Science and Technology.  
**Scientific interests:** different aspects of Internet of Things and Real-time Distributed Systems.



**Yixiang Chen, born on March 12, 1961, Xuzhou, China**

**Current position, grades:** Chair of the MoE Key Research Center for Software/Hardware Co-design Engineering, East China Normal University.  
**University studies:** PhD in Mathematics (1995) from Sichuan University.  
**Scientific interests:** different aspects of Real-time System, Formal Models, Formal Semantics of Programming, Foundation of Computations and Trustworthy Network.

# Analysis and research on the simulation and output of discrete event system with fuzzy parameters

Cuirong Zhao<sup>1\*</sup>, Honghai Wang<sup>2</sup>

<sup>1</sup>Anhui Wonder University of Information Engineering, Hefei, 231201, China

<sup>2</sup>The Department of Information and Communication Technology of Anhui Sanlian University, Hefei, 230601, China

Received 12 June 2014, www.tsi.lv

## Abstract

This paper discusses discrete event system simulation output analysis method with fuzzy input parameter. For a classic discrete event model, which contains randomness, once stimulation running is only a sampling according to systematic behaviour, which could not represent all features of the system. Hence, there should be a systematic analysis method, under the guidance of which to apply multi-times stimulation of model and analyse output data of stimulation. This paper provides a solution and introduces random fuzzy theory at last to improve traditional output analysis method. Result of stimulation experiment proves that the method could improve the reliability of stimulation output analysis.

*Keywords:* Discrete event system, Fuzzy Parameters

## 1 Introduction

This paper mainly contains two sections: single system output analysis method based on fuzzy discrete event stimulation and evaluation of influence of input parameter fuzziness on stimulation output result.

Section One analyses problems of stimulation output analysis method in discrete event system with fuzzy parameters, provides solutions and improves traditional output analysis method by introducing random fuzzy theory at last. Result of stimulation experiment proves that the method could improve reliability of stimulation output analysis.

Section Two provides influence index of evaluating fuzziness of input parameter on stimulation output, including absolute index and relative index. Absolute index could perform single evaluation of influence of some input parameter on stimulation output and relative index could compare the influence of each parameter on output. At last application of the method is illustrated through stimulation experiment.

## 2 Numerical characteristics of random fuzzy variable

Among computer algorithm, it always requires numerical characteristics such as mathematical expectation or variance of some random fuzzy variable. For random variable, it could sample in probability space, use method of statistical to evaluate corresponding numerical characteristics [1]. However, because law of large numbers does not exist in possible space, method of statistical could not be used to evaluate mathematical expectation or variance of some random fuzzy variable.

Hence, corresponding computer stimulation algorithm should be designed based on definition of each numerical characteristic [2, 3].

**Method 1:** Calculate expectation of random fuzzy variable  $\xi$  in possible space  $(\Theta, P(\Theta), Pos)$ .

- 1) Select  $N$   $\theta_k (k = 1, 2, \dots, N)$ , which satisfy  $Pos\{\theta_k\} > 0$  from  $\Theta$ ;
- 2) Make  $i = 1$ ;
- 3) Evaluate expectation  $E[\xi(\theta_i)]$  of random variable  $\xi(\theta_i)$ ;
- 4) If  $i < N$ , then  $i = i + 1$ , return 3);
- 5) Make,  $b = \max_{1 \leq i \leq N} E[\xi(\theta_i)]$ ;
- 6) Randomly generate  $M$  number  $r_k (k = 1, 2, \dots, M)$  from section  $[a, b]$ ;
- 7) Make  $sum = 0, j = 1$ ;
- 8) If  $r_j \geq 0$ , then  $sum = sum + Cr\{\theta \in \Theta | E[\xi(\theta)] \geq r_j\}$ , or  $sum = sum - Cr\{\theta \in \Theta | E[\xi(\theta)] \geq r_j\}$ ;
- 9) If  $j < M$ , then  $j = j + 1$ , return 8);
- 10) Calculate  $E[\xi] = a \vee 0 + b \wedge 0 + sum(b - a) / M$ .

Algorithm to calculate variance is to calculate expectation of  $\xi^2$  and  $\xi$  separately, then calculate  $V[\xi]$ .

**Method 2:** Calculate variance of random fuzzy variable  $\xi$ .

- 1) Use method 1 to calculate  $E[\xi^2]$  of random fuzzy variable  $\xi^2$ ;
- 2) Use method 1 to calculate  $E[\xi]$  of random fuzzy variable  $\xi$ ;

\* *Corresponding author* e-mail: zhaocuirong823@163.com

3) Calculate  $V[\xi] = E[\xi^2] - (E[\xi])^2$ .

**3 Stimulation data analysis of fuzzy discrete event**

**3.1 INFLUENCE OF INPUTTING INCORRECT PARAMETER ON STIMULATION OUTPUT ANALYSIS**

In order to explain the influence of inputting incorrect parameter on stimulation output analysis, next experiment is designed to illustrate this issue.

Let's consider an M/G/1queue model, in which process of customer coming is a simple Poisson process, process parameter  $\lambda = 1/15$  service table is independent identically distribution for each customer service, mean value is p time unit, variance is 1; performance index which users care is average waiting time  $W_q$  of each customer as system is stable.

Now, suppose that true value of p is 10, but stimulation modelling analyst does not know this value; and suppose that it is not allowed to take samples of service time in large quantities to evaluate value of p. There could only be 5 samples at most and mean value of 5 samples is value of pin the model.

Next, use true value p=10 and sampling evaluation value as input parameter separately to run 100 times independent stimulation and calculate confidence interval with confidence level of 90% [4, 5]:

$$\begin{aligned} \bar{W}_q(n) \pm t_{n-1, 1-\alpha/2} \sqrt{\frac{S_n^2(W_q)}{n}} \\ \bar{W}_q(100) \pm 0.166 \sqrt{\frac{S_{100}^2(W_q)}{100}} \end{aligned} \quad (1)$$

Then verify if this interval contains true value of  $W_q$ :

$$W_q = \frac{\lambda(1 - \mu^2 \sigma_s^2)}{2\mu(\mu - \lambda)} = \frac{1}{15} \left( 1 - \frac{1}{10^2} \right) = 9.9. \quad (2)$$

In order to verify if confidence interval built could reach confidence level of 90% which is required, 100 times independent experiments are applied for both conditions separately. Statistical stimulation result is "correct", i.e. confidence interval contains times and proportion of true value of 9.9. Experiment result is illustrated in table 1.

It can be told, that using inaccurate information to build system model causes severe deterioration of result of classic stimulation output analysis. And it lowers meaning of system stimulation. Hence, if model input parameter is inaccurate, it must be handled to improve method of stimulation output analysis and keep the result of stimulation reliable.

TABLE 1 Comparison of 90% confidence interval built by using accurate parameter and inaccurate parameter

Experiment order	Using accurate parameter		Using inaccurate parameter		
	Confidence interval	Contains true value	Parameter P value	Confidence interval	Contains true value
1	[9.6161,10.090]	Y	9.4337	[7.6949,8.1134]	N
2	[9.5746,10.146]	Y	10.569	[12.093,12.889]	N
3	[9.7620,10.291]	Y	10.400	[11.284,12.007]	N
4	[9.8680,10.339]	Y	10.062	[9.8531,10.451]	Y
5	[9.7634,10.209]	Y	10.121	[10.086,10.703]	N
6	[9.7908,10.282]	Y	9.9297	[9.3474,9.9017]	Y
7	[9.8805,10.408]	Y	10.622	[12.359,13.179]	N
8	[10.038,10.598]	N	10.096	[9.9864,10.595]	N
9	[9.7846,10.341]	Y	9.6414	[8.3448,8.8155]	N
10	[9.7657,10.274]	Y	9.5193	[7.9559,8.3952]	N
11	[9.7177,10.272]	Y	10.479	[11.656,12.412]	N
12	[9.7611,10.225]	Y	9.7320	[8.6466,9.1421]	N
.....	.....	.....	.....	.....	.....
96	[9.7648,10.309]	Y	9.8510	[9.0611,9.5912]	N
97	[9.9056,10.444]	N	9.3565	[7.4674,7.8681]	N
98	[9.7687,10.290]	Y	9.2652	[7.2071,7.5876]	N
99	[9.8506,10.306]	Y	9.7417	[8.6796,9.1779]	N
100	[9.6661,10.201]	Y	9.7611	[8.7460,9.2498]	N
Correct times (proportion)			89(99%)	12(12%)	

**3.2 TWO METHODS OF SOLVING DETERIORATION OF OUTPUT ANALYSIS RESULT**

Aim of supposed stimulation is to evaluate expectation  $E[y]$  of some output variable y, then run n times independent stimulation, using sampling mean value

$\bar{y}(n)$  to evaluate  $E[y]$  and build confidence interval of confidence level  $(1 - \alpha)$  as follow [6]:

$$\bar{y}(n) \pm t_{n-1, 1-\alpha/2} \sqrt{\frac{S_n^2(y)}{n}} \quad (3)$$

According to central-limit theorem, several mean values of random variables of IID are similar to normal distribution. Hence, distribution of  $\bar{y}(n)$  could be recognized as normal distribution and use  $E[y]$  as expectation. When input parameter of system model is inaccurate, expectation of  $\bar{y}(n)$  could deviate from  $E[y]$ . See figure 1.

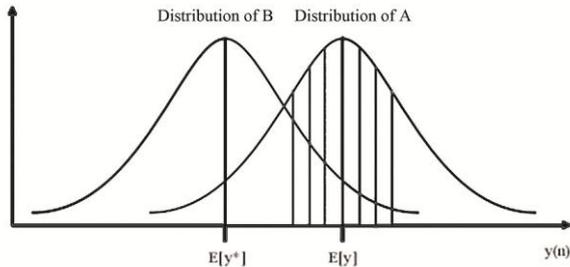


FIGURE 1 Probability density function of  $\bar{y}(n)$

Among those, distribution A is probability distribution of  $\bar{y}(n)$  when input parameter is accurate. If

$$\bar{y}(n) \text{ is in section of } E[y] \pm t_{n-1,1-\alpha/2} \sqrt{\frac{S_n^2(y)}{n}}, \text{ i.e.}$$

section under the shadow in figure 1, then confidence interval in (3) will contain  $E[y]$ . At this moment, probability of event “confidence interval contains true value” is area of shadow section in figure 1, set as  $(1-\alpha)$ .

When model input parameter is inaccurate, output variable  $y$  may deviate. For example, expectation of  $y$  is  $E[y^*]$ , and sample mean value  $\bar{y}(n)$  submits to distribution B in figure 1. In this condition, probability of

$$\bar{y}(n) \text{ falling into section } E[y] \pm t_{n-1,1-\alpha/2} \sqrt{\frac{S_n^2(y)}{n}} \text{ will}$$

decrease (to the area of curve of distribution B in the shadow), equally, probability of confidence interval

$$\bar{y}(n) \pm t_{n-1,1-\alpha/2} \sqrt{\frac{S_n^2(y)}{n}} \text{ containing } E[y] \text{ will decrease}$$

too which is lower than set up value  $(1-\alpha)$  [7, 8].

It can be told that, when input parameter is inaccurate, reliability of stimulation output analysis conclusion will lower. Hence, this paper shall solve this issue from two aspects: adjusting mid-point of confidence interval and enlarging half width of confidence interval.

### 3.2.1 Adjusting mid-point of confidence interval

Firstly, it will explain the first method. In order to make it simple, let us think about the condition that in system model, it only contains one inaccurate input parameter and one output performance index. When model has more than one inaccurate parameter, concept of solving the issue is similar; when several output performance index

need to be evaluated, this method could be applied to each index [9].

Suppose inaccurate parameter in model is  $p$ , output performance index is  $E[y]$ , and suppose true value of  $p$  is  $p_0$ , corresponding true value of  $E[y]$  is  $E[y_0]$ . Now suppose stimulation modelling staff and analyst could only get one inaccurate value  $p^*$  of  $p$ . When input parameter  $p^*$ , value of output performance index is  $E[y^*]$ . Normally, it could not be told true value  $p_0$  is bigger or smaller than mastered value  $p^*$  now, hence, it could not find out if  $E[y_0]$  is bigger than  $E[y^*]$ , or not. In fact, as information is insufficient, it could not make mid-point of confidence interval move along the direction to  $E[y_0]$ . Hence, the aim could not be achieved through the first method.

This paper applies another method of adjusting on this issue: through adjustment, it could make mid-point of confidence interval move along some direction (positive direction or negative direction) to decrease maximum error (i.e. confidence interval could not contain true value) probability. Idea of this method is explained as follow.

Suppose true value  $p_0$  of input parameter is different with  $p^*$  obtained at present, when  $p_0$  is separately smaller or bigger than an equality of  $p^*$  (i.e.  $p_0 = p^* - \Delta$  or  $p_0 = p^* + \Delta, \Delta > 0$ ), there are 3 conditions of difference between  $E[y_0]$  and  $E[y^*]$ .

When  $p_0$  is bigger than  $p^*$ , difference between  $E[y_0]$  and  $E[y^*]$  is huge, i.e.:

$$|E[y(p^* - \Delta)] - E[y(p^*)]| < |E[y(p^* + \Delta)] - E[y(p^*)]|. \quad (4)$$

When  $p_0$  is bigger or smaller than a equality of  $p^*$ , there is no difference between  $E[y_0]$  and  $E[y^*]$ , i.e.:

$$|E[y(p^* - \Delta)] - E[y(p^*)]| = |E[y(p^* + \Delta)] - E[y(p^*)]|. \quad (5)$$

When  $p_0$  is smaller than  $p^*$ , difference between  $E[y_0]$  and  $E[y^*]$  is huge, i.e.:

$$|E[y(p^* - \Delta)] - E[y(p^*)]| > |E[y(p^* + \Delta)] - E[y(p^*)]|. \quad (6)$$

See three conditions above in (a) (b) and (c). In figure 2, suppose that there is a positive correlation between  $E[y]$  and  $p$  (i.e.  $E[y]$  increases as  $p$  increases), but it does not influence effect of the method.

For condition 1, it should try to make mid-point of confidence interval move to positive direction to avoid when  $p_0$  is bigger than  $p^*$ , evaluated distance value of  $E[y_0]$  is too far away, which is out of confidence

interval; similarly, for condition 3, mid-point of confidence interval should move to the negative direction; for condition 2, it does not need to move.

Figure 3 takes condition 1 as example to explain the influence of moving mid-point of confidence interval on maintaining an acceptable confidence level.

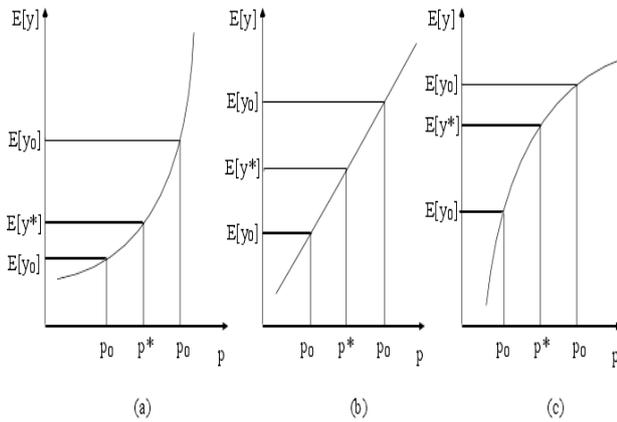


FIGURE 2 Response of  $E[y]$  to  $p$

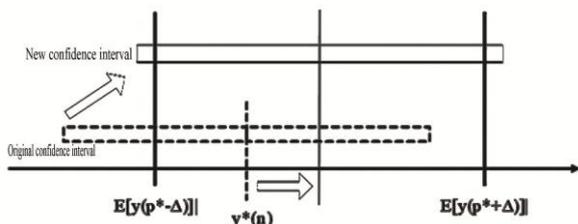


FIGURE 3 Influence of adjusting mid-point of confidence interval

It can be told from figure 3 when  $p_0$  is bigger than  $p^*$ , original confidence interval could miss true value  $E[y_0]$  with a big probability. As mid-point of confidence interval moves to positive direction, the probability shall decrease. Because in the condition 1,  $E[y(p^*-\Delta)]$  is closer to mid-point of confidence interval than to  $E[y(p^*+\Delta)]$ . As illustrated in figure 3, new confidence interval after moving confidence interval still contains  $E[y(p^*-\Delta)]$  [10, 11].

### 3.2.2 Enlarging half width of confidence interval

Second method of increasing confidence level is to enlarging half width of confidence interval. Obviously, this method could improve confidence level for sure.

Tough people prefer a smaller (which means accurate) confidence interval rather than a bigger section scope. But when there is not enough information, it is a safe choice for selecting a bigger confidence interval. It is more reasonable than determining a very thin confidence interval according to unreliable information.

### 3.3 USING RANDOM FUZZY THEORY TO SOLVE PROBLEM OF CLASSIC ANALYSIS METHOD

In order to achieve two improving methods above, this paper introduces fuzzy variable to input parameter, uses method of fuzzy discrete event system stimulation, and uses evaluation of expectation and variance of output random fuzzy variable to replace evaluation of expectation and variance of output random variable in classic method.

First of all, using mathematical expectation evaluation of output random fuzzy variable could adjust mid-point of confidence interval according to required direction. Suppose that fuzzy input parameter  $p$  has symmetrical triangle subordinating degree function, the centre of which does not locate at parameter value  $p^*$ , then subordinating degree function shape of system output performance index  $E[y(p)]$  (which is fuzzy variable at here) under three conditions illustrated in figure 2 are separately illustrated in (a) (b) and (c) in figure 4.

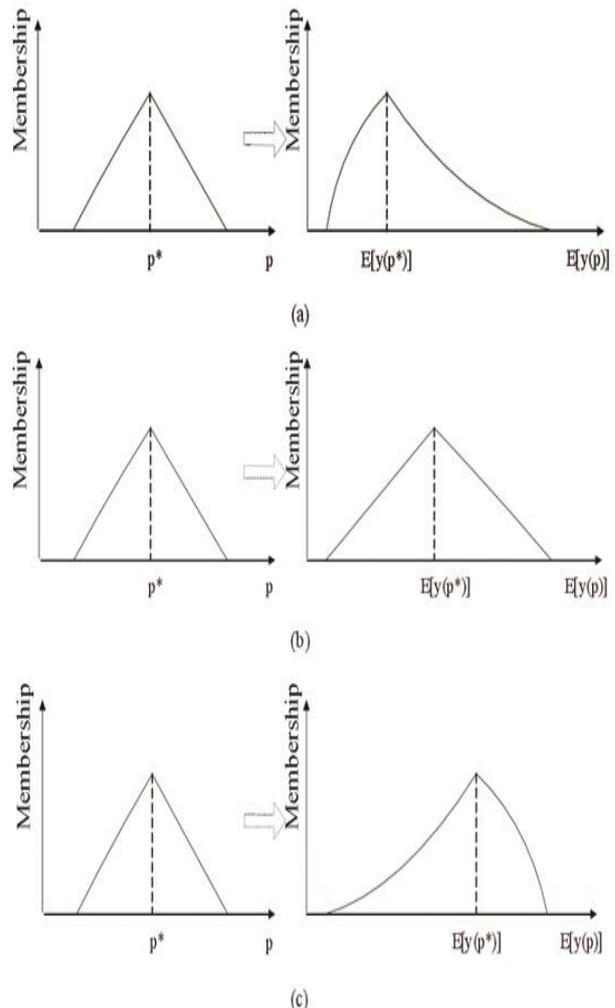


FIGURE 4 Subordinating degree functions of input parameter and output index

In figure 4 (a), expectation of  $E[y(p)]$  will be bigger than  $E[y(p^*)]$ , because right side of  $E[y(p^*)]$  has more possible value. Similarly, in figure 4 (c), expectation of  $E[y(p^*)]$  will be smaller than  $E[y(p^*)]$ .

But in figure 4 (b),  $E[y(p)]$  will be equal to  $E[y(p^*)]$ . So centre of confidence interval shall be moved to mathematical expectation estimator position of random fuzzy variable  $y$ .

Secondly, estimator of variance of random fuzzy variable  $y$  is used to calculate half width of confidence interval, because  $y$  has two uncertainties which are randomness and fuzziness at the same time. Increasing of uncertainty could be expressed in variance of  $y$ .

Influence evaluation of fuzzy input parameter on output result.

4.1 ABSOLUTE INDEX OF INFLUENCE OF FUZZY INPUT PARAMETER ON OUTPUT RESULT

Influence evaluation index here is initially applied through subordinating degrees, and they form a general influence index. Suppose that system has  $n$  fuzzy input parameters:  $p_1, p_2, \dots, p_n$ , output performance index is  $\tilde{q}$ . Detailed method is as follow:

- (1) Divide subordinating degree into several degrees, such as  $\alpha_1, \alpha_2, \dots, \alpha_k$ ;
- (2) Make  $j=1$ ;
- (3) On subordinating degree  $\alpha_j$ ,  $\alpha$  level set of each parameter  $p_i (i=1, 2, \dots, n)$  is calculated based on subordinating degree of all input parameters:

$$(p_i)_{\alpha_j} \sim I_i^j = [a_i^j, b_i^j], i=1, 2, \dots, n. \tag{7}$$

- (4) Perform discrete event stimulation on each parameter combination of subordinating degree  $\alpha_j$ , calculate value of performance index  $\tilde{q}$ . Make  $(p_1, p_2, \dots, p_n) = (x_1, x_2, \dots, x_n)$  in parameter combination, and mark the performance index calculated by stimulation as  $\hat{q}(x_1, x_2, \dots, x_n)$ .
- (5) Separately calculate absolute influence index of each parameter  $p_i (i=1, 2, \dots, n)$  on subordinating degree  $\alpha_j$ :

$$\lambda_i^j = \frac{\sum_{x_i=b_i^j} \hat{q}(x_1, x_2, \dots, x_n) - \sum_{x_i=a_i^j} \hat{q}(x_1, x_2, \dots, x_n)}{2^{n-1} (b_i^j - a_i^j)}, (i=1, 2, \dots, n). \tag{8}$$

- (6) If  $j < k$ , then  $j=j+1$ , return (3);
- (7) Calculate influence of fuzziness of each parameter  $p_i (i=1, 2, \dots, n)$  on output performance index:

$$\lambda_i^j = \frac{\sum_{j=1}^k \alpha_j \lambda_i^j}{\sum_{j=1}^k \alpha_j}, (i=1, 2, \dots, n). \tag{9}$$

4.2 RELATIVE INDEX OF INFLUENCE OF FUZZY INPUT PARAMETER ON OUTPUT RESULT

Absolute index of influence of fuzzy input parameter on output result is stated above, influence index of each parameter may has different dimension, hence, influence of each parameter could not be compared. Next, absolute index obtained based on method above shall be handled to get relative index, which could influence evaluation to meet requirement of comparison of parameters.

Mark variance of each fuzzy input parameter as  $V(p_i) (i=1, 2, \dots, n)$ , then relative index of influence of each parameter on output result is defined as follow:

$$\omega_i = \frac{|\lambda_i \sqrt{V(p_i)}|}{\sum_{l=1}^n |\lambda_l \sqrt{V(p_l)}|}, (i=1, 2, \dots, n). \tag{10}$$

After handling like that,  $\omega_i (i=1, 2, \dots, n)$  are all dimensionless indexes, and satisfy:

$$\sum_{i=1}^n \omega_i = 1. \tag{11}$$

At this time, compare each input parameter to determine fuzziness of which parameters have more influence on performance index.

5 Stimulation experiment

In order to testify effectiveness of method above, a stimulation experiment is designed. As for a system model, it is stimulated and analysed separately by classic discrete event system stimulation and output analysis method, and fuzzy discrete event system stimulation and output analysis method provided in this paper. And results of two methods are compared.

5.1 DESCRIPTION OF MODEL

For model of 3.1 which is still being used, according to classic discrete event stimulation method, use  $\bar{p}$  (5) (mean value of 5 samples) as value of parameter  $p$ ; and for fuzzy discrete event model, use  $\bar{p}$  (5) as centre value of fuzzy input parameter  $p$  (use symmetrical triangle subordinating degree function), and use 2 times of

standard deviation of 5 samples as half width of support set of  $p$ , i.e.  $p = [\bar{p}(5) - 2\sigma_p, \bar{p}(5), \bar{p}(5) + 2\sigma_p]$ .

Confidence intervals of 90% of  $W_q$  are built by two methods separately and times and proportion of confidence interval built by two methods, which contains true value 9.9 of  $W_q$  are compared.

5.2 STIMULATION OPERATION

For model above, perform stimulation separately by classic discrete event stimulation method and fuzzy discrete event stimulation method provided in this paper and build confidence interval of 90% by corresponding output analysis method separately. In the output analysis method taken by this paper, expectation of random fuzzy variable should be calculated.

In this algorithm, for model in this experiment, reliability measuring method is as follow: suppose  $E[\xi(\theta)]$  could reach minimum value a, and maximum value b, point which has a possibility of 1 is c, then if  $r < c$ :

$$\begin{aligned} Cr\{\theta \in \Theta [E[\xi(\theta)] \geq r]\} &= \frac{1}{2} (Pos\{E[\xi(\theta) \geq r]\} + Nec\{E[\xi(\theta)] \geq r\}) \\ &= \frac{1}{2} (Pos\{E[\xi(\theta) \geq r]\} + 1 - Pos\{E[\xi(\theta)] < r\}) \quad (12) \\ &= \frac{1}{2} (1 + 1 - Pos\{E[\xi(\theta)] = r\}) \\ &= 1 - \frac{1}{2} Pos\{E[\xi(\theta)] = r\} \end{aligned}$$

But if  $r \geq c$ , then:

$$\begin{aligned} Cr\{\theta \in \Theta [E[\xi(\theta)] \geq r]\} &= \frac{1}{2} (Pos\{E[\xi(\theta) \geq r]\} + Nec\{E[\xi(\theta)] \geq r\}) \\ &= \frac{1}{2} (Pos\{E[\xi(\theta) \geq r]\} + 1 - Pos\{E[\xi(\theta)] < r\}) \quad (13) \\ &= \frac{1}{2} (Pos\{E[\xi(\theta)] = r\} + 1 - 1) \\ &= \frac{1}{2} Pos\{E[\xi(\theta)] = r\} \end{aligned}$$

Use two methods to independently perform the experiment for 100 times, and separately make statistics of times and proportion of confidence intervals which contains true value 9.9.

5.3 RESULT ANALYSIS

Experiment result is illustrated in figure 2. It can be told that only 12 of 100 confidence intervals built based on classic stimulation and output analysis method which is not using fuzzy parameter contain true value 9.9; for those which fuzzy input parameter is introduced, 86 of 100 confidence intervals built based on stimulation analysis method provided in this paper contain true value 9.9. In addition, because relationship between input and output in model of this case matches condition 1 stated in section 3.2.1, it can be told from table 2 that for the same model input parameter information, new method could make mid-point of confidence interval move to positive direction and width of confidence interval is enlarged. Hence, this method could effectively achieve the two improving method provided in this paper and finally improve reliability of result. Though confidence interval obtained is wider than the one, which does not consider fuzziness, but when information is insufficient, it is necessary cost for maintaining confidence level.

TABLE 2 Comparison of 90% confidence interval considering parameter fuzziness and not considering parameter fuzziness

Experiment times	Fuzzy input parameter			Non-fuzzy input parameter		
	Fuzzy P value	Confidence interval	True value	P value	Confidence interval	True value
1	[8.601,9.434,10.27]	[6.6729,9.71]	Y	9.434	[7.695,8.113]	N
2	[9.987,10.57,11.15]	[10.15,3.92]	N	10.57	[12.11,12.89]	N
3	[9.293,10.40,11.51]	[9.553,3.35]	Y	10.40	[11.28,12.01]	N
4	[9.385,10.06,10.74]	[7.943,11.62]	Y	10.06	[9.753,10.45]	Y
5	[8.956,10.12,11.29]	[8.323,12.12]	Y	10.12	[10.09,10.70]	N
6	[8.586,9.930,11.27]	[7.718,11.52]	Y	9.930	[9.347,9.902]	Y
7	[10.00,10.62,11.24]	[10.39,14.18]	N	10.62	[12.36,13.18]	N
8	[9.189,10.10,11.00]	[8.072,11.63]	Y	10.10	[9.986,10.60]	N
9	[8.441,9.641,10.84]	[6.494,10.28]	Y	9.641	[8.345,8.816]	N
10	[8.431,9.519,10.61]	[7.205,10.18]	Y	9.519	[7.956,8.395]	N
11	[9.514,10.48,11.44]	[9.809,13.61]	Y	10.48	[11.66,12.41]	N
12	[9.180,9.732,10.28]	[6.641,10.44]	Y	9.732	[8.647,9.142]	N
.....	.....	.....	.....	.....	.....	.....
96	[8.265,9.851,11.44]	[7.737,11.54]	Y	9.851	[9.061,9.590]	N
97	[8.347,9.357,10.37]	[6.843,10.28]	Y	9.357	[7.467,7.868]	N
98	[8.558,9.265,9.972]	[5.184,8.976]	N	9.265	[7.134,7.587]	N
99	[8.502,9.742,10.98]	[6.883,10.68]	Y	9.742	[8.680,9.178]	N
100	[9.447,9.761,10.08]	[6.837,10.64]	Y	9.761	[8.746,9.250]	N
Correct times (Proportion)			86(86%)12(12%)			

Specifically speaking, suppose sample quantity  $n=5,10,20,50,100,200,500,1000$ , build 1000 confidence interval for each condition, and compare proportion occupied by confidence intervals (confidence level of 90%) of considering fuzziness of parameter and not considering fuzziness which cover true value. See Figure 5.

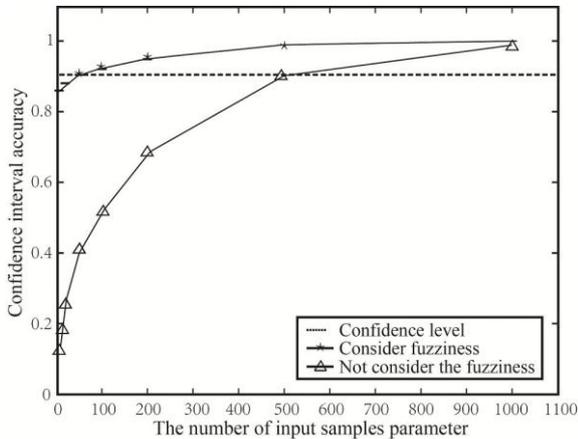


FIGURE 5 Relationship of accuracy and information quantity of confidence interval.

It could be told that when sample numbers are large, i.e. information is sufficient, both method of this paper and classic stimulation method could obtain acceptable result.

**6 Conclusion**

This paper uses stimulation experiment by testing influence of input parameter fuzziness on stimulation output analysis result, pointing out that fuzziness could severely deteriorate reliability of result, which is gotten by using traditional method; Furthermore, it analyses root of this deterioration and provide solution; finally, it provides new analysis method. Because numerical characteristics of statistical magnitude could often be used in statistical analysis, and stimulation method based on possible space in this paper is built based on random fuzzy theory, hence, analysis method provided in this paper uses numerical characteristics of random fuzzy variable.

**References**

[1] Goldsman D, Kim S H, Mmarshall W S, Nelson B L 2002 Ranking and selection for steady-state simulation: Procedures and perspectives *INFORMS Journal on Computing* 14(1) 2-19

[2] Koenig L W, Law A M 2010 A procedure for selecting a subset of size  $m$  containing the one best of  $k$  independent normal populations, with applications to simulation *Communications in Statistics-Simulation and Computation* 14(3) 719-34

[3] Santer T J 2003 A restricted subset selection approach to ranking and selection problems *The Annals of Statistics* 3(2) 334-49

[4] Sullivan D W, Wilson J R 2011 Restricted subset selection procedures for simulation *Operations Research* 37(1) 52-71

[5] Nelson B L, Swann J, Goldsman D, Song W M 2012 Simple procedures for selecting thebest simulated system when the number of alternatives is large *Operations Research* 49(6) 950-63

[6] Wilson J R 2011 A multiplicative decomposition property of the screening-and-selection procedures of Nelson et al. *Operations Research* 49(6) 964-6

[7] Bofinger E, Lewis G J 2012 Two-stage procedures for multiple comparisons with a control *American Journal of Mathematical and Management Sciences* 12(4) 253-75

[8] Nelson B L, Goldsman D 2012 Comparisons with a standard in simulation experiments *Management Science* 47(3) 449-63

[9] Hsu J C, Nelson B L 2008 Optimization over a finite number of system designs with one-stage sampling and multiple comparisons with the best *Proceedings of the Winter Simulation Conference*

[10] Yang W N, Nelson B L 2009 Optimization using common random numbers, control variates, and multiple comparisons with the best *Proceedings of the Winter Simulation Conference*

[11] Yang W N, Nelson B L 2011 Using common random numbers and control variates in multiple-comparison procedures *Operations Research* 39(4) 583-91

Authors	
	<p><b>Zhao Cuirong, born in 1979, Anhui, China</b></p> <p><b>Current position, grades:</b> a Lecturer with the Computer Engineering of Anhui Wonder University Of Information Engineering</p> <p><b>University studies:</b> Bachelor Degree of Engineering from the Anhui Construction Engineering College in 2005, the M.E. degree in School of Computer Science and Technology Anhui University in 2010.</p> <p><b>Scientific interests:</b> Data Mining, Fuzzy Logic and Pervasive Computing.</p>
	<p><b>Wang Honghai, born in 1975, Anhui, China</b></p> <p><b>Current position, grades:</b> an Associate Professor with The Department of Information and Communication Technology of Anhui Sanlian University.</p> <p><b>University studies:</b> B.E. degree from the Electronic Engineering Institute in 1999, and the M.E. degree in School of Computer Science and Technology Anhui University in 2012.</p> <p><b>Scientific interests:</b> Database and Data mining.</p>

# Research of simulation techniques based on rough set theory

**Weiliang Zhu\***

*Qingong College, Hebei United University, Tangshan, Hebei 063000, China*

*Received 1 March 2014, www.tsi.lv*

---

## Abstract

Rough set theory can effectively analyse and deal with incomplete information in simulation techniques. This paper studied the knowledge reduction problem and discrete continuous attributes and improved the BP neural network on the basis of rough set theory. Firstly, methods of attribute reduction of classical are analysis. This paper proposes a heuristic algorithm for reduction of knowledge based on information entropy. Subsequently, it studied the combination algorithm of rough set and neural network. Pre-treatment of sample data based on rough set in dealing with imprecision and uncertainty issues on the edge. The decision rules obtained after reduction in order to map to the training sample of neural network. Finally, the neuron number of hidden layer of neural network and hidden layer makes the neural network more logical. The simulation results show that the simulation technique of rough set and neural network has obvious complementary and reduce the time to train the neural network. It improved the training accuracy and generalization ability simulation techniques achieved satisfactory results.

*Keywords:* simulation techniques, rough set, knowledge reduction, BP neural network

---

## 1. Introduction

The computer simulation technology is the use of computer science and technology building is a comprehensive technology simulation model system and dynamic experiment of the model under some experimental conditions. Human thinking is the organic combination of image thinking and logical thinking. And the integration of rough set and neural network is not accidental, but rather reflect the qualitative and quantitative, human intelligence is clear and implied and serial and parallel cross the conventional thinking mechanism. Intelligent hybrid system is some intelligent technology to combine the use to overcome the shortcomings of single species technology, realize the complementary advantages. So as to facilitate the processing of information in different type in one integrated system. Combining rough set and neural network is the research focus in the field of intelligent information processing. Knowledge acquisition, knowledge expression and reasoning decision rules from a large number of observations and experimental data is an important task for intelligent information processing system. Especially for the inaccurate, incomplete knowledge, rough set theory and artificial neural network shows the infinite charm, rough set method to simulate the human abstract logical thinking, intuitive thinking simulation image of the neural network method.

The literature [1, 2] presented a neural network based on rough sets theory, which consists of rough neurons and conventional neurons. Rough neuron consists of a traditional neuron, the data in the upper and lower boundaries as input or output values of the network.

According to the literature of [3] neural network PID controller general difficult to obtain defect prediction system of output value, we put forward an improved rough set neural network PID controller. The literature [4-6] based on neural network put forward an evaluation method of rough set and neural network project, the method uses data to enrich the data of rough set theory. From the given learning found a set of rules in the sample, extraction rules as the input of neural network. This method simplifies the structure of the neural network, improve the training efficiency. Literature [7-9] presented a neural network model based on rough sets, the method using the data analysis method of rough set, extracted from the data input output mapping rules to the subspace. Analysis and modelling on the stability of rock slope. And compared with the traditional modelling method of neural network, illustrates the effectiveness of the method. The literature [10-13] and put forward a new method of discrete interval of a division of property, according to this classification, a method is proposed for reduction and the establishment of a joint pattern recognition system based on rough sets and LVQ neural network.

The computer simulation technology research starting from the rough set model and elaborates the basic theory of rough set. Study on the discretization method of continuous attributes reduction of knowledge. In view of the limitation of the rough set model, the extended model of fuzzy rough set model and variable precision rough set model are discussed in detail. Based on rough set theory and neural network in the processing logic of complementarily, the combining algorithm application in

---

\* *Corresponding author* e-mail: 915799718@qq.com

system modelling, and the corresponding simulation, and achieved good results in this paper.

**2 Rough set theory**

Rough set cannot confirm individuals vest in border areas, and the boundary area is defined as the upper approximation set and lower approximation set difference set. Because it has certain mathematical formula description, complete by data decision, so as to avoid the influence of subjective factors. In data pre-treatment process, the rough set theory can be applied to feature more accurate extraction. Data preparation process, using rough set theory, data reduction properties of data set for dimension reduction operation. In data mining stage, the rough set theory is used for classification rule discovery. In the interpretation and evaluation process, the rough set theory can be used for the result of statistical evaluation. The basic method of rough set has been in decision-making, forecasting, uncertainty reasoning, the network planning, and the ensemble.

A relational database can be viewed as an information system, the column to property and the line to object. Suppose  $P \subset A$ ,  $x_i, x_j \in U$ , definition two binary relation  $IND(P)$  called equivalence relation:

$$IND(P) = \{(x_i, x_j) \in U \times U \mid \forall p \in P, p(x_i) = p(x_j)\}$$

If and only if  $P(x_i) = P(x_j)$  for all  $p \in P$ , then  $x_i, x_j$  is equivalence relation about the property set  $P$ .

Suppose  $X \subset U$  is a subset of the individual domain in the information system  $S = \{U, A, V, f\}$ , then the lower approximation and upper approximation set and the boundary region of  $X$  are as follows:

$$\underline{PX} = \{Y \in U / P : Y \subseteq X\},$$

$$\overline{PX} = \{Y \in U / P : Y \cap X \neq \emptyset\},$$

$$Bnd_p(X) = \overline{PX} - \underline{PX},$$

where  $\underline{PX}$  are elements in  $X \subset U$  and must be classified collection, that is the maximum of definable sets in  $X$  and  $\overline{PX}$  are elements in  $U$  and must be classified collection, that is the minimum of definable sets in  $X$ .  $BndP(X)$  is the set of elements which cannot be classified in  $X \subset U$ .

Lower and upper approximation schema: assume that there is an information system, there are two attributes. Attributes of a 5 value, attribute two has 6 values. For example, the upper approximation and lower approximate relation as follows:

TABLE 1 Upper approximation and lower approximate relation

<i>U</i>	Headache	Temp.	Flu
<i>U</i> <sub>1</sub>	Yes	Normal	No
<i>U</i> <sub>2</sub>	Yes	High	Yes
<i>U</i> <sub>3</sub>	Yes	Very-high	Yes
<i>U</i> <sub>5</sub>	No	High	No
<i>U</i> <sub>6</sub>	No	Very-high	Yes
<i>U</i> <sub>7</sub>	No	High	Yes
<i>U</i> <sub>8</sub>	No	Very-high	No

Suppose  $S = \{U, A, V, f\}$  is a information system and  $X \subset U$ ,  $P \subseteq A$ , then the accuracy of approximation of  $X$  in  $S$  is as follows:

$$\mu_p(X) = \frac{\mu_{\underline{P}}(X)}{\mu_{\overline{P}}(X)} = \frac{card(\underline{PX})}{card(\overline{PX})}$$

Note:  $card(X)$  is the number of elements in set  $X$ .

Suppose  $S$  is a information system and  $P \subseteq A$ . Let  $\psi = \{x_1, x_2, \dots, x_n\}$  is a classification of  $U$  and  $X_i \subseteq U$ , then the  $P$ - lower approximations and the  $P$ - upper approximations of  $\psi$  can expressed as follows:

$$\underline{P}\psi = \{\underline{PX}_1, \underline{PX}_2, \dots, \underline{PX}_n\},$$

$$\overline{P}\psi = \{\overline{PX}_1, \overline{PX}_2, \dots, \overline{PX}_n\}.$$

The quality of classification of  $\psi$ , which conformed by the attribute subset  $P \subseteq A$  is as follows:

$$\gamma_P(\psi) = \frac{\sum_{i=1}^n card(\underline{PX}_i)}{card(U)}$$

The classification quality represented by the ratio of all object attribute subset  $P$  of correct classification number and the number of information system.

**3 BP neural network**

Neural network also can deal with inaccurate and incomplete knowledge. However rough sets theory and artificial neural network method of two different, rough set method to simulate human abstract thinking, neural network method is used to simulate the image intuitive thinking. Neural network general cannot handle with semantic form of input rough set theory can input qualitative and quantitative or mixed information neural network can be realized without guide.

Sum up the biological neurons in the process of transfer information biological neurons is a multiple input and single output unit. When the nerve cell  $j$  has many inputs  $x_i(i=1,2,\dots,m)$  and single output  $y_j$ , then the relationship between input and output can be expressed as follows:

$$\begin{cases} s_j = \sum_{i=1}^m w_{ij}x_i - \theta_j \\ y_j = f(s_j) \end{cases}$$

The  $w_{ij}$  is connection weights from neurons  $i$  to neurons  $j$ , and  $f$  is the transfer function or called excitation function. After determining the neuron model, topology structure and learning method of characteristics and the ability of a neural depends on the network.

BP Neural network is error back propagation neural network is referred to as "BP", it consists of an input layer, one or more of the hidden layer and an output layer structure, each time by a certain number of neurons composition. These neurons as a man of nerve cells are interrelated. The structure as shown in Figure 1 shows:

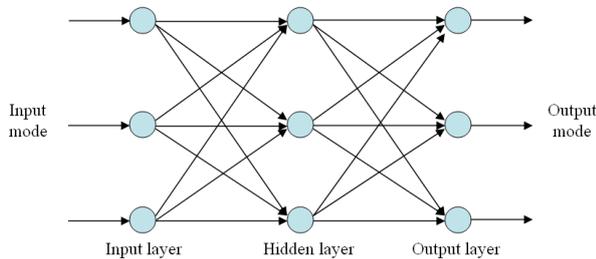


FIGURE 1 The BP neural network model

The traditional rough set of lack of semantic, rough neurons into semantic structure. Rough neural consists of a pair of overlapping normal neuron:  $\bar{r}$ (Upper Neural) and  $\underline{r}$ (Lower Neural), neurons output  $output \bar{r}$  is always greater than the lower neuron output  $output \underline{r}$ :

$$output \bar{r} = \max(f(input \bar{r}), f(input \underline{r})),$$

$$output \bar{r} = \min(f(input \bar{r}), f(input \underline{r})).$$

The function  $f$  is the transfer function.

#### 4 Algorithm design

This template provides authors with most of the formatting specifications needed for preparing their articles.

Neural network can deal with imprecise and incomplete knowledge. However, the rough set method both theory and artificial neural network simulation method is different, the abstract thinking of human rough set, neural network method simulation image intuitive thinking. It has different characteristics.

The basic BP algorithm includes two aspects: the counter-propagating signals prior to the dissemination and error. The calculation of actual output from input to output direction, while the weights and thresholds of correction from the output to the input direction.

#### 4.1 FORWARD PROPAGATION PROCESS SIGNAL

The input  $net_i$  of the first  $i$  node in hidden layer as follows:

$$net_i = \sum_{j=1}^M w_{ij}x_j + \theta_i.$$

The output  $y_i$  of the first  $i$  node in hidden layer as follows:

$$y_i = \phi(net_i) = \phi\left(\sum_{j=1}^M w_{ij}x_j + \theta_i\right).$$

The input  $net_k$  of the first  $k$  node in hidden layer as follows:

$$net_k = \sum_{i=1}^q w_{ki}y_i + a_k = \sum_{i=1}^q w_{ki}\phi\left(\sum_{j=1}^M w_{ij}x_j + \theta_i\right) + a_k.$$

The output  $o_k$  of the first  $k$  node in hidden layer as follows:

$$o_k = \psi(net_k) = \psi\left(\sum_{i=1}^q w_{ki}y_i + a_k\right) = \psi\left(\sum_{i=1}^q w_{ki}\phi\left(\sum_{j=1}^M w_{ij}x_j + \theta_i\right) + a_k\right).$$

#### 4.2 BACK PROPAGATION OF ERROR

Back propagation error, namely first by the output layer to output error calculation of each layer of neurons, and then to adjust the weights of each layer and the threshold error gradient descent method, so that the final output of the modified network can meet its expected value.

For the two type of error criterion, function  $E_p$  of each sample  $p$  as follows:

$$E_p = \frac{1}{2} \sum_{k=1}^L (T_k - o_k)^2.$$

The total system error criterion function on the training sample is as follows:

$$E = \frac{1}{2} \sum_{p=1}^P \sum_{k=1}^L (T_k^p - o_k^p)^2.$$

We can get the correction output layer threshold  $\Delta a_k$  and the correction of hidden layer weights  $\Delta w_{ij}$  and the correction of hidden layer threshold  $\Delta \theta_i$  according to the error gradient descent method are modified correction output layer weights  $\Delta w_{ki}$ . The output layer weights adjustment formula as follows:

$$\Delta w_{ki} = -\eta \frac{\partial E}{\partial w_{ki}} = -\eta \frac{\partial E}{\partial net_k} \frac{\partial net_k}{\partial w_{ki}} = -\eta \frac{\partial E}{\partial o_k} \frac{\partial o_k}{\partial net_k} \frac{\partial net_k}{\partial w_{ki}}$$

And formulas as follows:

$$\frac{\partial E}{\partial y_i} = -\sum_{p=1}^P \sum_{k=1}^L (T_k^p - o_k^p) \cdot \psi'(net_k) \cdot w_{ki}$$

So finally obtained the following formula:

$$\Delta w_{ki} = \eta \sum_{p=1}^P \sum_{k=1}^L (T_k^p - o_k^p) \cdot \psi'(net_k) \cdot w_{ki} \cdot \phi'(net_i) \cdot x_j$$

$$\Delta \theta_i = \eta \sum_{p=1}^P \sum_{k=1}^L (T_k^p - o_k^p) \cdot \psi'(net_k) \cdot w_{ki} \cdot \phi'(net_i)$$

Additional momentum method with the help of network in the correction of the weight, not only consider the role of error in the gradient, and considering the changing trends in the error surface. In the absence of additional momentum effect, the network may get into local minimum shallow value, using the additional momentum effect may be over these minima. The method is every weight in based on back-propagation method on changes on a proportional plus on previous weight or threshold changes in the amount of value, and to generate new weights based on the back propagation method or threshold changes. With additional momentum factor weights and threshold adjustment formula:

$$\Delta w_{ij}(k+1) = (1 - mc)\eta \delta_i p_j + mc \Delta w_{ij}(k),$$

where k is the number of training and MC is the momentum factor and generally around 0.95.

Usually adjust the learning rate criterion is: check whether the real weight reduces the error function, and if it does, then the learning rate is small, can increase an amount; if not, the overshoot, so would reduce the value of the learning rate. The formula gives an adaptive learning rate adjustment formula:

$$\eta(k+1) = \begin{cases} 1.05\eta(k) & \text{if } E(k+1) < E(k) \\ 0.7\eta(k) & \text{if } E(k+1) > 1.04E(k), \\ \eta(k) & \text{other} \end{cases}$$

where  $E(k)$  is the first k step error sum of square. The initial learning rate of the selected range has great arbitrariness.

The calculated properties of mutual information in the greedy algorithm to process it is difficult to judge the importance of the breakpoints. Then, consider the example of those breakpoints can be separated from the most. With the next breakpoint mutual information to a breakpoint to chose which one breakpoint when a breakpoint importance of the same. The improved algorithms are as follows:

Input: decision information system  $S = (U, A, d)$ .

Output: a discrete interval set  $CUT$ .

Step 1: according to the information system  $S$  construction of the new information system  $S^*$ .

Step 2: the initial breakpoint  $CUT = \Phi$ .

Step 3: in each column of  $S^*$ , calculation of the number 1 in the last row.

Step 4: the breakpoint to join in the nuclear  $CUT$ ,  $S_0 = S^*$ .

Step 5: breakpoint at will add to place  $CUT$  in the columns of  $S_0$  and on this column value is 1 line, get information table  $S_1$ .

Step 6: if  $S_1$  is empty, then turn to step 13.

Step 7: Re-calculate the number of column 1 in the last line of  $S_1$ .

Step 8: in the last row if only the maximum value of a column of  $S_1$ , then the row where the breakpoint is added to the  $CUT$ ,  $S_0 = S_1$ , turn to step 5.

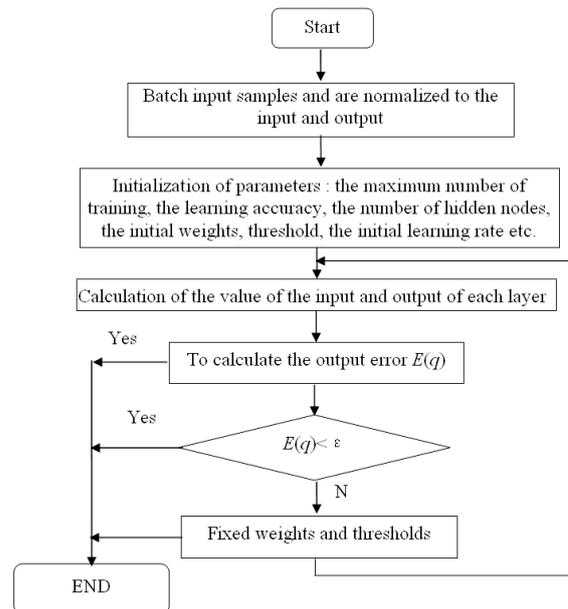


FIGURE 2 The flowchart of the improved BP algorithm program

Step 9: set the maximum value of the column has a maximum value  $n$ , respectively, where the column and the column value is 1 line, get  $n$  information tables  $S_{11}, S_{12}, \dots, S_{1n}$ .

Step 10: if there is any empty table in  $S_{11}, S_{12}, \dots, S_{1n}$ , have to break a table space to join in  $CUT$ , turn to step 13.

Step 11: in each column of  $S_{11}, S_{12}, \dots, S_{1n}$ , calculation of the number 1 in the last row. Were calculated for each table in the last line of column values corresponding to the maximum of the breakpoint of  $S_{1i}$  for mutual information in the table  $S$  are the breakpoint.

Step 12: take the minimum mutual information such as not only a column is the one corresponding to the breakpoint and get the breakpoint  $S_{1i}$  to join in  $CUT$ ,  $S_0 = S_{1i}$ , turn to step 9.

Step 13: the end.

5 Simulation procedure

The simulation results show that, the design method of the control system performance improved significantly. Information system decision table as shown in the table:

TABLE 2 The decision table information system

<i>U</i>	<i>a</i>	<i>b</i>	<i>d</i>
1	0.8	2.0	1
2	1.0	0.5	0
3	1.3	3.0	0
4	1.4	1.0	1
5	1.4	2.0	0
6	1.6	3.0	1
7	1.3	1.0	1

TABLE 3 *S*\* of Table 2

<i>U</i> *	<i>p</i> <sub>1</sub> <sup><i>a</i></sup>	<i>p</i> <sub>2</sub> <sup><i>a</i></sup>	<i>p</i> <sub>3</sub> <sup><i>a</i></sup>	<i>p</i> <sub>4</sub> <sup><i>a</i></sup>	<i>p</i> <sub>1</sub> <sup><i>b</i></sup>	<i>p</i> <sub>2</sub> <sup><i>b</i></sup>	<i>p</i> <sub>3</sub> <sup><i>b</i></sup>	<i>SUM</i>
( <i>u</i> <sub>1</sub> , <i>u</i> <sub>2</sub> )	1	0	0	0	1	0	0	3
( <i>u</i> <sub>1</sub> , <i>u</i> <sub>3</sub> )	1	1	0	0	0	0	1	3
( <i>u</i> <sub>1</sub> , <i>u</i> <sub>5</sub> )	1	1	1	0	0	0	0	3
( <i>u</i> <sub>2</sub> , <i>u</i> <sub>4</sub> )	0	1	1	0	1	0	0	3
( <i>u</i> <sub>2</sub> , <i>u</i> <sub>6</sub> )	0	1	1	1	1	1	1	6
( <i>u</i> <sub>2</sub> , <i>u</i> <sub>7</sub> )	0	1	0	0	1	0	0	2
( <i>u</i> <sub>3</sub> , <i>u</i> <sub>4</sub> )	0	0	1	0	0	1	1	3
( <i>u</i> <sub>3</sub> , <i>u</i> <sub>6</sub> )	0	0	1	1	0	0	0	2
( <i>u</i> <sub>3</sub> , <i>u</i> <sub>7</sub> )	0	0	0	0	0	1	1	2
( <i>u</i> <sub>4</sub> , <i>u</i> <sub>5</sub> )	0	0	0	0	0	1	0	1
( <i>u</i> <sub>5</sub> , <i>u</i> <sub>6</sub> )	0	0	0	1	0	0	1	2
( <i>u</i> <sub>5</sub> , <i>u</i> <sub>7</sub> )	0	0	1	0	0	1	0	2
<i>SUM</i>	3	5	6	3	4	6	5	

It exist number 1 identical at the column of *p*<sub>3</sub><sup>*a*</sup> and *p*<sub>2</sub><sup>*b*</sup> is the largest in Table 3. At this time were the maximum value of columns and the value is 1 line.

TABLE 4 Information table *S*<sub>11</sub>

<i>U</i> *	<i>p</i> <sub>1</sub> <sup><i>a</i></sup>	<i>p</i> <sub>2</sub> <sup><i>a</i></sup>	<i>p</i> <sub>4</sub> <sup><i>a</i></sup>	<i>p</i> <sub>1</sub> <sup><i>b</i></sup>	<i>p</i> <sub>2</sub> <sup><i>b</i></sup>	<i>p</i> <sub>3</sub> <sup><i>b</i></sup>
( <i>u</i> <sub>1</sub> , <i>u</i> <sub>2</sub> )	1	0	0	1	0	0
( <i>u</i> <sub>1</sub> , <i>u</i> <sub>3</sub> )	1	1	0	0	0	1
( <i>u</i> <sub>2</sub> , <i>u</i> <sub>7</sub> )	0	1	0	1	0	0
( <i>u</i> <sub>3</sub> , <i>u</i> <sub>7</sub> )	0	0	0	0	1	1
( <i>u</i> <sub>4</sub> , <i>u</i> <sub>5</sub> )	0	0	0	0	1	0
( <i>u</i> <sub>5</sub> , <i>u</i> <sub>6</sub> )	0	0	1	0	0	1
<i>SUM</i>	2	2	1	2	3	3

TABLE 5 Information table *S*<sub>12</sub>

<i>U</i> *	<i>p</i> <sub>1</sub> <sup><i>a</i></sup>	<i>p</i> <sub>2</sub> <sup><i>a</i></sup>	<i>p</i> <sub>3</sub> <sup><i>a</i></sup>	<i>p</i> <sub>4</sub> <sup><i>a</i></sup>	<i>p</i> <sub>1</sub> <sup><i>b</i></sup>	<i>p</i> <sub>3</sub> <sup><i>b</i></sup>
( <i>u</i> <sub>1</sub> , <i>u</i> <sub>3</sub> )	1	1	0	0	0	1
( <i>u</i> <sub>1</sub> , <i>u</i> <sub>5</sub> )	1	1	1	0	0	0
( <i>u</i> <sub>2</sub> , <i>u</i> <sub>4</sub> )	0	1	1	0	1	0
( <i>u</i> <sub>2</sub> , <i>u</i> <sub>7</sub> )	0	1	0	0	1	0
( <i>u</i> <sub>3</sub> , <i>u</i> <sub>6</sub> )	0	0	1	1	0	0
( <i>u</i> <sub>5</sub> , <i>u</i> <sub>6</sub> )	0	0	0	1	0	1
<i>SUM</i>	2	4	3	2	2	2

Table 4 shows that the column of *p*<sub>3</sub><sup>*a*</sup> and then remove column value is 1 rows. We calculate the mutual information in the table *S* for the breakpoint *p*<sub>3</sub><sup>*a*</sup> as follows:

$$I(p_3^a, p_2^b) = H(p_2^b) - H(p_2^b | p_3^a) = -(\frac{3}{7} \log_2 \frac{3}{7} + \frac{4}{7} \log_2 \frac{4}{7}) - [\frac{3}{7} (\frac{1}{3} \log_2 \frac{1}{3} + \frac{2}{3} \log_2 \frac{2}{3}) + \frac{4}{7} (\frac{2}{4} \log_2 \frac{2}{4} + \frac{2}{4} \log_2 \frac{2}{4})] = 1.9502$$

$$I(p_3^a, p_3^b) = H(p_3^b) - H(p_3^b | p_3^a) = -(\frac{5}{7} \log_2 \frac{5}{7} + \frac{2}{7} \log_2 \frac{2}{7}) - [\frac{3}{7} (\frac{1}{3} \log_2 \frac{1}{3} + \frac{2}{3} \log_2 \frac{2}{3}) + \frac{4}{7} (\frac{2}{4} \log_2 \frac{2}{4} + \frac{2}{4} \log_2 \frac{2}{4})] = 1.8282$$

Table 5 shows that the column of *p*<sub>2</sub><sup>*b*</sup> and then remove column value is 1 rows, the *p*<sub>2</sub><sup>*a*</sup> column number 1 to a maximum of 4. Calculation of the mutual information in the table *S* for the breakpoint *p*<sub>2</sub><sup>*b*</sup> as follows:

$$I(p_2^b, p_2^a) = H(p_2^a) - H(p_2^a | p_2^b) = -(\frac{5}{7} \log_2 \frac{5}{7} + \frac{2}{7} \log_2 \frac{2}{7}) - [\frac{3}{7} (\frac{1}{3} \log_2 \frac{1}{3} + \frac{2}{3} \log_2 \frac{2}{3}) + \frac{4}{7} (\frac{1}{4} \log_2 \frac{1}{4} + \frac{3}{4} \log_2 \frac{3}{4})] = 1.7203$$

Because breakpoint *p*<sub>2</sub><sup>*a*</sup> on the breakpoint *p*<sub>2</sub><sup>*b*</sup> of minimum mutual information, thus remove the breakpoint is *p*<sub>2</sub><sup>*a*</sup> and *p*<sub>2</sub><sup>*b*</sup>, the two breakpoint columns and on this column value is 1 of the line, and the information table is *S*<sub>1</sub>. In fact, because of the breakpoint *p*<sub>2</sub><sup>*b*</sup> is to distinguish objects 4, 5 only a breakpoint in this case the algorithm, and add the first breakpoint *CUT* and remove the breakpoint, the information table *S*<sub>1</sub> is shown as follows.

TABLE 6 Information table  $S_1$ 

$U^*$	$p_1^a$	$p_2^a$	$p_3^a$	$p_4^a$	$p_1^b$	$p_3^b$
$(u_1, u_3)$	1	1	0	0	0	1
$(u_1, u_5)$	1	1	1	0	0	0
$(u_2, u_4)$	0	1	1	0	1	0
$(u_2, u_7)$	0	1	0	0	1	0
$(u_3, u_6)$	0	0	1	1	0	0
$(u_5, u_6)$	0	0	0	1	0	1
<i>SUM</i>	2	4	3	2	2	2

From the Table 6, the column  $p_2^a$  number 1 to a maximum of 4. The breakpoint columns and on this column values for the 1 line, the table is in Table 5. The following process is no longer here.

This algorithm integrates the advantages of several algorithms. Because of the amount of information mutual information represents a source obtained from another source size, so take small mutual information column as a breakpoint makes its own more information. Satisfactory effect was obtained through the example. It solved the limitation of former algorithms effectively.

## References

- [1] Zdzislaw Pawlak, Andrzej Skowron 2007 Rough sets and Boolean reasoning *Information Sciences Publishers Zdzislaw Pawlak chapter 2*
- [2] HUANG Xian-xiang, LONG Yong 2010 Summarization of Distributed Visual Simulation Technology *Journal of System Simulation*. **22**(11) 201-7
- [3] Jian Meng, Minghui Cui, Kaiquan Shi 2006 Variation rough sets in the application of set pair analysis *Journal of Shandong University* **36**(3) 21-7
- [4] Mantas C J, Puche J M, Mantas J M 2006 Extraction of similarity based fuzzy rules from artificial neural networks *Journal of Approximate Reasoning* **43**(2) 111-7
- [5] Ganivada A, Dutta S, Pal S 2011 Fuzzy rough granular neural networks, fuzzy granules, and classification *Journal of Theoretical Computer Science* **41**(4) 53-61
- [6] Kwak K C 2011 A Development of Cascade Granular Neural Networks *Journal of IEICE Transactions on Information and Systems* **94**(7) 307-15
- [7] Fernandez Navarro F, Hervas-Martinez C, Sanchez-Monedero J 2011 MELM-GRBF A modified version of the extreme learning machine for generalized radial basis function neural networks *Journal of Neural Computing* **74**(16) 116-27
- [8] Martinez J M, Escandell-Montero P, Soria-Olivas E 2011 Regularized extreme learning machine for regression problems *Journal of Neural Computing* **74**(17) 71-7

## 6 Conclusions

The computer simulation technology has the advantages of high efficiency, safety, environmental conditions of constraint, can change the time scale etc. It has become an important tool for analysis, design, operation, evaluation, and training of complex systems. Neurons in the hidden layer of neural network and the hidden layer number in order to map to the training sample of neural network. To make neural network more logical, and reduce the time to train the neural network, improve the training accuracy and generalization ability. Based on the research of related literature, the rough set and weak coupling mode using the automatic control in the neural network system, the training sample pre-treatment of the neural network, improve the training speed. In pattern recognition, it is using the concept of rough membership function, constructing sub neural network pattern recognition system based on rough set. The output of the simulation computing improved the anti-noise ability and the accuracy of pattern recognition. It reduced the fuzzy neural network training time, and improving the precision of training.

- [9] Chong Wang, Li Zhao Intermediate 2010 View Synthesis Based on Adaptive BP Algorithm and View Interpolation *JCIT* **10**(5) 1068-71
- [10] Chen H Y, Leou J J 2012 Saliency-directed color image interpolation using artificial neural network and particle swarm optimization *Journal of Visual Communication and Image Representation* **23**(1) 35-41
- [11] Zhu Bing, Wang Hong-fang, Wang Wen-Sheng, Li Yao-Qing 2007 Analysis of relation between flood peak and volume based on set pair *Journal of Sichuan University* **39**(3) 121-30
- [12] Gao J, Gao X 2008 A New Model for Credibilitic Option Pricing *Journal of Uncertain systems* **4**(2) 33-41
- [13] C Robert, Wang L F, Alam M 2012 Training neural networks using Central Force Optimization and Particle Swarm Optimization: Insights and comparisons *Journal of Expert Systems with Applications* **1**(39) 25-37
- [14] Mohammed A, Minhas R, Wu Q M J 2011 Human face recognition based on multi dimensional PCA and extreme learning machine *Journal of Pattern Recognition* **44**(101) 55-67
- [15] Mohamed M H 2011 Rules extraction from constructively trained neural networks based on genetic algorithms *Journal of Neuronal computing* **17**(74) 112-19
- [16] Kahramanli H, Allahverdi N 2009 Rule extraction from trained adaptive neural networks using artificial immune systems *Journal of Expert Systems with Applications* **2**(36) 153-61

## Authors



**Weiliang Zhu, born in June 20, 1979, China**

**Current position, grades:** assistant professor, teacher of Qinggong College of Hebei United University.

**University studies:** mathematical model of information processing and application.

**Scientific interest:** Adaptive control, Rough set, Research on Intelligent Algorithm

**Publications:** 3 (EI)

**Experience:** The main author of the improved adaptive robust control problem

# A new formal representation of granules based on features

**Hong Li<sup>1, 2</sup>, Xiaoping Ma<sup>1</sup>, Zhenghua Xin<sup>2\*</sup>**

<sup>1</sup>*School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou 221008, Jiangsu*

<sup>2</sup>*Laboratory of Intelligent Information processing, School of Information Engineering, Suzhou University, Suzhou 234000, Anhui*

*Received 1 May 2014, www.tsi.lv*

---

## Abstract

In order to provide the unified representation of granules under Granular Computing, this paper studied the granules' basic meaning, descriptions and relationships between the features of granules, and described that the four elements of granules were closely related to granular features. Then the paper presented four-tuples formal representation on data levels based on granular features. The representation includes the object set, feature set, relationship sets and constraints set. Then the representations of several special granules were presented. They showed the unity and advantage of the representation. At last, this paper gave a specific example. By the example, the formal representation has important significance in methodology to solve the granular representation on data levels based on granular features well. And this method is conducive to solve problems and study of granular computing theory.

*Keywords:* granule, feature, formalization, representation, granular computing

---

## 1 Introduction

The world is composed of granules. People understand the world in the way of the granule. In this sense, the granule is the concept applied to practice accompanied with human's existence. It is a natural way to deal with the problem of human daily life [1]. The granule concept should be devoted to L.A.Zadeh who published a paper in 1979 [2]. In that paper, information granule is described to be a proposition.

In the fuzzy sets, quotient space, rough sets, concept lattice or other different models of granular computing, the meaning of granule is more and more clear but different. The different meanings are not conducive to the development of granular computing. The granular computing is based on the formation of granule which includes granule formalization and granule representation [3]. Due to different representations of granule, there is no unified basis and premise for the theory of granular computing. Therefore, a unified form and meaning interpretation of the granule will greatly facilitate the study in granule computing. And it is beneficial to establish the granular computing in a unified theory framework.

The granular computing theory has a long history. The granular computing theory formed a variety of granular representations relative to the specific background. In 1979, L. A. Zadeh studied the size of classified category or the granules [2]. In 1990, Bo Zhang studied the quotient space granular computing. He took quotient set as granules and used granules to calculate. In 1998, T. Y. Lin defined the granule through the binary relations at the view of the neighbourhood [5] in the

subsequent series of literatures. In 2001, Skowron described the information granule [6, 7] and its calculation. In 2002, Y. Y. Yao defined a fundamental granule through logical language [8]. In 2004, Qing Liu took the binary symmetric theory as the fundamental theory [9]. In 2006, Liang Zhang used Galois connection to describe granule [10]. In 2009, Hong Li proposed the four-tuples formal representation of granules [11].

From the paper [1] to the paper [10], all took the fuzzy sets, quotient space, rough set or concept lattice as background in an information system. And they have significant limitations. The formal four-tuples representation of the granule has a certain abstract meaning in [11]. But its application is not convenient. This paper gives the meaning, feature, and the other formal representation of granules based on the above work. In this paper, there is not the background of fuzzy sets, quotient space, rough set or concept lattice. Therefore, this method is conducive to solve the problems.

## 2 The feature of granules

### 2.1 DESCRIPTION OF THE GRANULAR FEATURES

Any object has a large number of basic features, which are independent of other objects. People abstract the common feature of some object and call the feature is part of this class object. Any granule has three basic properties. They are the intrinsic, external and environmental properties [12]. It is believed that granular features are the combination of these three basic properties and the formed values. These three basic

---

\* *Corresponding author* e-mail: begin0000@qq.com

properties include internal feature, external feature and environmental feature. The formed values are in the form of the property-value's binary pair. In addition to basic features, there are other features. The granular identification can be increased. There is Definition 1 and Definition 2 to understand the attribution characteristics of granule.

**Definition 1** (Granular Feature): A special logo of a granule is called granular feature to distinguish it from others. It describes the unit granular ingredients, which include both property and property value. The property is called feature items and the property value is called feature values.

**Definition 2** (Granular feature's description): The description of the formal granular feature is a three tuples:  $GF = \langle id, A, V \rangle$ . The  $id$  represents the granular unique identifier. The  $A$  represents the attribute sets or feature set of the granule. It is an  $n$ -dimensional vector  $(a_1, a_2, a_3, \dots, a_n)$  and  $n$  is a natural number. The  $V$  represents the set of attribute value or the feature value. They are  $n$   $m$ -dimensional vectors  $(b_{i1}, b_{i2}, b_{i3}, \dots, b_{im})$ . The  $m$  and  $n$  are both natural numbers. And  $m$  is feature value with the maximum number in  $n$  attributes  $(i=1, 2, \dots, n)$ .

In general, the description  $GF$  of granular feature can be reduced to tuple form. If  $F$  is the feature set of granule, i.e.  $F = \langle A, V \rangle$ , then  $GF = \langle id, F \rangle$ .

Any granular feature can be divided into the common and unique feature. The common feature contains all common properties in problem domain. It is common feature of all granules in problem domain. The unique feature is the property of part granules in problem domain.

Any feature of granules can be divided into two classes which are the essential and phenomena feature. The essential feature belongs to a particular granule which is the granule that contains all the inherent feature of individual objects. The inherent properties include the structure properties, behaviour, functions, etc. They are common features set of these class individual objects. The phenomenon feature is the external feature shown by the appearance of the class individual objects or the individuals.

Similarly, granular feature can be divided into atomic and complex feature. Atomic feature is no longer divided to other basic feature. The type of atom is real, character and Boolean and other basic data types. A complex feature is constructed from the other defined complex features based on atomic features.

In the decision-making problem, the granular feature can be divided into condition and decision-making feature. The feature describing the problem condition is called the conditional feature. And the feature describing the problem of making decisions is called the decision-making feature.

Feature value is an instance of feature item. The possible values reflect the variability of feature item itself in a different environment. Feature items with the feature value are cases of the relationship. They describe the

general and special relationship between the feature item and feature value.

**Definition 3** (The feature space of problem domain): All the characteristics are given by the related problem domain and structural forms. The structural forms are constructed by all the relationships between features. All the characteristics are named as the feature space of problem domain. The feature space of problem domain is expressed as  $\Omega = \langle FS, FR \rangle$ . The FS is the set of problem domain feature. And FR is a collection of feature relationship.

The feature space in the problem area usually uses the form of the tree structure to represent. It is shown as in the Figure 1. Each node of feature tree in the feature space represents the feature item. It can be identified by a unique feature mark. The feature leaf node represents atom feature. The non-leaf node represents complex features. The feature tree has only one root node to express the root of the problem domain feature space. The root node represents a concept or system. In the feature tree, the parent-child tree nodes are connecting directly. That represents the parent feature and child feature. In this way, the granule can be divided into atomic granule and complex granule corresponding to the granular feature level.

## 2.2 THE RELATIONSHIP AND EXPRESSION OF GRANULAR FEATURE

There are relations between any two things. There are various relationships between feature and feature of granules. The relationship between a pair of features usually has different type. A type of relationship can be applied to a different feature set, and format different semantic information. Therefore, it is particularly important to determine the type of relationship and study the association between various types of mechanisms.

In general, there are four types of relationships between two features. The first is parent-child relationship. The feature item of the parent relationship decides which sub-feature is necessary item to constitute the feature. For example, the "age" feature is such feature. And it decides "juvenile", "youth", "middle", "old" and so on. The second is separation-integration relationship. It is feature aggregation and decomposition. The complex feature can be formed through aggregating features. Simple feature can be formed through decomposing the feature and ultimately form atomic characteristics. The third is dependency relationship. That is an instance value of a set of features. On the other hand, it determines the unique or multiple instance values of another set of features. The "Age" and "title" instance features can decide "professional", "health" and other features instance value. The fourth is association relationship. It exists between the features of the semantic association. The timing relationship between operating features, the combination relationship between describing features, the mapping relationship between operating and describing

features and so on are the examples. At some view, the parent-child relationship and separation-integration relationship together form the combination relationship.

This type of relationship between features can be divided into the combination, dependency and association categories.

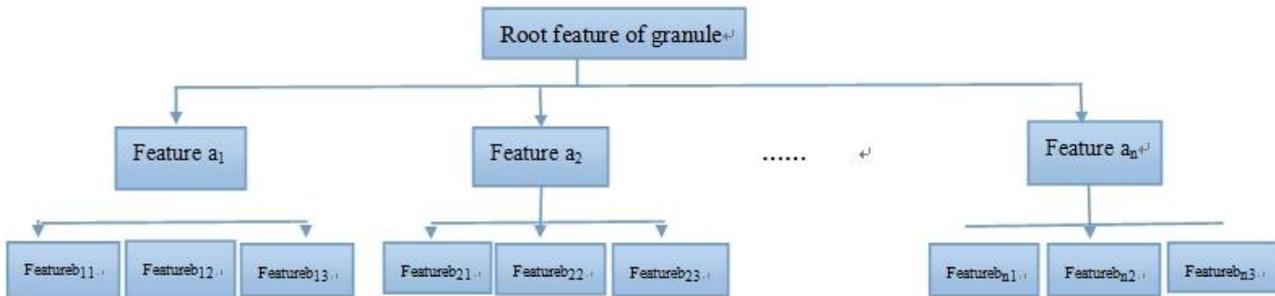


FIGURE 1 Tree diagram of granular features

Here study the expressions of these three type relationships.

**Definition 4** (Feature set is derivable): Suppose that the two granules' features are described as  $GF_1=(id_1, F_1)$  and  $GF_2=(id_2, F_2)$  in the same problem domain feature space  $\Omega$ . If all the granules' features in  $F_2$  can be defined by some granules' feature in  $F_1$  and the features' existing condition in  $F_1$  has no direct relationship whether the features in  $F_2$  exists or not, it is called that feature set  $F_2$  is derivable from the feather sets  $F_1$ . It is denoted as  $F_1 \subseteq F_2$ .

**Proposition 1:** suppose that  $F_1$  and  $F_2$  are the two features of the same problem domain  $\Omega$  set in feature space, and the feature set  $F_2$  is derivable from the feature set  $F_1$ . Any feature in  $F_2$  is the parent feature of all features that is used to define the feature self in  $F_1$ . And all the features in  $F_1$  are the sub feature of  $F_2$ .

**Definition 5** (Feature set is dependent): Suppose that the two granular features are  $GF_1=(id_1, F_1)$  and  $GF_2=(id_2, F_2)$  in the same problem domain feature space  $\Omega$ . If  $F_1 \subseteq F_2$ , we call that there exist dependent relationships between the feature sets  $F_2$  and  $F_1$ .  $F_2$  exists dependent of  $F_1$ . It is expressed as  $F_1 \Rightarrow F_2$ .

**Definition 6** (Feature sets are equivalent): Suppose that the description of the two granules feature are  $GF_1=(id_1, F_1)$  and  $GF_2=(id_2, F_2)$  in the same problem domain feature space  $\Omega$ . If  $F_1 \Rightarrow F_2$  and  $F_2 \Rightarrow F_1$ , we say that the feature set  $F_2$  and the feature set  $F_1$  are equivalent. Or it is interdependent between the feature set  $F_2$  and the feature set  $F_1$ . It is denoted as  $F_1 \equiv F_2$ .

**Definition 7** (Feature set is independent): Suppose that the description of the two granules feature are  $GF_1=(id_1, F_1)$  and  $GF_2=(id_2, F_2)$  in the same problem domain feature space  $\Omega$ . If there not exists  $F_1 \Rightarrow F_2$  and also not exists  $F_2 \Rightarrow F_1$ , then the feature sets  $F_2$  and  $F_1$  are independent each other. It is denoted as  $F_1 \neq F_2$ .

**Proposition 2:** Suppose  $F_1$  and  $F_1$  are two independent feature sets in the same problem domain feature space  $\Omega$ . Then any one feature in  $F_1$  and any one feature in  $F_2$  are characteristic independent on each other.

**Definition 8** (Feature relationships): Suppose that the description of the two granules feature are  $GF_1=(id_1, F_1)$

and  $GF_2=(id_2, F_2)$  in the same problem domain feature space  $\Omega$ . The formula  $R_F = \{ \langle GF_1, GF_2 \rangle = \{ \langle a_1, a_2 \rangle \mid a_1 \in F_1, a_2 \in F_2 \}$  denotes the relations between  $GF_1$  and  $GF_2$ .

If the feature sets  $F_2$  and  $F_1$  are independent separately, then there are no relations or exists empty relation between  $GF_1$  and  $GF_2$ . It is denoted as  $R_F = \Phi$ .

If there is dependent relation between the feature sets  $F_2$  and  $F_1$ , then there exists not empty relation between  $GF_1$  and  $GF_2$ . It is denoted as  $R_F \neq \Phi$ .

**Definition 9** (Feature relationship's representation): Suppose that the description of the two granules feature are  $GF_1=(id_1, F_1)$  and  $GF_2=(id_2, F_2)$  in the same problem domain feature space  $\Omega$ . And there is no empty relation between  $GF_1$  and  $GF_2$ . There are the following inferences.

- If  $F_1 \subseteq F_2$ , there exist  $a_1, a_2 \in F_1$  and  $a_3 \in F_2$ . And there are the features of father and son between  $a_3$  and  $a_1, a_2$ . However, there are brother characteristics between  $a_1$  and  $a_2$ . There exists  $a_3 = a_1 \cup a_2$ , then we say that there exists combination relations between  $GF_1$  and  $GF_2$ . That is denoted as  $GF_1 \rightarrow cGF_2$ .
- If there exist  $X \subseteq F_1$  and  $Y \subseteq F_2$  to make  $X \Rightarrow Y$  established in  $\Omega$ , then there is one-way dependent relation between  $GF_1$  and  $GF_2$ . It is expressed as  $GF_1 \rightarrow iGF_2$ . Especially, if there exist  $X \Rightarrow F_1$  and  $Y \Rightarrow F_2$ , and  $X \Rightarrow Y$  and  $Y \Rightarrow X$  are simultaneously established in  $\Omega$ , then there are bidirectional dependent relations:
- or interdependent relations. It is expressed into  $GF_1 \leftrightarrow iGF_2$ .
- If there exist  $X \subseteq F_1$  and  $Y \subseteq F_2$ , and there is a constrained condition  $J=f(X, Y)$  between the feature combination  $X$  and the feature combination  $Y$ , we call there is a link between  $GF_1$  and  $GF_2$ . It is expressed as  $GF_1 \rightarrow a GF_2$ . Generally, suppose that there is multiple granules features' description  $GF_i=(id_i, F_i)$  and  $X_i \subseteq F_i(i=1,2,\dots)$  in the same problem domain feature space  $\Omega$ . If there is a constrained condition  $J=f(X_1, X_2, \dots)$  among many granules feature combination  $J=f(X_1, X_2, \dots)$ , then there is multiple correlation between  $GF_i(i=1,2,\dots)$ . It is denoted as  $\rightarrow a(GF_1, GF_2, \dots)$ .

The combination, dependency and association relationships between granule features are shown in the Figure 2. The granule feature and relationships between granule features provide an important prerequisite for their formal representations.

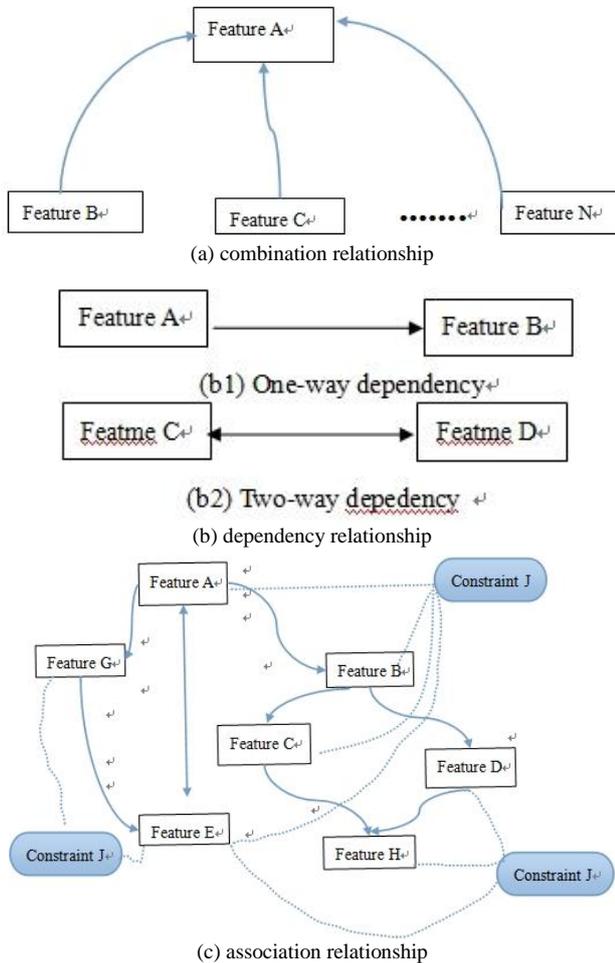


FIGURE 2 Relationships of granular features

### 3 The formal representation of granules

Since the granule is the model which is a shared objective of the various components [13]. In the process of the applying granular computing to solve real problems, the complex relationships of large original problem must be formed. Therefore, the large original problem needs to be decomposed into a number with complete information granule. The necessary merging operation and other treatments have been done according to the specific requirements and background. This requires the definition of internal restraint mechanism. The mechanism used the same granule structure of formal representation. The definition must include the complete information of granules.

Some references analysed the granular objects, features, relationships, and the state of the four basic elements [13]. Describing a granule should be used the four elements. The four elements are the object, the

feature, the relationship and the state. As objects, relations and states three elements are closely related with the feature elements, using these four elements as formal representation of granule is essentially feature-based representation.

### 3.1 THE FORMAL REPRESENTATION OF GRANULE

Information science uses four levels [14, 15]. These levels are in the following order, from the data, the information, the knowledge to the wisdom. They describe the different abstraction levels of the human-centred information procession. Each level needs to use the appropriate computer system to achieve. The information is commonly referred to general information. It includes the form, the carrier or the content [16]. In addition, general information and data is usually confused. Data is the information carrier. And the performance of information is data. The information in this paper indicates the general information with the indiscriminate data.

In the data layer, people use database management system (DBMS) to store and retrieve the structured data. The basic unit of storage and retrieval is the entity containing some values of the object attribute. The information storage and retrieval is the specific form of granule. The paper only discusses the formal representation of the granule in the data level here.

**Definition 10** (Granule formal representation): Suppose the given feature space problem domain is  $\Omega$ . A granule of this problem domain can be formally defined as the following four bytes in the following quadruple form:  $G=(O, F, R, J)$

The  $G$  is used to describe a granule. The  $O$  represents those studied individual objects set of granules. It is called the granules' extension or the domain. It can be a common set and also be a fuzzy set. In some aspects, it can be limited. It also can be infinite. For example, in a certain interval, the  $F$  is all the feature set of granules. It is named as the granules' connotation,  $F=(F_1, F_2, \dots, F_n)$ .  $F_i=\{ F_{i1}, F_{i2}, \dots, F_{im} \} (1 \leq i \leq n, m \geq 0)$  is the  $i$ -feature of each individual object in  $O$ . The  $F_{ij}$  is the  $j$ -feature of each individual object of granules in the  $i$ -feature. The  $F$  embodies the internal, the external and environmental features, etc. The  $R$  represents a collection of all relationships in granules, including the sequence structure, topology structure, and graph structure of various elements or individuals and so on. It also includes all kinds of relationships between features and various relations. The various relations first establishes between individual objects and the characteristics. Then they and features output all kinds of relationships. Of course, the output relationships are included in the  $F$ . The  $J$  is a set of constraints, which regulates the legal status of granules and timing sequence and so on. And it is used to represent the static, dynamic and unity properties of granules.  $J$  includes the static constraints and dynamic

constraints. The static constraints refer to the determining formulas of granules. They provide all the legal status of this kind of granules. Dynamic constraints include the timing constraints and real-time constraints. These constraints are the time constraints, distance constraints or the rules limitation.

**Definition 11** (Relationships between individual objects): Suppose the feature space of given problem domain is  $\Omega$ , a granule of the problem domain can be formally defined as  $G=(O, F, R, J)$ , let  $R_O=\{<o_1, o_2> | o_1, o_2 \in O\}$  indicate the relationship between the individual object  $o_1$  and the individual object  $o_2$ . If there is a feature  $a \in F$  to make the establishment of  $a(o_1) = a(o_2)$ , then we say that there is not distinguishable relationship between  $o_1$  and  $o_2$  on the feature  $a$ . It is denoted as  $R_O \neq \Phi$ . Otherwise, we call there exist distinguishable relationship between  $o_1$  and  $o_2$  on the feature  $a$ . It is denoted as  $R_O = \Phi$ .

According to definition 8, definition 10 and definition 11, there exists  $R = R_F \cup R_O$ . On the basis of definition 10, there are the following three propositions with the hypothesis that  $(O, F, R, J)$  is the primary granule.

**Proposition 3:** The granule  $(O', F', R', J')$  is atomic granule, if and only if  $O' \subseteq O \wedge |O'|=1, R'=\Phi, J'=\Phi$ .

Atomic granules have the smallest size. Because atomic granules do not contain any constraints, they can use every feature item of features at utmost. This makes the absolute multiplexing degree reach the maximum. But as the feature's degree is the smallest. The efficiency of reusing is the lowest.

**Proposition 4:** The granule  $(O', F', R', J')$  is elementary granule, if and only if  $F' \subset F, R' \subset R, J' \subset J$ .

The basic granules are implemented by non atomic features, while atomic granules can also be considered as a special kind of basic particle ( $R'=\Phi$ ). The two are essentially the same. They both are mapping that a feature directly to an independent particle. The basic granule's size is larger than the size of atomic granules. Thus the mining efficiency is improved. But because of too many demanding characteristics, the cost of mining massive data increases.

**Proposition 5:** The granule  $(O', F', R', J')$  is original granule, if and only if  $O' \subseteq O, F' \subseteq F, R' \subseteq R, J' \subseteq J$ .

### 3.2 THE UNITY ANALYSIS OF GRANULAR FOUR-TUPLES REPRESENTATION

The four-tuples  $(O, F, R, J)$  of definition 10 can be well unified fuzzy sets, quotient space and rough set etc. It can be used as a model of the granular computing

When taking the granules  $(O, F, R, J)$  as the whole study corresponding to the original problem space is the original granule. At this time, the element of granule can be in the elements of the original problem. And it can also be the sub-granule composed of elements of the original problem space.

The relationship  $R$  is an equivalence relation at the macro view of granule. A raw granule is composed of all

subsets or sub-granules corresponding to a certain equivalent division of the original problem. This granule description corresponds to a quotient space of the original problem space.

When the relationship  $R$  is an equivalence relation, a sub-granule is a subset of the original problem. This granule description corresponds to the granule of a rough set.

The relationship  $R$  is fuzzy relation at the fuzzy view of granule. It naturally corresponds to the granule of a fuzzy set.

Therefore, the granule representation form with four-tuples  $(O, F, R, J)$  of the definition 10, is convenient to converse at different views, such as macro and micro, the overall and partial, precise and vague. It is consistent with the features of human intelligence. It is also convenient to absorb some relatively mature theory, extract their common characteristic, and benefit to establish a unified theoretical framework for granular computing.

### 3.3 THE ADVANTAGES OF GRANULAR FOUR-TUPLES REPRESENTATION

Compared with the definition of granule in the literature [4-11], using the four-tuples  $(O, F, R, J)$  of definition 10 to describe the granule has strong advantages.

1) the granule representation of this paper reflects the unity of granule. The granules defined in literature [5-10] take fuzzy sets, quotient space, rough set or concept lattice as background in information systems. It has significant limitations. In addition, the formal representation of the four-tuples for granule has a certain abstract meaning in [11], but its application is not convenient. According to the above section B, the granular representation of this paper gets rid of the background of fuzzy sets, quotient space, rough set or concept lattice. It can be well unified the existing quotient space, rough sets and fuzzy sets etc. It is a typical granular computing model.

2) it reflects the concrete and abstract unity characteristic. In the four-tuples  $(O, F, R, J)$ , some specific individual objects, the specific feature and conditions together form a granule. Therefore, the granule is the set of entities, which reflects the specific nature of granules. At the same time, this granule representation solves problems based on the user's needs, and establishes a way to perceive the objective world. In this computing model, the objective world is granulated into granules. Because the virtual world in users' opinion is composed of the granules, it can support and help users understand the objective world. And this reflects the abstract characteristic of granule.

3) it reflects the unification of connotation and extension well. In the four-tuples  $(O, F, R, J)$ ,  $O$  indicates the individual object set of the granules. It is the extension of granule. It is also a universe term. It can be a normal set or a fuzzy set.  $F$  is the set of all the granule

features. It is the content of granules,  $F = (F_1, F_2, \dots, F_n)$ .  $F_i = \{F_{i1}, F_{i2}, \dots, F_{im}\}$  ( $1 \leq i \leq n, m \geq 0$ ) is the  $i$ -feature of each individual object in  $O$ .  $F_{ij}$  is the  $i$ - $j$ -feature of each individual object of the granule ( $1 \leq i \leq n, 0 \leq j \leq m$ ).  $F$  reflects the internal and external, environmental and other basic features of the granule. Compared with the granule definitions in the literature [4-11], the granule representation of this paper reflects preferable unification of the connotation and extension.

- Fourth, it reflects the unification of static and dynamic feature. In the four-tuples  $(O, F, R, J)$ ,  $J$  is a set of constraints. It provides the legal statuses of granules and the timing between the legal statuses with such constraints. It includes the static constraints and dynamic constraints. Static constraints determine the granule style. They provide all the legal status of such granules. Dynamic constraints include granule timing constraints, real-time constraints between the legal statuses. It may be time constraints or distance limitations, the rules limitations and so on. Compared the static feature with the definition of granule in the literature [4-10], it reflects the unification of static and dynamic feature well.
- Fifth, it reflects the unification of the original granule and the basic granules. When any problem is studied, the original granule takes the whole object as a granule. Atomic granule does not contain any constraints, with minimum size of the granule. It usually contains only a feature of each item. The basic granule is achieved by the non-atomic feature. Atomic granule is a special form of basic granules. The propositions from the proposition 3 to 5, it gets the original, the basic or atomic granule according to different selection of  $O, F, R, J$  in the four-tuples  $(O, F, R, J)$ . They reflect the unification of the original granule and the basic granule. But the granular representation forms in the literature [4-11] do not have this feature.

**4 A case study**

This paper gives a case to illustrate the formalization of granule. The Table 1 is a table of building block information to study the changes of granule.

TABLE 1 Building block information table

Element	Colour	Shape	Size	Stability
$x_1$	red	triangle	big	stability
$x_2$	red	triangle	big	stability
$x_3$	yellow	circle	small	instability
$x_4$	yellow	circle	small	instability
$x_5$	blue	square	big	stability
$x_6$	red	circle	middle	instability
$x_7$	blue	circle	small	instability
$x_8$	blue	square	middle	instability

Suppose 8 blocks form a set of objects  $O$ . It is noted that  $O = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ . Each building block has four kinds of feature items. That feature set is denoted

by  $A$ . Each item has different feature values. The value of each feature set is denoted by  $V_A$ .  $A$  and  $V_A$  constitute a building block feature. It is denoted that  $F = A \times V_A$ ,  $A = \{\text{colour, shape, size, stability}\}$ ,  $V_A = \{\text{values of each feature item}\}$ . There is such relationship,  $R: F \rightarrow O$ . The objects set  $O$  is decomposed into a smaller subset according to the given constraints  $J$  to select the feature item values. The original granule problem of building blocks can be expressed as  $G = (O, F, R, J)$ .

For example, according to the colour feature, these building blocks can be granulated into  $R_1 = \{\text{red, yellow, blue}\}$  three tablets. Then all the building blocks in the red colour constitute the kernel granule  $X_1 = \{x_1, x_2, x_6\}$ . The yellow building blocks constitute the kernel granule  $X_2 = \{x_3, x_4\}$ . The building blocks in the blue colour constitute the kernel granule  $X_3 = \{x_5, x_7, x_8\}$ . Then  $O/R_1 = \{X_1, X_2, X_3\} = \{\{x_1, x_2, x_6\}, \{x_3, x_4\}, \{x_5, x_7, x_8\}\}$  according to the colour granulation. Similarly,  $R_2 = \{\text{triangle, square, circle}\}$ ,  $R_3 = \{\text{large, medium, small}\}$  indicates the size.  $R_4 = \{\text{stable, unstable}\}$  indicates the stability. The equations are as follows according to the shape granulation, the size granulation and the stable granulation respectively.

$$O/R_2 = \{Y_1, Y_2, Y_3\} = \{\{x_1, x_2\}, \{x_5, x_8\}, \{x_3, x_4, x_6, x_7\}\},$$

$$O/R_3 = \{Z_1, Z_2, Z_3\} = \{\{x_1, x_2, x_5\}, \{x_6, x_8\}, \{x_3, x_4, x_7\}\},$$

$$O/R_4 = \{T_1, T_2, T_3\} = \{\{x_1, x_2, x_5\}, \{x_3, x_4, x_6, x_7, x_8\}\}.$$

If one wants to find the "large and triangular" building blocks, he/she can use  $R_2 \cap R_3$  to find the required result granule. Similarly, if one wants to find the "big or triangular" building blocks, he can use  $R_2 \cup R_3$  terms.

When mining massive data, it is complex and difficult to extract the valuable information. And it needs large calculation by operating data sets directly. This method presented in this paper can get knowledge in the higher level of abstraction. It may be more universal. It can construct the corresponding concept or identity-level tree by the abstract feature value in a given data table. Multiple abstraction levels can be achieved to solve granular computing problem. At the same time, the original data set can also be horizontal compression to reduce data size, and improve efficiency for solving problems. At this point, more abstract levels of granule-based knowledge representation can be introduced. Such granular representation can provide an important prerequisite for category theory based granular computing method.

The granular triple-tuples representation based on knowledge level would be study in other paper.

**5 Conclusion**

This paper provides a uniform granular representation for granular computing, which studies the feature of the granule. It illustrates that the four elements of granule are closely related to granular feature. It presents a new feature-based four-tuples representation of granule. By

comparison, this paper's granular four-tuples representation reflects the unity of granule, the specific and abstraction, the connotation and extension, static and dynamic, origination and basic granule. The formal representation has important methodology. It is a good solution to the data level feature-based granular representation problem. It is conducive to solve problem and granular computing theory. Of course, this representation is only a preliminary exploration, which still needs further test for its effectiveness. In addition, the abstract representation of granules needs further study.

The introduction of category theory provides an important prerequisite for the granular calculation.

### Acknowledgment

This work is supported by National Nature Science Foundation under Grant 60974126, Key Project of Anhui Education Department Nature Science Foundation under Grant KJ2012A263, Anhui Provincial Natural Science Foundation under Grant 10040606Q64.

### References

- [1] Liu Q, Sun H, Wang H F 2008 The Present Studying State of Granular Computing and Studying of Granular Computing Based on the Semantics of Rough Logic *Chinese Journal of Computers* 31(4) 543-55
- [2] Zadeh L A 1979 Fuzzy sets and information granularity *Advances in Fuzzy Set Theory and Applications North-Holland Amsterdam* 3-18
- [3] Zadeh L A 1998 Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems *Soft Computing Berlin Springer-Verlag* 23-5
- [4] Zhang B, Zhang L 2007 Problem solving theory and application (The 2<sup>nd</sup> Edition) *Tsinghua University Press* 4-14
- [5] Lin T Y 1998 Granular Computing on Binary Relations II Rough Set Representations and Belief Functions *Rough Sets In Knowledge Discovery Physica -Verlag* 121-40
- [6] Skowron A 2001 Toward Intelligent Systems Calculi of Information Granules *Proceedings of International Workshop on Rough Set Theory and Granular Computing Bulletin of International Rough Set Society* 5(1/2) 9-30
- [7] Skowron A, Stepaniuk J, Peters James F 2001 Extracting Patterns Using Information Granules *Proceedings of International Workshop on Rough Set Theory and Granular Computing Bulletin of International Rough Set Society* 5(1/2) 135-42
- [8] Yao Y Y, Yao J T 2002 Granular Computing as a Basis for Consistent Classification problems *Proceedings of PAKDD'02 Workshop on Foundations of Data Mining CIICM 2002(5)* 101-6
- [9] Liu Q, Huang Z H 2004 G -Logic and Its Resolution Reaching *Chinese Journal of Computers* 27(7) 865-73
- [10] Zhang L 2006 The Research on Model of Knowledge Discovery Based on Granularity and Concept Lattice *Hefei University of Technology*
- [11] Li H 2009 Granule, Granular Set and Granular System *Proceedings of the IEEE 4<sup>th</sup> International Conference for Granular Computing Los Alamitos California IEEE Computer Society CPS* 340-5
- [12] Yao Y Y 2006 Granular computing for data mining *The SPIE Conference on Data Mining, Intrusion Detection, Information Assurance and Data Networks Security Kissimmee USA* 1-12
- [13] Li H 2010 Research on Granule and Its Characteristics *Journal of Suzhou University* 25(11) 8-11 29
- [14] Bellinger G, Castro D, Mills A 2010 Data, Information, Knowledge and Wisdom [EB/OL]. [www.systems-thinking.org/dikw/dikw.htm](http://www.systems-thinking.org/dikw/dikw.htm)
- [15] Jing N N, Cheng J Y 2005 Data, Information, Knowledge and Wisdom *Information Science* 23(12) 1786-90
- [16] Zhu X L 2000 Fundamentals of applied information theory *Tsinghua University Press* 1-2 4888-900

### Authors



**Hong Li, born in 1965, Anhui province, China**

**Current position, grades:** professor China University of Mining and Technology, Xuzhou in Jiangsu province.

**University studies:** master degree in computer science and technology from Hefei University of Technology, China.

**Scientific interest:** intelligent computing, algorithm design and analysis, data mining.

**Publications:** [1] Further Study on Granulating Thinking. *Journal of Suzhou university*. 2013.28(5). [2] Research on four-element model of granular computing. *Computer Engineering and Applications*. 2013.49(4). [3] Formal Method of Granular Synthesis Based on Category Theory. *Computer Engineering*. 2011.37(14).



**Xiao-ping Ma**

**Current position, grades:** professor, China University of Mining and Technology, Xuzhou in Jiangsu province.

**University studies:** doctor degree from China University of Mining and Technology.

**Scientific interest:** control theory and application, computer application.

**Publications:** 25

**Experience:** The executive dean of School of information and electrical engineering, CUMT



**Zhenghua Xin**

**Current position, grades:** lecture of Suzhou University.

**University studies:** master degree from Hebei University of Technology, China.

**Scientific interest:** cloud computing, internet of things.

**Publications:** 17

# Interleaving semantics and action refinement in event structures

**Weidong Tang<sup>1, 2, 3</sup>, Jinzhao Wu<sup>2, 3\*</sup>, Meiling Liu<sup>2, 4</sup>**

<sup>1</sup>Chengdu Institute of Computer Applications, Chinese Academy of Sciences, Chengdu 610041, China

<sup>2</sup>School of Information Science and Engineering, Guangxi University for Nationalities, Nanning 530006, China

<sup>3</sup>Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis, Nanning 530006, China

<sup>4</sup>Science Computing and Intelligent Information Processing of GuangXi higher education key laboratory, Nanning 530023, China

Received 1 March 2014, www.tsi.lv

---

## Abstract

An event structure acts as a denotational semantic model of concurrent systems. Action refinement is an essential operation in the design of concurrent systems. But there exists an important problem about preserving equivalence under action refinement. If two processes are equivalent with each other, we hope that they still can preserve equivalence after action refinement. In linear time equivalence and branching time equivalence spectrum, interleaving equivalences, which include interleaving trace equivalence and interleaving bisimulation equivalence are not preserved under action refinement [9-11, 14, 16, 21]. In this paper, we define a class of concurrent processes with specific properties and put forward the concept of clustered action transition, which ensures that interleaving equivalences are able to preserve under action refinement.

*Keywords:* event structure, action refinement, concurrency, interleaving equivalence, clustered equivalence

---

## 1 Introduction

In order to model concurrent systems, we hope to have formal method for hierarchical structure. Action refinement is the core operation of the hierarchical method, which interprets an action in higher abstract layer with a process in lower layer, hence reduces the level of abstraction and eventually reaches its implementation layer. In the development course from top to bottom of concurrent system, we must first build models, which depict the system with description language of top layer; subsequently, according to these descriptions, we complete its implementation. This course often requires equivalence notion to verify the correctness of implementation of system. More concretely, assuming that  $\mathcal{P}$  represents the descriptions of system and  $\mathcal{Q}$  represents its implementation, if  $\mathcal{P}$  is equivalent with  $\mathcal{Q}$  (expressed as  $\mathcal{P} \approx \mathcal{Q}$ ), then this shows that  $\mathcal{Q}$  is correct. In development, the description  $\mathcal{P}$  of a system can be refined layer-by-layer, accordingly its implementation  $\mathcal{Q}$  can be converted from framework into code or electronic components. Only the description and its implementation at all levels are required to maintain equivalence so as to ensure correctness of its implementation. This leads to an important question what kind of equivalence is maintained under action refinement, that is, if two concurrent systems are equivalent with each other, we hope that they still can preserve equivalence after action refinement.

Vogler [14, 15] first raised the basic thought of preserving equivalence under action refinement. Czaja, Van Glabbeek and Goltz [21] demonstrated that if interleaving bisimulation equivalence doesn't produce choice operations or action self-concurrences after actions are refined then it can preserve equivalence under action refinement, but interleaving trace equivalence still cannot preserve. Goltz and Wehrheim [20] proved that history preserving bisimulation is consistent with global causal dependencies, but they did not further discuss about the problem how to preserve equivalence under action refinement, and did not discuss that there are other situations under environment of action independencies. Van Glabbeek and Goltz [11, 21] summarized the research results of action refinement, gave a detailed explanation for preserving equivalence problem under action refinement, and proved that interleaving equivalence cannot preserve under action refinement in general, but did not discuss further. Moreover, no work further discusses preserving problem under action refinement of interleaving trace equivalence and interleaving bisimulation equivalence. In this paper, we define a class of concurrent processes with specific properties and put forward the concept of clustered action transition, which ensures that interleaving equivalences are able to preserve under action refinement in the absence of constraints.

---

\* *Corresponding author* e-mail: himrwujz@126.com

**2 Event structures and action refinement**

Let Act be a set of actions.

**Definition 2.1** [2, 12, 18]: A event structure  $\mathcal{P}$  is a 5-tuple  $(E, <, \#, \Delta, l)$ , where

- $E$  is the set of events;
- $< \subseteq E \times E$  is irreflexive partial relation, and satisfy the rule of finite causes that  $\forall e \in E : \{e1 \in E | e1 < e\}$  is finite; In addition, its inverse " $<$ " is expressed as " $>$ ";
- $\# \subseteq E \times E$  is irreflexive and finite conflicting relation, and satisfy the rule of inheriting of conflict that  $\forall e1, e2, e3 \in E : e1 < e2 \wedge e1 \# e3 \Rightarrow e2 \# e3$ ;
- $\Delta \subseteq E \times E$  is irreflexive concurrent relation, altogether with  $<$  and  $\#$  to satisfy the principle of partition that  $< \cup \# \cup \Delta = E \times E$ ,  $e1 \Delta e2 \Leftrightarrow \neg(e1 = e2 \vee e1 < e2 \vee e2 < e1 \vee e1 \# e2)$ ;
- $l : E \rightarrow Act$  is a label function of actions.

In this paper, let  $\mathbb{S}$  denote the set of all event structures.

**Definition 2.2** [2, 11]: Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . A relation between  $\mathcal{P}$  and  $\mathcal{Q}$  is called isomorphic (expressed as  $\mathcal{P} \cong \mathcal{Q}$ ) if there exists an bijection between their sets and preserves corresponding relations with  $<$ ,  $\#$ ,  $\Delta$  and same corresponding labels.

Unless specified, we do not discriminate isomorphic event structures.

The behaviours of a event structure are expressed with its configurations which are considered as all possible states of the system. The following is its definition.

**Definition 2.3** [2, 12, 18]: Let  $X$  be a subset of the set  $E_{\mathcal{P}}$  of all events in event structure  $\mathcal{P}$ .

- (1)  $X$  is left closed if  $\forall e1, e \in E : e \in X \wedge e1 < e \Rightarrow e1 \in X$ ;
- (2)  $X$  is conflicted-free if  $\mathcal{P}|_X$  is conflicted-free;
- (3)  $X$  is a configuration if  $X$  is not only left closed but also conflicted-free.

Here, let  $C(\mathcal{P})$  represent the set of all configurations in event structure  $\mathcal{P}$ .

A configuration  $X$  ( $X \in C(\mathcal{P})$ ) is called (successfully) terminated configuration if  $\forall e \in E : e \notin X \Rightarrow \exists e1 \in X : e1 \# e$ .

The event structure is also often represented with graph, where  $\rightarrow, \dots$  stands for casual relation and immediately conflict relation in event structure respectively, inherited conflict relation is not considered and indenpent relation is not explicitly expressed.

**Example 2.1:** A system  $\mathcal{P} = (a|b) + (c;b;d)$ , executing either  $a, b$  concurrently or  $c, b$  and  $d$  sequentially, can be described by the event structure with

events  $e1, e2, e3, e4, e5$  with  $l(e1) = a, l(e3) = c, l(e2) = l(e4) = b, l(e5) = d$ , where  $e1 \Delta e2, e3 < e4 < e5$ , each of  $e1, e2$  is in conflict with each of  $e3, e4, e5$ . This event structure is expressed in Figure 1.

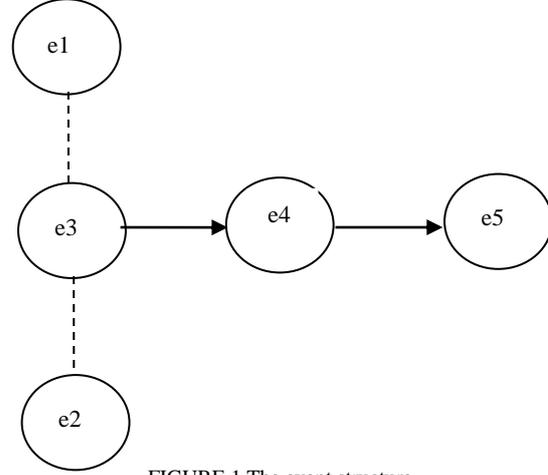


FIGURE 1 The event structure

Its configurations are  $\emptyset, \{e1\}, \{e2\}, \{e1, e2\}, \{e3\}, \{e3, e4\}, \{e3, e4, e5\}$ , where  $\{e1, e2\}, \{e3, e4, e5\}$  are terminated configurations.

Main thought of action refinement [1, 3, 5, 8, 22] is: replace an action in higher layer with a process in lower layer, do in the same way layer by layer, until get detailed design or implementation of the system.

**Definition 2.4** [13, 14, 17, 22]: A function  $ref : Act \rightarrow E - \{\emptyset\}$  is called a refinement function of event structure, if  $\forall a \in Act : ref(a)$  is not empty, finite and conflict-free. Let  $\mathcal{P} \in \mathbb{S}$ .  $ref(\mathcal{P})$  is an event structure defined as follows:

$$E_{ref(\mathcal{P})} = \left\{ (e, e') \mid e \in E_{\mathcal{P}}, e' \in E_{ref(l_{\mathcal{P}}(e))} \right\},$$

$$(e1, e1') <_{ref(\mathcal{P})} (e2, e2') \text{ iff } e1 <_{\mathcal{P}} e2 \text{ or}$$

$$e1 = e2 \wedge e1' <_{ref(l_{\mathcal{P}}(e1))} e2'.$$

$$(e1, e1') \#_{ref(\mathcal{P})} (e2, e2') \text{ iff } e1 \#_{\mathcal{P}} e2,$$

$$(e1, e1') \Delta_{ref(\mathcal{P})} (e2, e2') \text{ iff } e1 \Delta_{\mathcal{P}} e2 \text{ or}$$

$$e1 = e2 \wedge e1' \Delta_{ref(l_{\mathcal{P}}(e1))} e2',$$

$$l_{ref(\mathcal{P})}(e, e') = l_{ref(l_{\mathcal{P}}(e))}(e').$$

**Example 2.2:** Continue with Example 2.1. Assuming that  $ref(b) = (b1;b2) + b3$ , the event structure after action refinement is expressed in Figure 2.

Under action refinement, each event  $e$  labelled by the action  $b$  is replaced by a disjoint copy,  $\mathcal{P}_e$ , of  $ref(b)$ , i.e., the event  $e2$  is replaced by  $(e22 < e23) \# e21$  and the event  $e4$  is replaced by  $(e41 < e42) \# e43$  [17]. The causality and conflict structure is inherited from  $\mathcal{P}$ : all events which were casually before  $e$  will be casually before all events of  $\mathcal{P}_e$ , every event which casually followed  $e$  will casually follow all events of  $\mathcal{P}_e$ , and all events in conflict with  $e$  will be in conflict with all events of  $\mathcal{P}_e$ .

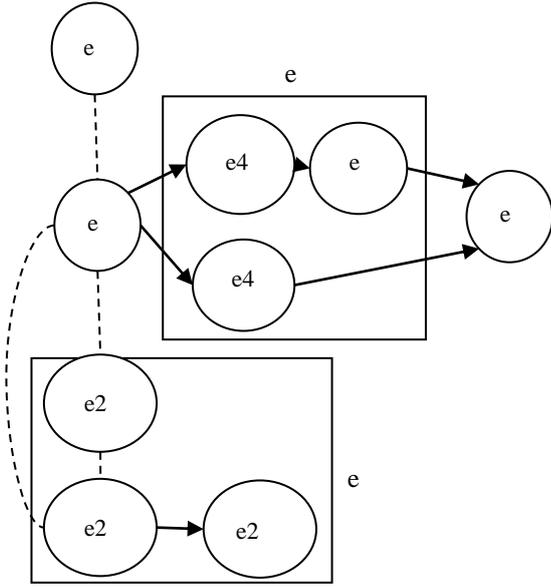


FIGURE 2 The event structure after action refinement

### 3 Interleaving equivalences

Interleaving equivalences can be divided into interleaving trace equivalence and interleaving bisimulation equivalence [9, 10, 11, 21]. To begin with, we give the definition of single action transition.

**Definition 3.1:** Let  $\mathcal{P} \in \mathbb{S}$ . A transition relation  $X \xrightarrow{a} \mathcal{P} X'$  is called single action transition if  $a \in Act, X, X' \in C(\mathcal{P}), X \subseteq X'$ , and  $\exists e \in E_{\mathcal{P}} : X' - X = e, l_{\mathcal{P}}(e) = a$ .

Here,  $X \xrightarrow{a} \mathcal{P} X'$  denotes that the state expressed by configuration  $X$  turns into the one expressed by configuration  $X'$  after performing action  $a$  in event structure  $\mathcal{P} \in \mathbb{S}$ .

Then, we define trace and interleaving trace equivalence.

**Definition 3.2:** Let  $\mathcal{P} \in \mathbb{S}$ . A word  $w = a_1 \cdots a_n \in Act^*$  is called a trace of event structure  $\mathcal{P}$  if

$$\exists X_0, \dots, X_n \in C(\mathcal{P}) : X_0 = \emptyset \text{ and } X_{i-1} \xrightarrow{a_i} X_i, \quad i = 1, \dots, n.$$

Here,  $trs(\mathcal{P})$  represents the set of all traces in event structure  $\mathcal{P}$ .

**Definition 3.3:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . A relation between  $\mathcal{P}$  and  $\mathcal{Q}$  is called interleaving trace equivalence (expressed as  $\mathcal{P} \approx_{it} \mathcal{Q}$ ) if  $trs(\mathcal{P}) = trs(\mathcal{Q})$ .

Furthermore, we define interleaving bisimulation equivalence.

**Definition 3.4:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . A relation  $R \subseteq C(\mathcal{P}) \times C(\mathcal{Q})$  is called a interleaving bisimulation between  $\mathcal{P}$  and  $\mathcal{Q}$  if  $(\emptyset, \emptyset) \in R$  and if  $(X, Y) \in R$  then

$$X \xrightarrow{a} \mathcal{P} X', a \in Act \Rightarrow \exists Y' : Y \xrightarrow{a} \mathcal{Q} Y' \wedge (X', Y') \in R, \\ Y \xrightarrow{a} \mathcal{Q} Y', a \in Act \Rightarrow \exists X' : X \xrightarrow{a} \mathcal{P} X' \wedge (X', Y') \in R.$$

A relation between  $\mathcal{P}$  and  $\mathcal{Q}$  is called interleaving bisimulation equivalence (expressed as  $\mathcal{P} \approx_{ib} \mathcal{Q}$ ) if there exists a interleaving bisimulation between them.

Finally, we introduce another kind of equivalence named pomset trace equivalence which depends on partial order theory. we will still study whether quotient event structure and original event structure are pomset trace equivalences or not.

**Definition 3.5:** Let  $\mathcal{P} \in \mathbb{S}$ .

(1) Let  $R1 = \langle X1, <_{X1}, l|_{X1} \rangle$  and  $R2 = \langle X2, <_{X2}, l|_{X2} \rangle$  be two partial sets labelled in Act  $R1$  and  $R2$  is isomorphic with each other (expressed as  $R1 \cong_p R2$ ) if there exists a bisection  $f : X \rightarrow Y$  such that  $\forall e1, e2 \in X1 : e1 <_{X1} e2 \Leftrightarrow f(e1) <_{X2} f(e2)$ , and  $l|_{X1} = l|_{X2} \circ f$ . A isomorphic class of partial set in the set Act is called pomset.

(2) Partial sets of configuration  $X$  is  $pomset(X) = \left[ (X, <_X, l|_X) \right]_{\cong_p}$ . The set of all pomsets in event structure  $\mathcal{P}$  is  $pomsets(\mathcal{P}) = \{ pomset(X) | X \in C(\mathcal{P}) \}$ .

**Definition 3.6:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . Assume that  $pomsets(\mathcal{P}) = \{ pomset(X) | X \in C(\mathcal{P}) \}$  and  $pomsets(\mathcal{Q}) = \{ pomset(Y) | Y \in C(\mathcal{Q}) \}$ .

A relation between  $\mathcal{P}$  and  $\mathcal{Q}$  is called pomset trace equivalence (expressed as  $\mathcal{P} \approx_{pt} \mathcal{Q}$ ) if  $pomsets(\mathcal{P}) = pomsets(\mathcal{Q})$ .

Article [17] has proved that pomset trace equivalence is able to preserve under action refinement.

**Proposition 3.1:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ , let  $ref$  be a refinement function. If  $\mathcal{P} \approx_{pt} \mathcal{Q}$  then  $ref(\mathcal{P}) \approx_{pt} ref(\mathcal{Q})$ .

The proof confers theorem 8.1 in article [11].

Pomset trace equivalence requires that casual and independent relations between events must be consistent. Apparently, it can directly reach a conclusion that pomset trace equivalence implies interleaving trace equivalence.

**Proposition 3.2:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . If  $\mathcal{P} \approx_{pt} \mathcal{Q}$  then  $\mathcal{P} \approx_{it} \mathcal{Q}$ .

Subsequently, we introduce distinct kind of equivalence called history preserving bisimulation which is finer than interleaving bisimulation and is able to preserve under action refinement.

**Definition 3.7:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . A relation  $R \subseteq C(\mathcal{P}) \times C(\mathcal{Q}) \times \mathcal{P}(E_{\mathcal{P}} \times E_{\mathcal{Q}})$  is called a history preserving bisimulation between  $\mathcal{P}$  and  $\mathcal{Q}$  if  $(\emptyset, \emptyset, \emptyset) \in R$ , there always exists  $(X, Y, f) \in R$ , then  $f$  is a bijection between  $X$  and  $Y$ ;

$$X \xrightarrow{a} \mathcal{P} X', a \in Act \Rightarrow \exists Y', f' :$$

$$Y \xrightarrow{a} \mathcal{Q} Y', (X', Y', f') \in R \wedge f'|_X = f ;$$

$$Y \xrightarrow{a} \mathcal{Q} Y', a \in Act \Rightarrow \exists X', f' :$$

$$X \xrightarrow{a} \mathcal{P} X', (X', Y', f') \in R \wedge f'|_X = f .$$

Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . A relation between  $\mathcal{P}$  and  $\mathcal{Q}$  is called history preserving equivalence (expressed as  $\mathcal{P} \approx_h \mathcal{Q}$ ) if there exists a history preserving bisimulation between them.

History preserving equivalence is able to preserve under action refinement. The following gives the conclusion by means of Proposition 3.3.

**Proposition 3.3:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ , let  $\text{ref}$  be a refinement function. If  $\mathcal{P} \approx_h \mathcal{Q}$  then  $\text{ref}(\mathcal{P}) \approx_h \text{ref}(\mathcal{Q})$ .

The proof can confer theorem 9.1 and proposition 9.2 in article [11].

According to definition 3.4 and definition 3.7, we immediately attain a result that history preserving equivalence can deduce interleaving bisimulation equivalence.

**Proposition 3.4:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . If  $\mathcal{P} \approx_h \mathcal{Q}$  then  $\mathcal{P} \approx_{ib} \mathcal{Q}$ .

Interleaving equivalences cannot preserve under action refinement. However, if restrictions with some properties based on event structure model are placed, some useful results will be attained. The following proposition shows that if no independency exists in event structure then interleaving equivalences (include interleaving trace equivalence and interleaving bisimulation equivalence) are able to preserve under action refinement.

**Proposition 3.5:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ , let  $\text{ref}$  be a refinement function  $\Delta_{\mathcal{P}} = \Delta_{\mathcal{Q}} = \emptyset$ .

(1) If  $\mathcal{P} \approx_{it} \mathcal{Q}$  then  $\text{ref}(\mathcal{P}) \approx_{it} \text{ref}(\mathcal{Q})$ .

(2) If  $\mathcal{P} \approx_{ib} \mathcal{Q}$  then  $\text{ref}(\mathcal{P}) \approx_{ib} \text{ref}(\mathcal{Q})$ .

**Proof:**

(1) For any configuration  $X \in C(\mathcal{P})$ , by definition 2.2, there is  $\#_{\mathcal{P}|_X} = \emptyset$ . Moreover, because  $\Delta_{\mathcal{P}} = \emptyset$ ,  $X$  is left closed, there exists a total order relation in  $X$ . By definition 3.2,  $X$  corresponds to unique trace (expressed as  $w_X$ ), namely,

$$\forall X, Y \in C(\mathcal{P}) \wedge X \neq Y \Rightarrow \exists w_X, w_Y \in \text{trs}(\mathcal{P}) \wedge w_X \neq w_Y .$$

On the condition that actions names cannot give rise to confusion, we represent corresponding event with its action name,  $\text{pomsets}(\mathcal{P}) = \text{trs}(\mathcal{P})$ . Similarly,  $\text{pomsets}(\mathcal{Q}) = \text{trs}(\mathcal{Q})$ . Given  $\mathcal{P} \approx_{it} \mathcal{Q}$ , namely,  $\text{trs}(\mathcal{P}) = \text{trs}(\mathcal{Q})$ , we obtain

$$\text{pomsets}(\mathcal{P}) = \text{pomsets}(\mathcal{Q}), \text{ hence } \mathcal{P} \approx_{it} \mathcal{Q} \Rightarrow \mathcal{P} \approx_{pt} \mathcal{Q} .$$

By proposition 3.1,  $\mathcal{P} \approx_{pt} \mathcal{Q} \Rightarrow \text{ref}(\mathcal{P}) \approx_{pt} \text{ref}(\mathcal{Q})$ . Also, by proposition 3.2, we obtain that  $\text{ref}(\mathcal{P}) \approx_{pt} \text{ref}(\mathcal{Q}) \Rightarrow \text{ref}(\mathcal{P}) \approx_{it} \text{ref}(\mathcal{Q})$ .

(2) Based on the above proof, For any configuration  $X \in C(\mathcal{P})$ , there only exists a total order relation in  $X$ .

Similarly, for any configuration  $Y \in C(\mathcal{Q})$ , So is  $Y$ . Due

to  $\mathcal{P} \approx_{ib} \mathcal{Q}$ , we obtain  $X \xrightarrow{a} \mathcal{P} X', a \in Act \Rightarrow \exists Y' : Y \xrightarrow{a} \mathcal{Q} Y' \wedge (X', Y') \in R$  and  $(\emptyset, \emptyset)$ ,

$(X, Y), (X', Y') \in R$ . Now, we want to prove  $X \cong_p Y$  and  $X' \cong_p Y'$ . Distinctly, the bisimulation between  $\mathcal{P}$

and  $\mathcal{Q}$  starts from initial configuration  $\emptyset$  and gets to states of bisimulation (expressed as  $X, Y$ ) by performing same actions. Also,  $X, Y$  is total order and they execute corresponding events of same actions, hence  $X \cong_p Y$ .

Similarly,  $X' \cong_p Y'$ . According to symmetry [6, 19], for any

$$Y \xrightarrow{a} \mathcal{Q} Y', a \in Act \Rightarrow \exists X' : X \xrightarrow{a} \mathcal{P} X' \wedge (X', Y') \in R ,$$

there also exist  $Y \cong_p X$  and  $Y' \cong_p X'$ . Following the

above method, we are able to find a history preserving bisimulation between  $\mathcal{P}$  and  $\mathcal{Q}$ , i.e.,  $\mathcal{P} \approx_h \mathcal{Q}$ . Also, by

proposition 3.3,  $\mathcal{P} \approx_h \mathcal{Q} \Rightarrow \text{ref}(\mathcal{P}) \approx_h \text{ref}(\mathcal{Q})$ . By

proposition 3.4,  $\text{ref}(\mathcal{P}) \approx_h \text{ref}(\mathcal{Q}) \Rightarrow \text{ref}(\mathcal{P}) \approx_{ib} \text{ref}(\mathcal{Q})$ .

#### 4 Clustered action transition

We present a new class of action transition, where  $A$  is a multiple set in action set  $Act$  and all actions within  $A$  independently perform with each other. We call this multiple set  $A$  as a clustered action and call this class transitions as clustered action transitions. With clustered action transitions, we construct two new types of equivalence.

**Definition 4.1:** Let  $\mathcal{P} \in \mathbb{S}$ . A transition  $X \xrightarrow{A} X'$  is called a clustered action transition if  $A \in N^{Act}$  (i.e.,  $A$  is multiple set in  $Act$ )

$X, X' \in C(\mathcal{P}), X \subseteq X', X' - X = G$ , the events in set  $G$  satisfy:

(1) entire independency of causes:

$$\forall d, e \in G: (d \Delta_{X'} e) \wedge (\{e_1 \in E_{\mathcal{P}} | e_1 \Delta_{\mathcal{P}} e\} \cup \{e\} = \{e_2 \in E_{\mathcal{P}} | e_2 \Delta_{\mathcal{P}} d\} \cup \{d\} = G) \text{ and } l_{\mathcal{P}}(G) = A.$$

(2) same causality:

$$\forall e_1 \in E_{\mathcal{P}} \setminus G, \exists e_2 \in G: (e_1 < e_2 \Rightarrow \forall e_3 \in G: e_1 < e_3) \\ \vee (e_1 > e_2 \Rightarrow \forall e_3 \in G: e_1 > e_3);$$

(3) same conflict relation:

$$\forall e_1 \in E_{\mathcal{P}} \setminus G, \exists e_2 \in G: (e_1 \# e_2 \Rightarrow \forall e_3 \in G: e_1 \# e_3).$$

Here,  $l_{\mathcal{P}}(G) \in N^{Act}$  results from

$$l_{\mathcal{P}}(G)(a) = |\{e \in G | l_{\mathcal{P}}(e) = a\}|.$$

Clustered action transition  $X \xrightarrow{A} X'$  represents that after independently executing all actions of set  $A$ , the state expressed by configuration  $X$  is changed into the one expressed by configuration  $X'$  in event structure  $\mathcal{P}$ .

**Definition 4.2:** suppose that  $e$  is an event in event structure  $\mathcal{P}$ . Independency set of cause on  $e$   $\varphi(e)$ , including  $e$  itself, is defined as  $\varphi(e) = \{e_1 \in E_{\mathcal{P}} | e_1 \Delta_{\mathcal{P}} e\} \cup \{e\}$ .

In order to study the follow-up problems, we give a proposition in advance.

**Proposition 4.1:** Let  $\mathcal{P} \in \mathbb{S}$ . If all transitions in  $\mathcal{P}$  are clustered action transitions then

(1)  $\forall d \in E_{\mathcal{P}} \Rightarrow \exists (X \xrightarrow{A} X') : \varphi(d) = G$ , where  $A \in N^{Act}$ ,  $X, X' \in C(\mathcal{P}), X \subseteq X', X' - X = G$  and  $l_{\mathcal{P}}(G) = A$ .

(2)  $\forall (X \xrightarrow{A} X') \Rightarrow \exists d \in E_{\mathcal{P}} : G = \varphi(d)$ , where  $A \in N^{Act}$ ,  $X, X' \in C(\mathcal{P}), X \subseteq X', X' - X = G$  and  $l_{\mathcal{P}}(G) = A$ ;

(3)  $\bigcup_{e \in E_{\mathcal{P}}} \varphi(e) = E_{\mathcal{P}}$ ;

(4)  $\forall e_1, e_2 \in E_{\mathcal{P}} : e_1 \neq e_2 \wedge \neg(e_1 \Delta e_2) \Rightarrow \varphi(e_1) \cap \varphi(e_2) = \emptyset$ .

**Proof:** Conclusions (1), (2) and (3) is very easy to reach, omitted here. Only give proof of (4). If  $\varphi(e_1) \cap \varphi(e_2) \neq \emptyset$  then  $\exists e \in \varphi(e_1) \cap \varphi(e_2)$ . By definition 4.1, we obtain  $\varphi(e) = \varphi(e_1) = \varphi(e_2)$ , consequently obtain either  $e_1 \Delta e_2$  or  $e_1 = e_2$ . This contradict with  $e_1 \neq e_2 \wedge \neg(e_1 \Delta e_2)$ . Accordingly, arrive at the conclusion that

$$\forall e_1, e_2 \in E_{\mathcal{P}} : e_1 \neq e_2 \wedge \neg(e_1 \Delta e_2) \Rightarrow \varphi(e_1) \cap \varphi(e_2) = \emptyset.$$

The above proposition shows that if all transitions in an event structure are clustered action transitions, then all independent actions involved in a clustered action transition are seen as a “big” action, its corresponding events can be thought of as a “big” event. Hence, not only no independency between events exists in this event structure but also independency sets of cause divide set of events into different parts, which induces an equivalence relation in set of events of this event structure.

**Definition 4.3:** Let  $\mathcal{P} \in \mathbb{S}$ . A sequence  $w = A_1 \cdots A_n$  ( $A_i \in N^{Act}$  ( $i = 1, \dots, n$ )) is called a clustered trace of event structure  $\mathcal{P}$  if  $\exists X_0, \dots, X_n \in C(\mathcal{P}) : X_0 = \emptyset$  and  $X_{i-1} \xrightarrow{A_i} X_i, i = 1, \dots, n$  is a clustered action transition.

Here,  $Clusteredtr(\mathcal{P})$  represents the set of all clustered traces of event structure  $\mathcal{P}$ .

Subsequently, we define clustered trace equivalence.

**Definition 4.4:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ , all transitions in  $\mathcal{P}$  and  $\mathcal{Q}$  be clustered transitions. Let  $Clusteredtr(\mathcal{P})$  and  $Clusteredtr(\mathcal{Q})$  be the sets of all clustered traces of  $\mathcal{P}, \mathcal{Q}$  respectively. A relation between  $\mathcal{P}$  and  $\mathcal{Q}$  is called clustered trace equivalence (expressed as  $\mathcal{P} \approx_{ct} \mathcal{Q}$ ) if  $Clusteredtr(\mathcal{P}) = Clusteredtr(\mathcal{Q})$ .

The following proposition shows that clustered trace equivalence can imply interleaving trace equivalence.

**Proposition 4.2:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . If all transitions in event structures  $\mathcal{P}$  and  $\mathcal{Q}$  are clustered action transitions then  $\mathcal{P} \approx_{ct} \mathcal{Q} \Rightarrow \mathcal{P} \approx_{it} \mathcal{Q}$ .

**Proof:** By definition 4.1 and proposition 4.1,  $\forall A_i, A_j$  ( $i = 1, \dots, n, j = 1, \dots, n, i \neq j$ ) in a clustered trace, then all actions in  $A_i$  are independent of those in  $A_j$ . Performing of any action in  $A_i$  does not interfere with those in  $A_j$ , vice versa. Consequently,  $Clusteredtr(\mathcal{P}) = Clusteredtr(\mathcal{Q}) \Rightarrow trs(\mathcal{P}) = trs(\mathcal{Q})$ , i.e.  $\mathcal{P} \approx_{ct} \mathcal{Q} \Rightarrow \mathcal{P} \approx_{it} \mathcal{Q}$ .

Provide that  $W = A_1, \dots, A_n$ , where  $A_i \in N^{Act}$  ( $i = 1, \dots, n$ ) be a clustered trace, and  $|A_i|$  stand for the number of elements, then the clustered trace  $W$  corresponds to  $|A_1|! \times |A_2|! \times \dots \times |A_n|!$  general traces.

After discussing clustered trace equivalence, we begin with another new class of equivalence and study whether they can maintain under action refinement or not. This class of equivalence is designated clustered bisimulation equivalence. The following provides for its definition.

**Definition 4.5:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ , let all transitions in  $\mathcal{P}$  and  $\mathcal{Q}$  be clustered action transitions. A relation  $R \subseteq C(\mathcal{P}) \times C(\mathcal{Q})$  is called a clustered bisimulation

between  $\mathcal{P}$  and  $\mathcal{Q}$  if  $(\emptyset, \emptyset) \in R$  and if  $(X, Y) \in R$  then

$$X \xrightarrow{A} \mathcal{P} X', A \in N^{Act} \Rightarrow \exists Y' : Y \xrightarrow{A} \mathcal{Q} Y' \wedge (X', Y') \in R,$$

$$Y \xrightarrow{A} \mathcal{Q} Y', A \in N^{Act} \Rightarrow \exists X' : X \xrightarrow{A} \mathcal{P} X' \wedge (X', Y') \in R.$$

A relation between  $\mathcal{P}$  and  $\mathcal{Q}$  is called clustered bisimulation equivalence (expressed as  $\mathcal{P} \approx_{cb} \mathcal{Q}$ ) if there exists a clustered bisimulation between them.

Obviously, by definition 4.5, we come to a decision that  $\mathcal{P} \approx_{cb} \mathcal{Q} \Rightarrow \mathcal{P} \approx_{ct} \mathcal{Q}$ .

**Proposition 4.3** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . If all transitions in event structures  $\mathcal{P}$  and  $\mathcal{Q}$  are clustered action transitions then  $\mathcal{P} \approx_{cb} \mathcal{Q} \Rightarrow \mathcal{P} \approx_{ib} \mathcal{Q}$ .

Procedure of proof is similar to that of proposition 4.2, omitted here.

## 5 Preserving of interleaving equivalences

These papers [9-11, 14, 16, 21] have demonstrated that interleaving equivalence cannot preserve under action refinement in the general case. However, proposition 3.5 shows that interleaving equivalence without concurrency can preserve under action refinement. In this part, we will testify that in the presence of concurrency interleaving equivalence is able to preserve under action refinement in given conditions. In order to prove the conclusion, we first discuss the relationship between clustered trace equivalence and partial order trace equivalence.

**Proposition 5.1:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . If all transitions in event structures  $\mathcal{P}$  and  $\mathcal{Q}$  are clustered action transitions, then  $\mathcal{P} \approx_{ct} \mathcal{Q} \Rightarrow \mathcal{P} \approx_{pt} \mathcal{Q}$ .

**Proof:** For any configuration  $X \in C(\mathcal{P})$ , since all transitions in event structures  $\mathcal{P}$  are clustered action transitions, there exists unique one clustered trace, let it be  $W$ , construction of  $W$  starts from initial configuration  $\emptyset$  and evolves into configuration  $X$ . Also, given  $\mathcal{P} \approx_{ct} \mathcal{Q}$ , we obtain  $Clusteredtr(\mathcal{P}) = Clusteredtr(\mathcal{Q})$ , then there exists a configuration  $Y \in C(\mathcal{Q})$  reached by the implementation of clustered trace  $W$ . Obviously,  $X \cong_p Y$ , i.e.  $\forall X \in C(\mathcal{P}) \Rightarrow \exists Y \in C(\mathcal{Q}) \wedge X \cong_p Y$ . Similarly, we obtain  $\forall Y \in C(\mathcal{Q}) \Rightarrow \exists X \in C(\mathcal{P}) \wedge Y \cong_p X$ . Isomorphism ensures they have same labels, therefore  $pomsets(\mathcal{P}) = pomsets(\mathcal{Q})$  i.e.  $\mathcal{P} \approx_{pt} \mathcal{Q}$ .

At present, we discuss about the problem that interleaving equivalence can preserve under action refinement in given conditions.

**Proposition 5.2:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . Assume that all transitions in event structures  $\mathcal{P}$  and  $\mathcal{Q}$  be clustered

action transitions,  $ref$  be a refinement function. If  $\mathcal{P} \approx_{ct} \mathcal{Q}$  then  $\mathcal{P} \approx_{it} \mathcal{Q} \Rightarrow ref(\mathcal{P}) \approx_{it} ref(\mathcal{Q})$ .

**Proof:** By proposition 5.1,  $\mathcal{P} \approx_{ct} \mathcal{Q} \Rightarrow \mathcal{P} \approx_{pt} \mathcal{Q}$ . Also, by proposition 3.1,  $\mathcal{P} \approx_{pt} \mathcal{Q} \Rightarrow ref(\mathcal{P}) \approx_{pt} ref(\mathcal{Q})$ . And, by proposition 3.2,  $ref(\mathcal{P}) \approx_{pt} ref(\mathcal{Q}) \Rightarrow ref(\mathcal{P}) \approx_{it} ref(\mathcal{Q})$ . Consequently, we arrive at a conclusion that  $\mathcal{P} \approx_{it} \mathcal{Q} \Rightarrow ref(\mathcal{P}) \approx_{it} ref(\mathcal{Q})$ .

This proposition shows that if two event structures satisfy clustered trace equivalence and their all transitions are clustered action transitions, then interleaving equivalence can hold under action refinement. This is an important conclusion and is one of research goals to be achieved.

Subsequently, we discuss the topic that interleaving bisimulation equivalence can preserve under action refinement in given conditions.

**Proposition 5.3:** Let  $\mathcal{P}, \mathcal{Q} \in \mathbb{S}$ . If all transitions in event structures  $\mathcal{P}$  and  $\mathcal{Q}$  are clustered action transitions then  $\mathcal{P} \approx_{cb} \mathcal{Q} \Rightarrow \mathcal{P} \approx_h \mathcal{Q}$ .

**Proof:** Assume that the relation  $R \subseteq C(\mathcal{P}) \times C(\mathcal{Q})$  be a clustered bisimulation between  $\mathcal{P}$  and  $\mathcal{Q}$ . We construct a history preserving bisimulation  $\tilde{R} \subseteq C(\mathcal{P}) \times C(\mathcal{Q})$  between  $\mathcal{P}$  and  $\mathcal{Q}$  by inductive method:

(1) First let the element  $(\emptyset, \emptyset)$  add into the set  $\tilde{R}$ ;

(2) Now assume that  $(X, Y)$  have already been an element of  $\tilde{R}$ , and given  $\mathcal{P} \approx_{cb} \mathcal{Q}$ , also assume that  $X \xrightarrow{A} \mathcal{P} X', Y \xrightarrow{A} \mathcal{Q} Y', A \in N^{Act}$  and  $(X, Y) \in R$ , also because of  $\mathcal{P} \approx_{cb} \mathcal{Q} \Rightarrow \mathcal{P} \approx_{ct} \mathcal{Q}$ , by definition 4.3, we obtain  $\mathcal{P} \approx_{pt} \mathcal{Q}$ . Hence,  $X \cong_p Y$  and  $X' \cong_p Y'$ . If there are  $n$  ( $n > 1$ ) elements in multiple action set  $A$  and the execution among these elements is interleaving then there exist  $n!$  sequences of single action transition:  $X \xrightarrow{a_1} \mathcal{P} X_1 \xrightarrow{a_2} \mathcal{P} X_2 \cdots X_{n-1} \xrightarrow{a_{n-1}} \mathcal{P} X'$ .

Accordingly, there exists the corresponding sequence of single action transition:  $Y \xrightarrow{a_1} \mathcal{Q} Y_1 \xrightarrow{a_2} \mathcal{Q} Y_2, \dots, Y_{n-1} \xrightarrow{a_{n-1}} \mathcal{Q} Y'$ , and meets  $X_1 \cong_p Y_1, X_2 \cong_p Y_2, \dots, X_{n-1} \cong_p Y_{n-1}$ . Apparently, add these  $n$  elements  $(X_i, Y_i) (i=1, \dots, n)$  to  $\tilde{R}$ , meanwhile, add  $(X', Y')$  to  $\tilde{R}$ , where  $X_i \cong_p Y_i$ . Along this route, we construct all elements of the set  $\tilde{R}$  which are configuration pairs acquired by same single action transitions from  $X$  and  $Y$  respectively. Therefore,  $\tilde{R}$  is a history preserving bisimulation, i.e.  $\mathcal{P} \approx_h \mathcal{Q}$ .

This proposition shows that clustered bisimulation equivalence  $\approx_{bb}$  can imply history preserving equivalence  $\approx_h$ .

Subsequently, we discuss the relationship under action refinement between clustered bisimulation equivalence and interleaving bisimulation equivalence.

**Proposition 5.4:** Let  $\mathcal{P}, \mathcal{Q} \in \mathcal{S}$ . Assume that all transitions in event structures  $\mathcal{P}$  and  $\mathcal{Q}$  be clustered action transitions,  $ref$  be a refinement function. If  $\mathcal{P} \approx_{cb} \mathcal{Q}$  then  $\mathcal{P} \approx_{ib} \mathcal{Q} \Rightarrow ref(\mathcal{P}) \approx_{ib} ref(\mathcal{Q})$ .

**Proof:** By proposition 5.3,  $\mathcal{P} \approx_h \mathcal{Q}$ . Also, by proposition 3.3,  $ref(\mathcal{P}) \approx_h ref(\mathcal{Q})$ . And, by proposition 3.4,  $\mathcal{P} \approx_{ib} \mathcal{Q} \Rightarrow ref(\mathcal{P}) \approx_{ib} ref(\mathcal{Q})$ . Therefore, we can obtain  $\mathcal{P} \approx_{ib} \mathcal{Q} \Rightarrow ref(\mathcal{P}) \approx_{ib} ref(\mathcal{Q})$ .

This proposition shows that if two event structures satisfy clustered bisimulation equivalence and their all transitions are clustered action transitions, then interleaving bisimulation equivalence also can hold under action refinement. Proposition 5.2 and proposition 5.4 extend proposition 3.5, which demonstrates that in the presence of concurrency interleaving equivalence is able to preserve under action refinement in given conditions.

## 6 Results and Discussion

The paper has demonstrated that

(1) In event structures without independency between events, interleaving trace equivalence and interleaving

bisimulation equivalence can preserve under action refinement.

(2) In event structures, if all transitions are clustered action transitions, then with clustered trace equivalence between event structures, we can reach that interleaving trace equivalence can preserve under action refinement; likewise, with clustered bisimulation equivalence between event structures, we can reach that interleaving bisimulation equivalence can preserve under action refinement. Therefore, we find a class of concurrent processes with specific properties, which enable interleaving equivalence to preserve under action refinement in the absence of constraint.

Our next work is to introduce the thought of clustered action transition into model checking so as to deal with states explosion problem [4, 7, 15] in the process of system verification.

## Acknowledgements

This Work is partly supported by Grants No. HCIC201306 of Guangxi HCIC lab Open Fund, the National Natural Science Foundation of China under Grant No. 11371003, the Natural Science Foundation of Guangxi under Grant No. 2011GXNSFA018154 and No. 2012GXNSFGA060003, and Science Computing and Intelligent Information Processing of GuangXi higher education key laboratory No. GXSCIP201201.

## References

- [1] Wu J 2001 Action refinement in timed LOTOS *Proc Of ASCM'01 World Scientific Publ* 183-92
- [2] Wang Y, Wu J, Jiang J 2007 Process algebra-symmetry and action refinement, *Chinese Science Press*06
- [3] Majster-Cederbaum M, Wu J 2003 Towards action refinement for true concurrent real time *Acta Informatica* **39** 1-47
- [4] Clarke E M, Grumberg O, Minea M, Peled D 1999 State space reduction using partial order techniques *STTT* **2**(3) 279-87
- [5] Darondeau P, Degano P 1993 Refinement of actions in event structures and casual trees *Theoretical Computer Science* **118** 21-48
- [6] Jiang J, Wu J 2005 Symmetry and autobisimulation *Proceedings of the 6th International Conference on Parallel and Distributed Computing, Applications and technologies IEEE Computer Society Press* 866-70
- [7] Jiang J, Wu J, Yan W 2005 Structural reductions in process algebra languages *Proceedings of the 11th Joint International Computer Conference World Scientific Publishing Co* 596-600
- [8] Degano P, Gorrieri R 1995 A causal operational semantics of action refinement *Information and Computation* **122** 97-119
- [9] van Glabbeek R J, Goltz U 1989 Equivalence notions for concurrent systems and refinement of actions *Mathematical Foundations of Computer Science Lecture Notes in Computer Science* **379** 237-48
- [10] van Glabbeek R J, Goltz U 1990 A deadlock-sensitive congruence for action refinement *Institut fuer Informatik Technische Universitaet Munchen SFB-Bericht* 342/23/90 A
- [11] van Glabbeek R J, Goltz U 2001 Refinement of actions and equivalence notions for concurrent systems *Acta Informatica* **37**(4/5) 229-327
- [12] Tang W, Wu J, Zheng D 2014 On fuzzy rough sets and their topological structures *Mathematical Problems in Engineering* 01
- [13] Fecher H, Majster-Cederbaum M, Wu J 2002 Refinement of actions in a real-time process algebra with a true concurrency model *Electronic Notes in Theoretical Computer Science* **70**(3) 620-40
- [14] Vogler W 1991 Bisimulation and action refinement *STACS'91 Lecture Note in Computer Science* **480** 309-21
- [15] Vogler W 1992 Modular construction and partial order semantics of Petri nets *Lecture Note in Computer Science* **625**
- [16] Huhn M 1996 Action refinement and property inheritance in systems of sequential agents *Concur'96 Lecture Notes in Computer Science* **1119** 639-54
- [17] Majster-Cederbaum M, Wu J, Yue H 2006 Refinement of actions for real-time concurrent systems with causal ambiguity *Acta Informatica* **42**(6/7) 389-418
- [18] Winskel G 1989 An Introduction to Event structures *Berlin Springer LNCS* **354** 364-97
- [19] Wu J, Fecher H 2004 Symmetric structure in logic programming *Journal of Computer Science and Technology* **19**(6) 803-11
- [20] Goltz U, Wehrheim H 1996 Modelling causality by dependency of actions in branching time semantics *Information Processing Letters* **59**(4) 179-84
- [21] Czaja I, van Glabbeek R J, Goltz U 1992 Interleaving semantics and action refinement with atomic choice *Advances in Petri Nets Lecture Notes in Computer Science* **609** 89-107
- [22] Gorrieri R, Rensink A 2001 Action Refinement *Bergstra J A Ponse A and Smolka S A editors Handbook of Process Algebra New York Elsevier Science* 1047-47

Authors	
	<p><b>Weidong Tang, born in February, 1968, Nanning China</b></p> <p><b>Current position, grades:</b> Teacher, Associate professor <b>University studies:</b> Computer Software and Theory <b>Scientific interest:</b> Symbolic computation, formal verification</p>
	<p><b>Jinzhao Wu, born in October, 1965, Nanning China</b></p> <p><b>Current position, grades:</b> Professor, Ph.D. <b>University studies:</b> Computer Software and Theory <b>Scientific interest:</b> Symbolic computation, automated reasoning, formal methods</p>
	<p><b>Meiling Liu, born in January, 1979, Nanning China</b></p> <p><b>Current position, grades:</b> Teacher, Lecturer <b>University studies:</b> Computer Software and Theory <b>Scientific interest:</b> Data Mining, formal verification</p>

# Research on growth opportunity and liquidity monitoring by mathematical optimization

**Xiani Yang<sup>1</sup>, Yaqin Lu<sup>2\*</sup>, Cunzhi Tian<sup>1</sup>**

<sup>1</sup>*Economic Research Center, Kunming University of Science and Technology, 650093, Kunming, Yunnan, China*

<sup>2</sup>*Economic Research Institute, Yunnan University of Finance and Economics, 650221, Kunming, Yunnan, China*

*Received 4 May 2014, www.tsi.lv*

---

## Abstract

Based on the assumption of variable-investment, this paper introduces growth opportunity into the model of liquidity needs (Tirole, 2006). Through the establishment of mathematical optimization model, we analyse the influence of growth opportunities on liquidity needs and liquidity investment decisions. Both of mathematical derivation and numerical simulation show that, the entrepreneur tends to overinvest in illiquid assets if the growth opportunity is small; otherwise, he will overhold of liquid assets. In addition, the agency costs due to information asymmetry may also affect the entrepreneur's decisions of liquid assets investment.

*Keywords:* Growth opportunity, liquidity monitoring, variable-investment model, mathematical optimization

---

## 1 Introduction

As the vitality of enterprises, liquidity is necessary to ensure the normal operation of enterprises. As ongoing entities, firms are concerned that they may in the future be deprived of the funds that would enable them to take advantage of exciting growth prospects, strengthen existing investments, or simply stay alive. Enterprises in the process of actual operation often face a liquidity shocks, and then the liquidity demand will happen. Liquidity is the most important part of the business, whether purchasing, production, sales and other business sectors, or investment in new projects, distribution of profits to shareholders, as well as the repayment of debt principal and interest, all of these require a lot of liquidity.

As the first one explained the meaning of mobility, Keynes believes that liquidity refers to how easy is it to convert asset into payment [1]. In the definition of enterprises' liquidity, however, the academia has yet formed a clear definition. The study of [2] points out that the liquidity of enterprises not only includes the solvency, but also contains the whole cash needs. However, the author of [3] holds that corporate liquidity is the total liquidity of each single asset. The research for liquidity assets describes enterprise liquidity as frequency of fund flows [4]. Liquidity risk is divided into market liquidity risk and financing liquidity risk [5]. The market liquidity risk can be further divided into two categories, exogenous and endogenous [6].

As the vitality of enterprises, liquidity is necessary to ensure the normal operation of enterprises. The occurrence of liquidity risk accidents tends to endanger the survival of the enterprise, and even spread to the

entire community. Unsurprisingly, liquidity planning is central to the practice of corporate finance and consumes a large fraction of chief financial officers' time. Studies for liquidity monitoring have never stopped.

There are lots of factors that affect liquidity demand and liquidity risk, and growth prospect is also one of them. The writer of [7] firstly puts forward the concept of growth opportunity, and he thinks growth opportunity refers to the part of enterprise's value depending on discretionary expenditures in the future. Growth opportunities, in other words, not only including traditional investment opportunities, also including discretionary spending, which can bring greater success for the enterprise [8]. Growth opportunities depend not only on the external environment, but also depend on the enterprise itself [9]. And different enterprises may own different growth opportunities [10].

The essence of liquidity supervision is liquidity risk management, and lots of scholars study the factors that influence the liquidity and risk management. The purpose of liquidity risk management is to seek anticipated cost trade-off in the ample liquidity and abundant liquidity [11]. The major factors that affect the liquidity are company size, growth opportunities, physical assets investment yields, the company's ability to create operating cash flow, working capital management efficiency as well as the debt ratio and so on [12]. Otherwise, maintaining high liquidity can improve the value of the inter-temporal investment options [13].

At present, research on liquidity monitoring is concentrated on definition, classification and measurement through qualitative analysis. Some scholars have discussed the impact and causes of liquidity monitoring by empirical analysis. However, few scholars

---

\*Corresponding author e-mail: xianiyang@126.com

have made theoretical study for liquidity monitoring, from the perspective of demand for liquidity.

On the basis of liquidity needs model [14-17], this paper analyses the impact of growth opportunities on liquidity needs. Illiquidity may affect the normal operation, make it impossible to repay maturing debt, and even endanger the survival of the enterprise; conversely, excess liquidity will reduce the profits. Therefore, we try to study the influence of growth opportunities on the firm's liquidity investment decisions by establishing and solving the mathematical optimization model. Ultimately, we want to prove the need for monitoring liquidity assets.

The result shows that the entrepreneur will overinvest in illiquid assets when the growth opportunity is relatively small. Conversely, the entrepreneurs will overboard liquid assets when the growth opportunity is relatively big. In addition, agency costs will also affect the entrepreneur's investment decisions on liquid assets. Therefore, in order to prevent entrepreneurs make some investment decision not conducive to the development of enterprises, the company's liquidity ratio must be formulated in loan agreement.

**2 Assumption**

This paper introduces growth prospects on the basis of the liquidity risk management model, in the contest of the variable-investment framework. At date 0 , the

entrepreneur has a project requiring fixed investment  $I$  , he initially has "assets"  $A$  and needs to borrow  $I - A$  from investors. At date 1 , the firm meets a new investment chance requiring an amount  $\rho I$  , where  $\rho$  is ex ante unknown and has cumulative distribution function  $F(\rho)$  with density  $f(\rho)$  on  $\rho \in [0, \infty)$  . The realization of  $\rho$  is learned at date 1 .

The probability of success  $p$  is affected by the effort degree of the entrepreneur, which is unobservable. Behaving yields probability  $p = p_H$  of success, and misbehaving results in probability  $p = p_L < p_H$  of success and private benefit  $BI > 0$  .

Let  $\Delta p = p_H - p_L > 0$  . If the firm does not reinvest  $\rho I$  , then it yields, at date 2 ,  $RI$  with probability  $p$  and 0 with probability  $1 - p$  . If the firm reinvests  $\rho I$  , then it yields, at date 2 ,  $RI$  with probability  $p + \tau$  and 0 with probability  $1 - (p + \tau)$  , where  $\tau > 0$  .

The investment has positive NPV. Both the entrepreneur and investors are risk neutral. The entrepreneur is protected by limited liability. Investors behave competitively in the sense that the loan, if any, makes zero profit. We summarize the timing in Figure 1:

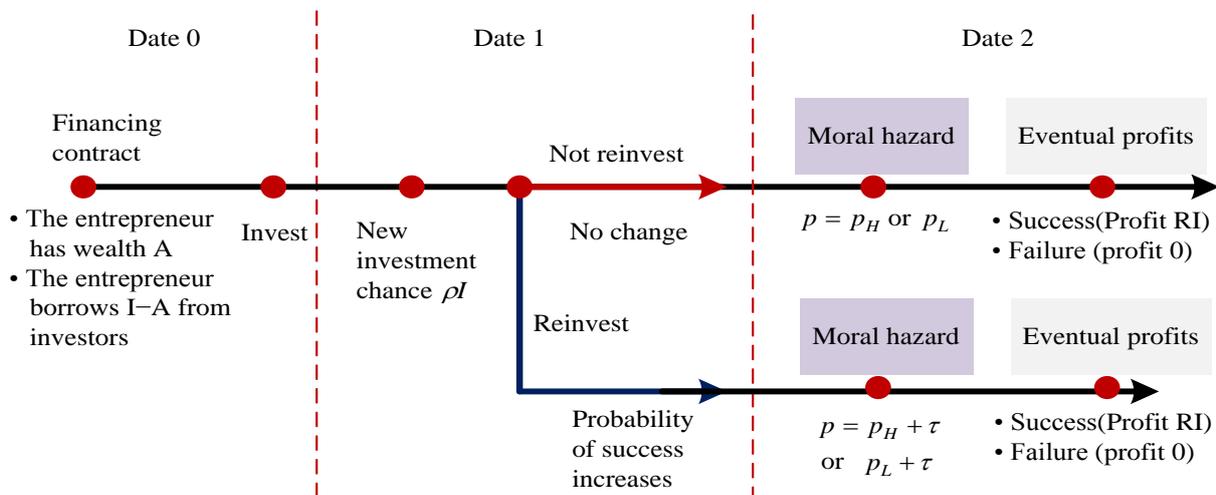


FIGURE 1 Figure of the timing

**3 Optimal models**

Suppose that the financing contract takes the following state-contingent form:  $\{I; \rho^c; (R_b, 0); (RI - R_b, 0)\}$ .

The contract specifies that  $\rho^c$  is a cut-off of reinvestment: only if  $\rho \leq \rho^c$  , the firm reinvests and the investment level is  $I$  . If the project success, the entrepreneur and investors get  $R_b$  and  $RI - R_b$  respectively; if the project fail, both of them get 0 . For

any cut-off of reinvestment  $\rho^c$  :

$$0 < \frac{[\tau F(\rho^c) + p_H]R - [1 + \int_0^{\rho^c} \rho f(\rho)]}{\tau F(\rho^c) + p_H} < \frac{B}{\Delta p}$$

The probability of reinvestment is  $\Pr ob\{\rho \leq \rho^c\} = F(\rho^c)$ .

So the optimization problem becomes

$$\left\{ \begin{array}{l} \max_{R_b, \rho^c, I} F(\rho^c)(p_H + \tau)R_b + \\ \quad [1 - F(\rho^c)]p_H R_b - A \\ \text{s.t. (a1)} F(\rho^c)(p_H + \tau)R_b \\ \quad \geq F(\rho^c)[(p_L + \tau)R_b + BI], \\ \text{(b1)} [1 - F(\rho^c)]p_H R_b \\ \quad \geq [1 - F(\rho^c)](p_L R_b + BI), \\ \text{(c1)} F(\rho^c)(p_H + \tau)(RI - R_b) + [1 - F(\rho^c)] \\ \quad \times p_H (RI - R_b) \geq I + \int_0^{\rho^c} \rho f(\rho) d\rho - A \end{array} \right. \quad (1)$$

(a1) is the objective function is the entrepreneur's utility, when the entrepreneur's incentive-compatibility constraint if the firm can come up with enough cash to reinvest. (b1) is the incentive-compatibility constraint if not, and both of them could be simplified as:  $R_b \geq BI/\Delta p$ . (c1) is the investors' individual-rationality constraint and it holds with equality. So the optimal model (1) will be simplified as:

$$\left\{ \begin{array}{l} \max_{R_b, \rho^c, I} \{F(\rho^c)(p_H + \tau) + [1 - F(\rho^c)]p_H\}RI \\ \quad - [I + \int_0^{\rho^c} \rho f(\rho) d\rho], \\ \text{s.t. (a2)} R_b \geq BI / \Delta p, \\ \text{(b2)} \{F(\rho^c)(p_H + \tau) + [1 - F(\rho^c)]p_H\} \\ \quad \times (RI - R_b) = I + \int_0^{\rho^c} \rho f(\rho) d\rho - A \end{array} \right. \quad (2)$$

The optimal solution of the model (2) can be got through the following three steps.

Firstly, get  $R_b$  for a given  $\rho^c$  and  $I$ . We may illustrate the "feasible contract set" of (2) by Figure 2, and it may be constituted by the shaded area OEF.

First of all, it is easy to get that the intercept of the line (b2) is  $A/[\tau F(\rho^c) + p_H]$  and the slope is:

$$k_1 = \frac{[\tau F(\rho^c) + p_H]R - [I + \int_0^{\rho^c} \rho f(\rho) d\rho]}{\tau F(\rho^c) + p_H}$$

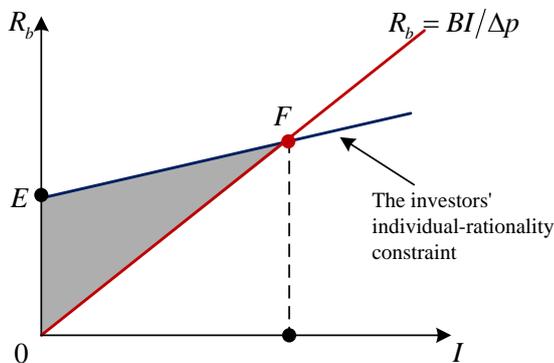


FIGURE 2 The solution for the model

The intercept of the line (b1) is 0 and the slope is  $B/\Delta p$ . Since  $0 < k_1 < B/\Delta p$ , the "feasible contract set" is not an empty set if  $A \geq 0$ . And then, it is the bigger the better for  $R_b$ . So the point F constitutes the optimal contract:  $R_b^* = BI/\Delta p$ .

Secondly, consider the optimal  $I$  for a given  $\rho^c$ . Actually, take  $R_b^*$  into the problem (2) and it can be further simplified as that

$$\left\{ \begin{array}{l} \max_{\rho^c, I} m(\rho^c)I \\ \text{s.t. } \{F(\rho^c)(p_H + \tau) + [1 - F(\rho^c)]p_H\} \\ \quad \times (RI - R_b) = I + \int_0^{\rho^c} \rho f(\rho) d\rho - A \end{array} \right. \quad (3)$$

where

$$m(\rho^c) = [p_H + \tau F(\rho^c)]R - [I + \int_0^{\rho^c} \rho f(\rho) d\rho],$$

$$I^*(\rho^c) = k(\rho^c)A$$

$$= \frac{A}{[1 + \int_0^{\rho^c} \rho f(\rho) d\rho] - [\tau F(\rho^c) + p_H](R - B/\Delta p)}$$

$$U_b(\rho^c) = m(\rho^c)k(\rho^c)A.$$

Finally, we could solve  $\rho^c$ .

**Proposition 1:** If  $\rho^{c*}$  is equal to the expected unit cost of effective investment,  $U_b(\rho^c)$  reaches its

$$\text{maximum, that is: } \rho^{c*} = c(\rho^{c*}) = \frac{1 + \int_0^{\rho^{c*}} \rho f(\rho) d\rho}{p_H / \tau + F(\rho^{c*})}.$$

*Proof:* In fact

$$U_b(\rho^c) = \frac{\{[p_H + \tau F(\rho^c)]R - [I + \int_0^{\rho^c} \rho f(\rho) d\rho]\}A}{[1 + \int_0^{\rho^c} \rho f(\rho) d\rho] - [\tau F(\rho^c) + p_H](R - B/\Delta p)} \cdot \\ = \frac{\tau R - c(\rho^c)}{c(\rho^c) - \tau(R - B/\Delta p)} A$$

So  $U_b(\rho^c)$  reaches its maximum means that

$$\min_{\rho^c} c(\rho^c) = [1 + \int_0^{\rho^c} \rho f(\rho) d\rho] / [p_H / \tau + F(\rho^c)].$$

The first-order condition is:

$$\rho^c f(\rho^c)[p_H / \tau + F(\rho^c)] - f(\rho^c) - f(\rho^c) \int_0^{\rho^c} \rho f(\rho) d\rho = 0$$

That is

$$\rho^{c*} \frac{p_H}{\tau} + \int_0^{\rho^{c*}} F(\rho) d\rho = 1. \quad (4)$$

Assume that  $Z = 1 - \rho^{c^*} (p_H / \tau) - \int_0^{\rho^{c^*}} F(\rho) d\rho = 0$ , then from (4) we can get that  $c(\rho^{c^*}) = \frac{Z + \rho^{c^*} [p_H / \tau + F(\rho^{c^*})]}{p_H / \tau + F(\rho^{c^*})} = \rho^{c^*}$ .

**Proposition 2:**  $\rho^{c^*}$  is increasing with the increase of  $\tau / p_H$ .

*Proof:* Because  $\rho^{c^*} = c(\rho^{c^*})$

$$\Leftrightarrow \rho^{c^*} F(\rho^{c^*}) + \rho^{c^*} (p_H / \tau) = 1 + \int_0^{\rho^{c^*}} \rho f(\rho) d\rho.$$

Solving the partial derivative of  $\rho^{c^*}$  with respect to  $\tau / p_H$  on both sides of the above formula:

$$[\rho^{c^*} f(\rho^{c^*}) + F(\rho^{c^*})] \frac{\partial \rho^{c^*}}{\partial (p_H / \tau)} + \rho^{c^*} + \frac{p_H}{\tau} \frac{\partial \rho^{c^*}}{\partial (p_H / \tau)} = \rho^{c^*} f(\rho^{c^*}) \frac{\partial \rho^{c^*}}{\partial (p_H / \tau)}$$

$$\text{That means } \rho^{c^*} = -[F(\rho^{c^*}) + \frac{p_H}{\tau}] \frac{\partial \rho^{c^*}}{\partial (p_H / \tau)}.$$

It is easy to get that  $\rho^{c^*} > 0$ ,  $F(\rho^{c^*}) + p_H / \tau > 0$ , and then  $\frac{\partial \rho^{c^*}}{\partial (p_H / \tau)} < 0 \Leftrightarrow \frac{\partial \rho^{c^*}}{\partial (\tau / p_H)} > 0$ .

#### 4 Monitoring Overinvestment in illiquid assets

In most instances, loan agreements do not focus solely on the borrower's solvency, but also strictly constrain the borrower's liquidity. For example, many loan agreements require that the borrower maintain a minimum level of working capital. It is not a priori clear why this is so. Let us bring one answer to this puzzle, and show that it may be optimal for lenders to simultaneously impose gearing and liquidity ratios.

In the absence of a liquidity requirement, suppose that the borrower invests the full  $I^* (1 + \rho^{c^*}) = I'$  in illiquid assets; despite the lack of cash left for reinvestment, the project will often be continued. An interesting issue relates to whether the investors should renegotiate the borrower's compensation scheme so as to account for the unexpectedly high scale of operations. The answer to this question depends on the way the managerial compensation contract was initially drawn.

✓ If the borrower owns a share in the firm's final profit, then managerial compensation scales up with investment, and the initial incentive scheme remains incentive compatible as investment increases and is not renegotiated by lenders to account for the altered firm size.

✓ If the entrepreneur gets a fixed reward for "success"; because the private benefit scales up with investment, the initial incentive scheme is then no longer

incentive compatible. Lenders then offer to increase the borrower's reward in the case of "success" and so they raise the borrower's payoff in the case of success to  $BI' / \Delta p$  in order to make sure the borrower behaves.

It should be noted that the high payoff can be achieved only if investors agree that. Now suppose that investor may get more when the entrepreneur behaving than misbehaving, the entrepreneur's high payoff will not be cancelled. Then the following proposition can be drawn.

**Proposition 3:** The entrepreneur's high payoff will not be cancelled if  $\tau$ ,  $p_L$  and  $B$  are relatively small,  $p_H$  and  $R$  are relatively big.

*Proof:* In fact, as long as

$$\begin{aligned} & \{F(\rho^{c^*})(p_H + \tau) + [1 - F(\rho^{c^*})]p_H\}(R - B / \Delta P)I' \\ & \geq \{F(\rho^{c^*})(p_H + \tau) + [1 - F(\rho^{c^*})]p_L\}(RI' - B / \Delta PI^*) \\ & (\Delta P)(R - B / \Delta P) \\ & \geq [p_L + \tau F(\rho^{c^*})](B / \Delta P)[\rho^{c^*} / (1 + \rho^{c^*})]. \end{aligned} \tag{5}$$

This means  $Y_1 \equiv (\Delta P)(R - B / \Delta P) -$

$$[p_L + \tau F(\rho^{c^*})](B / \Delta P)[\rho^{c^*} / (1 + \rho^{c^*})] \geq 0.$$

Then the high payoff may not be cancelled.  $\rho^{c^*}$  will reduce as  $\tau$  decreases and  $p_H$  increases, and  $\rho^{c^*} / (1 + \rho^{c^*})$  will also reduce. Furthermore, with the decreasing of  $p_L$  and  $B$  and the increasing of  $R$ , the left side of (5) will increase and the right side will decrease, that means (5) will be set up.

**Proposition 4:** The entrepreneur will overinvestment in illiquid assets if  $\tau$ ,  $p_L$  and  $B$  are relatively small,  $p_H$  and  $R$  are relatively big.

*Proof:* Indeed, the borrower, who, regardless of the design of her initial compensation contract, receives expected rent  $p_H B / \Delta p$  per unit of illiquid assets, prefers investing  $I'$  rather than  $I^*$  if:

$$\begin{aligned} & \{F(\rho^{c^*})(p_H + \tau) + [1 - F(\rho^{c^*})]p_H\}(BI^* / \Delta P) \\ & < \{F(\hat{\rho}_0)(p_H + \tau) + [1 - F(\hat{\rho}_0)]p_L\}(BI' / \Delta P), \\ & F(\rho^{c^*}) - F(\hat{\rho}_0) < [(p_H / \tau)F(\hat{\rho}_0)]\rho^{c^*}. \end{aligned} \tag{6}$$

That means

$$Y_2 \equiv F(\rho^{c^*}) - F(\hat{\rho}_0) - [(p_H / \tau)F(\hat{\rho}_0)]\rho^{c^*} < 0 \text{ and if}$$

$\tau / p_H$  is relatively small,  $p_H / \tau$  will be relatively big, then (6) will be set up. Furthermore, with the decreasing

of  $p_L$  and  $B$  and the increasing of  $R$ ,  $\hat{\rho}_0$  will increase and (6) will be set up more easily.

**Corollary 1:** The overinvestment in illiquid assets will be achieved if  $\tau$ ,  $p_L$  and  $B$  are relatively small,  $p_H$  and  $R$  are relatively big.

From proposition 3 and proposition 4, we may get that the entrepreneur will overinvestment in illiquid assets and the investors will not cancel his high payoff if  $\tau$ ,  $p_L$  and  $B$  are relatively small,  $p_H$  and  $R$  are relatively big.

Alternatively, the entrepreneur may have been granted in the initial agreement a fixed reward for “success”. Because the private benefit scales up with investment, the initial incentive scheme is then no longer incentive compatible. Because the borrower is then strictly better off overinvesting, the lender should rationally anticipate to lose money overall (That the lender loses money results from the facts that the borrower deviates from investment  $I^*$  to obtain more than  $U_b$ , and that  $U_b$  is the maximum utility for the borrower consistent with a nonnegative profit for the lender.). Hence, the optimal investment lever is not  $I^*$  and the rationale for a liquidity requirement.

**5 Monitoring overhoarding of liquid assets**

As mentioned earlier, lenders may also need to verify that the borrower does not underinvest in illiquid assets in order to overinsure against liquidity shocks. We select a specific set of assumptions for the sole purpose of illustrating a possible incentive to underinvest in illiquid assets.

- ✓ The borrower can use the excess liquidity in order to withstand the liquidity shock;
- ✓ The borrower and investors receive shares of the date 2 profit with share  $(B/\Delta p)R$  held by the borrower and share  $(R-B/\Delta p)R$  held by the investors;
- ✓ Unused liquidity is returned to investors and the firm only issue stock.

**Proposition 5:** The entrepreneur will overhoard of liquid assets if  $\tau/p_H$  is relatively big.

*Proof:* Suppose further that the borrower invests  $I < I^*$  in illiquid assets and thus hoards liquidity equal to  $\rho^* I^* + (I^* - I)$ . She can then withstand liquidity shocks

$$\rho I \leq \rho^* I^* + (I^* - I)$$

$\rho$  such that:

$$\Leftrightarrow \rho \leq [\rho^* I^* + (I^* - I)] / I \equiv \bar{\rho}$$

Letting  $\varepsilon = (I^* - I) / I$ , and using the all-equity-firm assumption, the borrower prefers to underinvest only if:

$$\begin{aligned} & \{F(\bar{\rho})(p_H + \tau) + [1 - F(\bar{\rho})]p_H\} (BI^* / \Delta P) \\ & \geq \{F(\rho^*)(p_H + \tau) + [1 - F(\rho^*)]p_H\} (BI^* / \Delta P) \end{aligned} \tag{7}$$

For small underinvestments, that is  $\varepsilon \rightarrow 0$

$$\begin{aligned} & \{F[\rho^* + (1 + \rho^*)\varepsilon] - F(\rho^*)\} \\ & = f(\rho^*)[\rho^* + (1 + \rho^*)\varepsilon - \rho^*] \\ & = f(\rho^*)(1 + \rho^*)\varepsilon \end{aligned}$$

and (7) will be like that

$$\begin{aligned} & \varepsilon f(\rho^*)(1 + \rho^*) \geq [F(\rho^*) + p_H / \tau] \varepsilon \\ & \Leftrightarrow f(\rho^*)(1 + \rho^*) \geq [F(\rho^*) + p_H / \tau] \end{aligned} \tag{8}$$

That means  $Y_3 \equiv f(\rho^*)(1 + \rho^*) \geq [F(\rho^*) + p_H / \tau] \geq 0$  and if  $\tau/p_H$  is relatively small,  $p_H/\tau$  will be relatively big, then (8) will be set up.

Roughly, if liquidity shocks around the threshold  $\rho^{c*}$  are quite likely, hoarding a bit more liquidity than allowed is privately profitable for the borrower. The borrower would always prefer underinvesting to investing  $I^*$  if she had a fixed claim (namely,  $BI^* / \Delta p$  in the case of success).

**6 Numerical simulation and conclusions**

As mentioned earlier, lenders may also need to verify that the borrower does not underinvest in illiquid assets in order to overinsure against liquidity shocks. We select a specific set of assumptions for the sole purpose of illustrating a possible incentive to underinvest in illiquid assets.

Now, we make some numerical calculations on the theoretical results. Table 1 shows the influence of  $\tau$  on financing. Where, basic parameters are  $R = 3$ ,  $B = 0.8$ ,  $p_L = 0.07$ ,  $p_H = 0.36$ ,  $\rho \in U[0,1]$ . From table 1, it can be get that  $\rho^{c*}$  increases with the increasing of  $\tau$  and the entrepreneur will overinvestment in illiquid assets if  $\tau$  is relatively small. (In all of the following tables, “Ent.” means the entrepreneur and “Inv.” Means investors).

Table 2 shows the influence of  $\tau/p_H$  on financing. Where, basic parameters are  $R = 3$ ,  $B = 0.8$ ,  $p_L = 0.07$ ,  $\rho \in U[0,1]$ . From table 2, we may get that  $\rho^{c*}$  is increasing with the increase of  $\tau/p_H$  and the entrepreneur’s incentive to overboard of liquid assets increases with the increasing of  $\tau/p_H$ .

Table 3 shows the influence of  $p_L$  on financing. Where, basic parameters are  $R = 3$ ,  $B = 0.8$ ,

$p_H = 0.38, \tau = 0.55, \rho \in U[0,1]$ . From Table 3 some endings may be got: (i)  $\rho^{c*}$  has nothing to do with  $p_L$ ; (ii) The entrepreneur's incentive to overinvest in illiquid assets decreases with the increasing of  $p_L$ ; (iii) Investors will gradually become opposed to their high salaries with the increasing of  $p_L$ . Therefore, overinvestment in illiquid assets would be achieved only if  $p_L$  is comparatively small.

Table 4 shows the influence of  $B$  on financing. Where, basic parameters are  $\tau = 0.5, p_L = 0.05, R = 3, p_H = 0.4, \rho \in U[0,1]$ . From table 3 we may get the following conclusions:

- (i)  $\rho^{c*}$  has nothing to do with  $B$ ;

- (ii) The entrepreneur's incentive to overinvest in illiquid assets decreases with the increasing of  $B$ ;
- (iii) Investors will gradually become opposed to their high payoff with the increasing of  $B$ . Therefore, overinvestment in illiquid assets would be achieved only if  $B$  is relatively small.

Generally speaking, this paper analysed the influence of growth opportunities on liquidity needs and liquidity investment decisions. Firstly, the "first-best cutoff" of reinvestment increases with the increasing of growth opportunity. Secondly, with the increase of growth opportunities, the entrepreneur's incentive to overinvest of illiquid assets decreases and investors will gradually become against the high payoff. As long as one participant does not agree, overinvesting will not happen. In other words, overinvesting of illiquid assets becomes possible only if the growth opportunity is small.

TABLE 1 The growth opportunity  $\tau$  and overinvesting of illiquid assets

$\tau$	$\rho^{c*}$	$I^*$	$Y_1$	Ent.	$Y_2$	Inv.	Overinvest
0.01	0.03	1.09	-0.97	√	0.06	√	√
0.06	0.16	1.08	-0.84	√	0.04	√	√
0.11	0.29	1.05	-0.70	√	0.01	√	√
0.16	0.41	1.02	-0.56	√	-0.04	×	×
0.21	0.51	0.98	-0.44	√	-0.09	×	×
0.26	0.59	0.95	-0.33	√	-0.16	×	×
0.31	0.67	0.92	-0.23	√	-0.24	×	×
0.36	0.73	0.89	-0.15	√	-0.32	×	×
0.41	0.79	0.87	-0.08	√	-0.41	×	×
0.46	0.83	0.86	-0.02	√	-0.50	×	×
0.51	0.87	0.84	0.03	×	-0.59	×	×
0.56	0.91	0.83	0.07	×	-0.69	×	×

TABLE 2 The ratio of growth opportunity  $\tau/p_H$  and overhoarding of liquid assets

$p_H$	$\tau$	$\rho^{c*}$	$I^*$	$Y_3$	Overhoard
0.36	0.21	0.51	0.98	-0.71	×
0.36	0.26	0.59	0.95	-0.38	×
0.36	0.31	0.67	0.92	-0.16	×
0.36	0.36	0.73	0.89	0.00	√
0.36	0.41	0.79	0.87	0.12	√
0.36	0.46	0.83	0.86	0.22	√
0.44	0.40	0.69	1.57	-0.10	×
0.42	0.40	0.71	1.33	-0.05	×
0.40	0.40	0.73	1.15	0.00	√
0.38	0.40	0.75	1.00	0.05	√
0.36	0.40	0.78	0.88	0.10	√
0.34	0.40	0.80	0.77	0.15	√

TABLE 3 The success probability of misbehaving  $p_L$  and overinvesting of illiquid assets

$p_L$	$\rho^{c*}$	$I^*$	$I'$	$Y_2$	Ent.	$Y_1$	Inv.	Overinvest
0.01	0.883	5.085	9.575	-1.155	√	0.133	√	√
0.03	0.883	3.612	6.801	-1.059	√	0.035	√	√
0.05	0.883	2.726	5.134	-0.951	√	-0.067	×	×
0.07	0.883	2.135	4.021	-0.830	√	-0.174	×	×
0.09	0.883	1.713	3.225	-0.691	√	-0.289	×	×
0.11	0.883	1.396	2.628	-0.533	√	-0.411	×	×
0.13	0.883	1.149	2.163	-0.348	√	-0.543	×	×
0.15	0.883	0.951	1.791	-0.132	√	-0.688	×	×
0.17	0.883	0.790	1.487	0.125	×	-0.849	×	×
0.19	0.883	0.655	1.234	0.436	×	-1.031	×	×

TABLE 4 The private benefit  $B$  and overinvesting of illiquid assets

$B$	$\rho^{**}$	$I^*$	$I'$	$Y_2$	Ent.	$Y_1$	Inv.	Overinvest
0.60	0.825	3.383	6.173	-1.008	√	0.092	√	√
0.65	0.825	2.429	4.432	-0.878	√	0.012	√	√
0.70	0.825	1.895	3.458	-0.747	√	-0.068	×	×
0.75	0.825	1.553	2.834	-0.617	√	-0.148	×	×
0.80	0.825	1.316	2.401	-0.487	√	-0.228	×	×
0.85	0.825	1.142	2.083	-0.356	√	-0.308	×	×
0.90	0.825	1.008	1.840	-0.226	√	-0.387	×	×
0.95	0.825	0.902	1.647	-0.096	√	-0.467	×	×
1.00	0.825	0.817	1.491	0.035	×	-0.547	×	×
1.04	0.825	0.759	1.386	0.139	×	-0.611	×	×

Finally, if the growth opportunity is big, the entrepreneur tries to overhold of liquid assets. In addition, agency costs, which may be measured by  $B/\Delta p$  will affect the entrepreneur's investment decisions on liquid assets. In order to prevent the entrepreneur to make wrong investment decisions, it is optimal for lenders to simultaneously impose gearing (leverage) and liquidity ratios.

**Acknowledgments**

The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper. The authors acknowledge the financial support of the project “strategic trading behaviour of institutional investors and stock price volatility”, project No. 2011FZ016

**References**

[1] Keynes J M 1936 *The General Theory of Employment, Interest, and Money* London: Interscience

[2] Liu R J 1998 Liquidity Risk Management of Enterprises *Foreign Trade Accounting* 1 6-7

[3] Zhang J R 1999 *The Theory of Assets Realization* Dalian: Dongbei University of Finance and Economics Press Ch1 (in Chinese)

[4] Zhang J Q, Li X 2008 *Systems Engineering Theory and Practice* 28(6) 14-21

[5] Jorion P 2002 *Journal of Risk* 5 75-96

[6] Bangia A F, Diebold T, Schuerman J 2002 *Risk Management: State of the Art* 8 3-13

[7] Myers S 1977 *Journal of Financial Economics* 5(2) 147-75

[8] Mason S P and Merton R C 1985 *The Role of Contingent Claims Analysis in Corporate Finance* Massachusetts: Harvard University Press

[9] Kester W C 1986 *Handbook of Corporate Finance* 5(2) 5-35

[10] Christie A 1989 Equity Risk, the Opportunity Set, Production Costs and Debt *University of Rochester*

[11] Wang Y 2012 The Influence Factors of Corporate Liquidity Analysis *Shandong Textile Economy* 3 15 (in Chinese)

[12] Yang H C 2007 Analysis of Factors That Influence Holding Liquid Assets for Listed Companies in China *Economist* 1 58 (in Chinese)

[13] Lian Y J, Peng F Y, Su Z 2010 *Financial Research* 10 158-71

[14] Holmström B, Tirole J 1998 *Journal of Political Economy* 106(1) 1-40

[15] Holmström B, Tirole J 2000 Liquidity and Risk Management *Journal of Money, Credit and Banking* 32(3) 295-319

[16] Holmström B, Tirole J 2002 Domestic and International Supply of Liquidity *American Economic Review* 92(2) 42-5

[17] Tirole J 2006 *The Theory of Corporate Finance* Princeton: Princeton University Press

**Authors**

	<p><b>Xiani Yang, born in 1988, Jincheng, Shanxi, China</b></p> <p><b>Current position:</b> Xiani Yang is a Ph.D. student in School of Management and Economics at Kunming University of Science and Technology in China</p> <p><b>Research interests:</b> corporate finance and mathematical economics</p>
	<p><b>Yaqin Lu, born in 1972, Dayao, Yunnan, China</b></p> <p><b>Current position:</b> Associate Professor in Economic Research Institute at Yunnan University of Finance and Economics</p> <p><b>Research interests:</b> international trade theory and inter-regional trade theory</p>
	<p><b>Cunzhi Tian, born in 1969, Heqing, Yunnan, China</b></p> <p><b>Current position:</b> Professor in Economic Research Center at Kunming University of Science and Technology of China.</p> <p><b>University study:</b> Ph.D. degree in Economics from Nankai University in 2001, Tianjin, China. Dr., postdoctoral research in applied mathematics at Yunnan University, China, the postdoctoral in finance at Shanghai University of Finance and Economics, China</p> <p><b>Research interests:</b> security market microstructure theory, financial engineering and corporate finance</p>

# A computer-assisted analysis of literary texts: a sample study

I Keshabyan-Ivanova\*

Departamento de Filología Inglesa, Universidad de Murcia, Facultad de Economía y Empresa Campus de Espinardo, 30100 Murcia. España (Spain)

Received 1 March 2014, www.tsi.lv

## Abstract

The overall aim pursued in this work is to demonstrate how quantitative data and a range of different corpus-based analytical techniques can be used in assessing an author's literary originality in relation to his texts' structures and meanings. With this in mind, the present study provides a sample of quantitative analysis of the two literary texts –Shakespeare's *Hamlet* [1685] and Sumarokov's *Gamlet* [1787]. Prior research has explored *Hamlet* and *Gamlet* in terms of historical, philosophical, language-based, etc. approaches that have existed to date. Taking into consideration the aforementioned visions of both plays, a special perspective on *Hamlet* and *Gamlet* is adopted herein. Given the importance assigned to computer-assisted analysis of literary texts, the current study is based on the idea that the texts under examination contain a certain number of particular characters that are distributed in a special way within and among the acts and intervene with a particular frequency specified by the authors. To achieve this aim, the texts are closely read and, then, computational and quantitative resources are applied. In general, the relevant findings unveil substantial structural deviations of the presence and interventions of all main characters, leading to noticeable diversions in the role and weight assigned by the authors to them per different acts inter-plays.

*Keywords:* Corpus-based, computational, quantitative, presence, interventions

## 1 Introduction

Corpus-based analytical techniques and specific ways in which corpus analysis has been applied to the study of literature have become more widespread over the recent decades. Interestingly, the works of such scholars as Stubbs [2005], Wynne [2006], Biber [2011], Johnson [2011], etc. demonstrate that it is increasingly becoming possible to test empirically claims about the language of literature, to search for and provide evidence from texts, to establish the norms of literary and non-literary style, and to have in-depth insights into the texts' structures and meanings.

The present research addresses a fundamental question concerning how quantitative data and a range of different computational and quantitative tools can be used in assessing an author's literary originality in relation to his texts' structures and meanings. It should not only benefit research on computer-assisted analysis of literary texts but also be of interest more generally to scholars of translation and comparative literature, leading to a more far-reaching understanding of many aspects of literature. At the same time, it should be noted that translation *stricto sensu* is not the main topic of the present investigation.

The overall aim pursued in this paper is to provide an example of quantitative analysis of the two literary texts – the Fourth Folio Edition of *The Tragedy of Hamlet Prince of Denmark* [1685] by Shakespeare and the English translated version of *Gamlet* [1787] by

Sumarokov, rendered into English by Richard Fortune in 1970.

For ease of reference, the abbreviations are used instead of the complete titles of the plays. For example, *Hamlet* or SH for Shakespeare's text, SG-R for Sumarokov's text, whereas *Gamlet* or SG for the English translation of the Russian text. However, one should keep in mind that in this research SG-R and SG are interchangeable, even though the general parameters of structural distinctions are explored between SH and SG and not between SH and SG-R. The selected texts are shown in Table 1.

TABLE 1 Texts selected for the analysis

Genre	Author	Title	Abbreviation
Drama	Shakespeare	<i>The Tragedy of Hamlet Prince of Denmark</i> (1685), the Fourth Folio Edition	SH
	Sumarokov	<i>Gamlet</i> (1787), in Russian (for reference) <i>Hamlet</i> (1970), in English	SG-R SG

The rationale behind my selection of Sumarokov's *Gamlet* for the current study is based on the fact that it was the first appearance of any of Shakespeare's plays in Russian culture, literature and theatre, although Shakespeare's name was nowhere mentioned [Lang 1948: 67] and the author himself denied any resemblance to Shakespeare's tragedy, 'apart from the monologue at the end of the third act and Claudius' falling down on his knees' [in Levitt 1994: 320]. The Fourth Folio Edition of *The Tragedy of Hamlet Prince of Denmark* [1685] was

\*Corresponding author e-mail: kirina@um.es

selected for the comparison due to the fact that Sumarokov may have been acquainted with this edition before writing his *Gamlet* and, thus, may have been influenced by it [Levitt 1994: 322].

Prior research has analysed *Hamlet* and *Gamlet* in terms of historical, philosophical, language-based, etc. approaches that have existed to date. With respect to *Hamlet*, Bradley [1904], Wilson [1959], Johnson [1960], Eagleton [1986], Kermode [2000], to name but few, have provided studies by using previously mentioned perspectives. With regard to *Gamlet*, Tynianov [1929], Lang [1948], Billington [1970], Levitt [1994 and Gukovskii [2003], among others, have contributed relevant works based on the above-mentioned approaches.

Lately, original scholarly contributions within Shakespeare studies have appeared. Murphy [2007: 67] suggests that a popular interest in Shakespeare has been paralleled in the last decade by a rise of computer-assisted textual analyses of his plays, specifically in the field of literary stylistics, although met with resistance by some literary critics such as Fish [1996], Louw [1997], etc. From within the field of corpus linguistics, Stubbs [2005: 22] calls to combine the findings of corpus stylistics with close reading of texts.

Recent corpus-based studies on Shakespeare have examined distinct aspects of his plays such as characterisation in *Romeo and Juliet* [Culpeper 2002], the morpho-syntactic variability of the second person pronouns in the Shakespeare Corpus [Busse 2002], key semantic domains and metaphor in love tragedies and love comedies [Archer, Culpeper and Rayson 2009], the rhetoric of suicide in *Hamlet* [Anderson and Crossley 2011], etc.

Having taken the aforementioned visions of both plays into consideration, a special perspective on *Hamlet* and *Gamlet* has been adopted herein. Given the importance assigned to computer-assisted analysis of literary texts, the current study is based on the idea that the analysed texts contain a certain number of particular characters that are distributed in a special way within and among the acts and intervene with a particular frequency specified by the authors. Hence, to achieve this aim, the texts are closely read and, then, computational and quantitative resources are applied.

To identify the dimensions of structural divergences that are especially characteristic of *Hamlet* and *Gamlet*, this investigation focuses on those aspects of the plays that could be easily located, extracted, quantified and computerized –in other words, on the distribution patterns of the presence and interventions of all main characters per distinct acts inter-plays. In so doing, the researcher seeks to reveal probable commonalities and/or deviations in the weight the authors put on all main characters within the plays that have led Sumarokov to introduce substantial changes into the structure of his play compared to Shakespeare's original play *Hamlet*.

The two phases of the present research include the analysis and discussion of the data associated with the distribution patterns of the presence and total interventions of all main characters, namely Hamlet, Claudius, Polonius, Gertrude and Ophelia, per different acts inter-plays.

## 2 Methodology

Two variables (including presence and intervention categories) are chosen in the current work under the criteria that they are quantitative and require certain computational tools. Each text is analysed with respect to the occurrences of these features that are quantified. The quantification of presence and intervention variables is carried out manually by exploring the two text files. After, the extracted data is computerised, tabulated (intra-play), cross-tabulated (inter-plays) and presented in tables and figures. The tools used for the computational quantification and presentation of the data in tables and figures are SPSS V.15 and Excel (Office 2007).

For the purposes of analysis, the data shown as a percentage throughout various phases of the present investigation are considered more valid than the data given in figures.

However, the aim is to normalise the data quantitatively in order to provide more precise identification and comparison of the general trends employed by the two playwrights in relation to the distribution patterns of the presence and interventions of all main characters per distinct acts inter-plays.

The instrument used for the standardisation of the data and their presentation in figures is Pearson's Correlation Test [Altman 1991: 285-288]. Correlation is a kind of technique that summarises the strength of the connection between two variables. For example, in this study, the presence and intervention variables in one text (SG) are compared with the presence and intervention variables in another text (SH) –in other words, there is one variable in each text separately that is compared between two texts.

The general requirement for Pearson's correlation coefficient is the observation of the two variables, which are measured on an interval or ratio scale and the calculation is based on the actual values [Altman 1991: 285-288]. In the present research, Pearson's Correlation Test computes the quantitative correlation between the presence and intervention variables per act inter-plays.

## 3 Findings and discussion

### 3.1 PRESENCE VARIABLES

Table 2 and Figure 1 incorporate the data linked to the distribution patterns of the presence of all main characters per act intra-play and inter-plays. However, as has already been commented in section 2, only the data

presented as a percentage inter-plays are analysed and discussed.

TABLE 2 Distribution Patterns of the Presence of All Main Characters per Act

Act	SH	SH %	SG	SG %
	Presence	Presence	Presence	Presence
I	5	22.73	2	13.33
II	5	22.73	3	20.00
III	5	22.73	3	20.00
IV	4	18.18	3	20.00
V	3	13.64	4	26.67

Comparing these data, the following diversions can be observed: the distribution pattern varies a little in acts II, III and IV inter-plays, whereas, in acts I and V, the difference is the greatest as in act I it equals 22.73 % in SH against 13.33 % in SG and, in act V, 13.64 % in SH against 26.67 % in SG.

As a result, the difference in the distribution is not statistically significant inter-plays ( $\chi^2 = 1.293$ ;  $df = 4$ ;  $p = 0.862$ ), which is a clear sign of divergence.

The data in Figure 1 display a probable quantitative correlation between the distribution patterns of the presence of all main characters per act inter-plays. These data show that the patterns of the presence of all main characters are negatively correlated per act inter-plays. The statistically non-significant correlation ( $\rho = -0.790$ ;  $df (8)$ ;  $p = 0.111$ ) might indicate distinct distribution patterns.

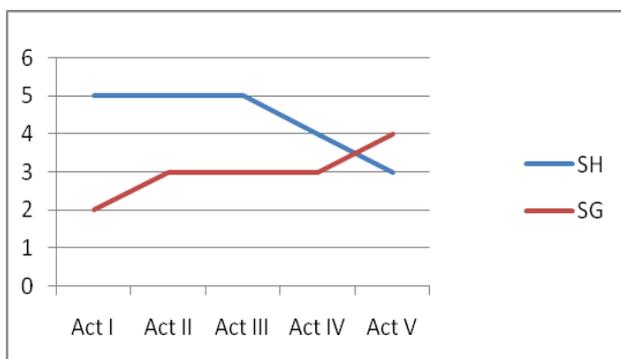


FIGURE 1 Quantitative Correlation between the Distribution Patterns of the Presence of All Main Characters per Act

It should be observed that Figure 1 clearly illustrates that the line, which corresponds to SH, remains in the same position in acts I-III as Shakespeare distributes all main characters equally in these acts whilst in acts IV and V it drops dramatically due to the reduction of the number of all main characters. Sumarokov behaves differently as the line goes up from act I to act II, remains in the same position in acts II-IV and again goes up in act

V. Even though the lines cross in act V inter-plays, the movement is downward in SH as opposed to the upward movement in SG.

As a result, the data in Table 2 and Figure 1 possibly demonstrate that both playwrights follow diverse distribution patterns of all main characters per act inter-plays, specifically in acts I and V.

### 3.2 INTERVENTION VARIABLES

Table 3 and Figures 2 and 3 comprise the data linked to the distribution patterns of the total interventions of all main characters per act and per full text intra-play and inter-plays.

TABLE 3 Distribution Patterns of the Total Interventions of All Main Characters per Act and per Full Text

Act	Frequency of Interventions	Percentage of Interventions	Frequency of Interventions	Percentage of Interventions
	SH	SH %	SG	SG %
I	100	9.17	32	16.67
II	145	13.30	22	11.46
III	197	18.07	47	24.48
IV	98	8.99	22	11.46
V	111	10.18	37	19.27
I-V	651	59.72	160	83.33

Comparing the data expressed as a percentage inter-plays, it can be seen that the frequency of total interventions of all main characters per act is higher in SH than in SG. As can be appreciated in Table 3 and Figure 2, the frequency of occurrence of total interventions is significantly higher in SH, although the percentage of total interventions is greater in SG, especially in acts I, III and V. For example, the difference is the highest in acts I and V as it equals 9.17 % against 16.67 % and 10.18 % against 19.27 % in SH versus SG, respectively. However, the distribution pattern is more or less alike in acts II and IV inter-plays.

Although the divergence in the distribution is statistically significant inter-plays ( $\chi^2 = 9.407$ ;  $df = 4$ ;  $p = 0.049$ ), which might indicate similar distribution patterns, there appears to be a partial dissimilarity, particularly in acts I and V.

To identify a possible quantitative correlation between the distribution patterns of the total interventions of all main characters per act inter-plays, Figure 3 is generated. According to the data shown in this figure, the quantitative correlation between the distribution patterns of the total interventions of all main characters is statistically not significant per act inter-plays ( $p = 0.608$ ;  $df (8)$ ;  $p = 0.276$ ); a clear sign of a distinct inter-play behaviour.

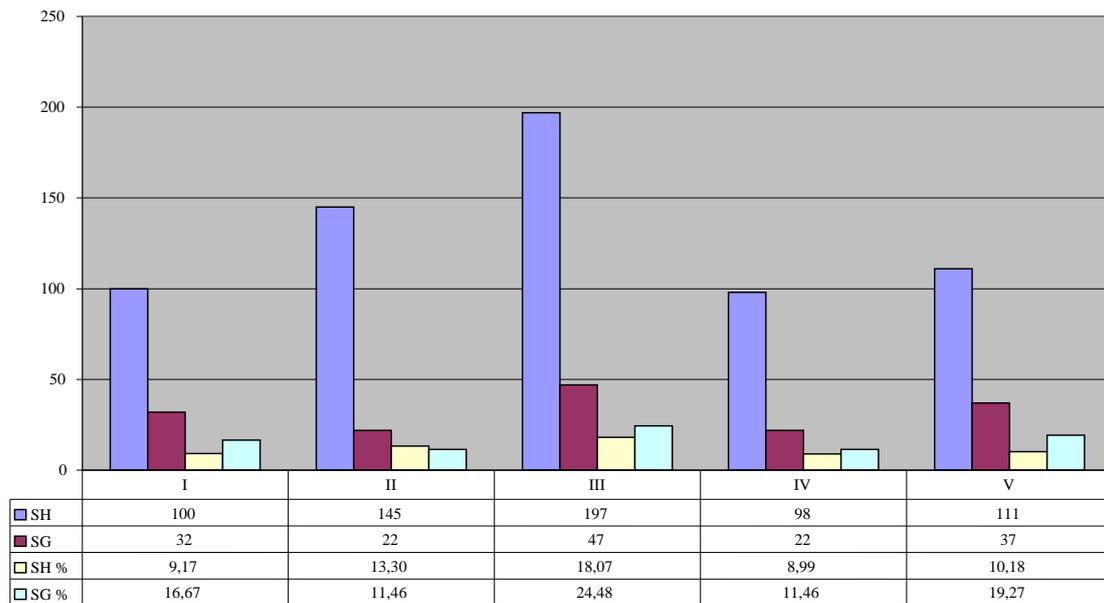


FIGURE 2 Distribution Patterns of the Total Interventions of All Main Characters per Act

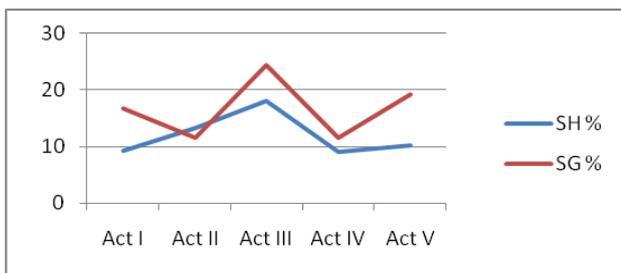


FIGURE 3 Quantitative Correlation between the Distribution Patterns of the Total Interventions of All Main Characters per Act

In fact, the line that displays the total interventions of all main characters in SH rises from act I to act II, goes up considerably in act III, falls in act IV and rises a little in act V. In SG, the line falls and rises more strikingly. Moreover, in act III, the upward movement is more significant in SG than in SH.

As a result, Shakespeare and Sumarokov apparently follow dissimilar distribution patterns of the total interventions of all main characters in all acts, especially in acts I and V.

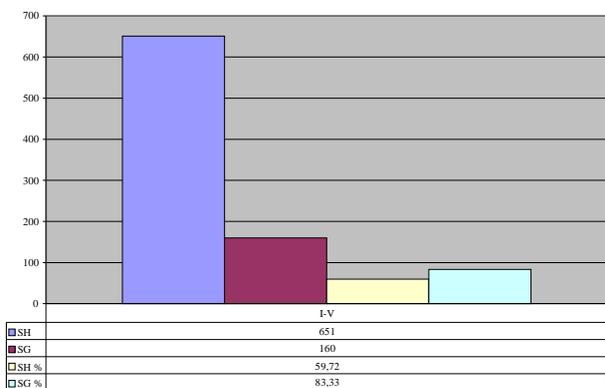


FIGURE 4 Distribution Patterns of the Total Interventions of All Main Characters per Full Text

To better illustrate the distribution patterns of the total interventions of all main characters per full text, Figure 4 is designed. The data in this figure show that the frequency of occurrence of total interventions of all main characters per full text equals 651 in SH in contrast to 160 in SG. However, the total percentage of interventions of all main characters is substantially higher in SG than in SH, that is, 83.33 % against 59.72 %, respectively.

Indeed, the data in Table 3 and Figures 2-4 seemingly point towards the fact that all main characters are assigned a more notable role in SG as opposed to SH.

### 3.3 SUMMARY OF THE MAIN FINDINGS

With regard to the distribution patterns of the presence, I have observed that Shakespeare keeps more or less the same number of all main characters from act I to act III, decreasing their presence dramatically in acts IV and V. By contrast, Sumarokov gradually increases their number, reaching its peak in act V.

With respect to the distribution patterns of the total interventions of all main characters, I have found out that all main characters intervene more frequently and, therefore, carry more weight, specifically in acts I and V in SG. In act III, Sumarokov, as opposed to Shakespeare, distributes them in a more striking way, which proves that there exists a considerable partial divergence in this act inter-plays.

The aforementioned findings fit with the results in the previous research linked to the analysis of the presence, interventions and interactions of the main and secondary characters as well as of the topics dealt with in SH and SG [Keshabyan-Ivanova 2011].

To give an example of clear correlations between the results of the quantitative approach and Sumarokov’s literary originality in relation to his text’s structure and

meaning, I shall compare the findings from this work as well as from the above-mentioned investigation linked to act I in SH and SG.

As a result of these studies, in act I Shakespeare's Hamlet intervenes much more than the other main characters, namely Claudius, Polonius, Gertrude and Ophelia; however, not more than the secondary characters [Keshabyan-Ivanova 2011: 92-95, 128, 129, 137 and 138]. Moreover, he interacts only two times with Claudius and three times with Gertrude [Keshabyan-Ivanova 2011: 137].

In comparison to Shakespeare, only two main characters –Hamlet and Gertrude appear and socialise with each other in act I in SG. At the same time, Hamlet interacts, although with a lower frequency, with his confidant Armands –a secondary character that is present only in SG [Keshabyan-Ivanova 2011: 157].

Consequently, the interrelation between Hamlet and secondary characters is more prominent in act I in SH than in SG where the linkage between two main characters –Hamlet and Gertrude, who represent all main characters in this act, is of major interest.

Furthermore, the results, obtained through the quantitative and qualitative comparison of the distribution patterns of the most frequent content words in act I, provide evidence to the fact that the topic of religion, with its traditional moral values, represents a greater appeal for Sumarokov as opposed to Shakespeare who is drawn to it to only some extent [Keshabyan-Ivanova 2011: 290-291].

Indeed, Sumarokov, through his main character Hamlet, tries to improve the morality of another main character Gertrude –in other words, in line with Sumarokov, to save the sinner's soul. Shakespeare also links the human soul to God's instructions, although he is less straightforward and explicit than Sumarokov is in his moral plan [Keshabyan-Ivanova 2011: 270].

## References

- [1] Stubbs M 2005 *Language and Literature* 14(1) 5–24
- [2] Wynne M 2006 *Encyclopedia of Language & Linguistics* 12 ed Keith Brown Oxford: Elsevier 223-6
- [3] Biber D 2011 *Scientific Study of Literature* 1(1) 15-23
- [4] Johnson J 2011 *Target* 23(1) 62-76
- [5] Shakespeare W 1685 *The Tragedy of Hamlet Prince of Denmark* Mr William Shakespeare's Comedies, Histories, and Tragedies London: H Herringman, E Brewster, R Chiswell and R Bentley 59-86
- [6] Sumarokov A 1787 *Gamlet Tragediya Polnoe sobranie vseh sochineny v stikhakh i proze v 10-ti tomakh Aleksandra Petrovicha Sumarokova Vol 3* ed N Novikov 10 vols V Moskva: V Universitetskaya Tip 61-134 (in Russian)
- [7] Sumarokov 1970 *Hamlet A Tragedy Selected Tragedies of A P Sumarokov* ed Henry M Nebel Jr transl R Fortune Evanston: Northwestern UP 87-134
- [8] Lang D 1948 *Modern Language Review* 43(1) 67-72
- [9] Levitt M 1994 *Slavic and East European Journal* 38(2) 319-41
- [10] Bradley A C 1904 *Shakespearean Tragedy* London: Macmillan
- [11] Wilson J D 1959 *What Happens in Hamlet* Cambridge: Cambridge UP

## 4 Conclusions

The current study has explored the frequency of the distribution patterns of the presence and total interventions of all main characters in *Hamlet* and *Gamlet*. Obviously, the analysis indicates that Shakespeare and Sumarokov set rather distinct aims associated with all main characters.

Compared to Shakespeare, Sumarokov pays greater attention to the main characters, that is, people of a high social rank. At the same time, these aims had a great impact on the structure of the plays, leading to noticeable diversions in the role and weight assigned by the authors to them per different acts inter-plays.

In this respect, it should be noted that to produce a critical analysis of the structure or of the subject of study, that is, to answer the question what quantitative differences mean with regard to literary analysis, more variables should be incorporated. For example, the method could measure presence alone or along other characters, at which points within acts the characters occur or intervene, the relevance of the intervention and/or their speech/actions in the development of each act, the different importance of each character for a distinct purpose, etc.

Finally, it should be highlighted that the present research represents only a sample of quantitative analysis of literary texts as, with this paper, I have aimed to suggest that this kind of empirical work is needed to underpin qualitative literary analysis. However, limitations of computer-assisted textual analysis should always be taken into consideration.

## Acknowledgments

I would like to thank Dr Pascual Cantos-Gómez who was the source of the invaluable knowledge needed to carry out this study.

- [12] Johnson S 1960 *Samuel Johnson on Shakespeare* ed W K Wimsatt Jr New York: Hill
- [13] Eagleton T 1986 *William Shakespeare* Re-reading Literature Series Oxford: Oxford UP
- [14] Kermode F 2000 *Shakespeare's Language* Harmondsworth: Penguin
- [15] Tynianov Yu 1929 *Oda kak oratorsky zhanr* Arkhaisty i novatory Moskva Leningrad: Priboy (in Russian)
- [16] Billington J 1970 *The Icon and the Axe: An Interpretative History of Russian Culture* New York: Vintage
- [17] Gukovskii G 2003 *Russkaya literatura XVIII veka* Moskva: Aspekt P (in Russian)
- [18] Murphy S 2006 *Papers from the Lancaster University Postgraduate Conference in Linguistics & Language Teaching* 1 66-85
- [19] Fish S 1996 1973 *The Stylistics Reader. From Roman Jakobson to the Present* ed J J Weber London: Arnold 94-116
- [20] Louw B 1997 *Teaching and Language Corpora* ed A Wichman, S Fligelstone, T McEnery and G Knowels London: Longman 240-51
- [21] Culpeper J 2002 *Conversation in Life and in Literature: Papers from the ASLA Symposium 15* ed U Melander-Marttala, C Ostman and M Kyto Uppsala: Association Suedoise de Linguistique Appliquee 11-30

- [22] Busse U 2002 *Linguistic Variation in the Shakespeare Corpus: morpho-syntactic variability of second person pronouns*. Amsterdam: John Benjamins
- [23] Archer D, Culpeper J, Rayson P 2009 *What's in a Word-list? Investigating word frequency and keyword extraction* ed A Dawn Farnham: Ashgate 137-158
- [24] Anderson Th, Crossley S 2011 *Stylistics and Shakespeare: Transdisciplinary Approaches* ed M Ravassat and J Culpeper London: Continuum 192-214
- [25] Altman D 1991 *Practical Statistics for Medical Research*. London: Chapman & Hall 285-8
- [26] Keshabyan-Ivanova I 2011 *Shakespeare's Hamlet versus Sumarokov's Gamlet: A Corpus-Based Perspective* Saarbrücken: LAP LAMBERT

## Authors

I Keshabyan-Ivanova, born on November 25, 1952, Sochi (Russia)



**Current position, grades:** Lecturer in the English Department at the University of Murcia (Spain)

**University studies:** She studied English and German at the Pyatigorsk State University of Foreign Languages (Russia) and at the Yerevan Linguistics University (BA Hons). She holds a PhD in English/Corpus and Applied Linguistics from the University of Murcia (Spain).

**Scientific interest:** Her research interests focus on the application of computer-aided corpus-based methodologies to literary texts – specifically to structural and lexical analyses. Recently, she has also developed an interest in frame semantics. Since 2009 she has been a member of the LACELL Research Group at the University of Murcia.

**Publications:** Her publications include a book, entitled *Shakespeare's Hamlet versus Sumarokov's Gamlet: A Corpus-Based Perspective*, published by LAP LAMBERT Academic Publishing; articles in journals such as *IJES*, *RCEI*, etc. In collaboration with Dr Angela Almela, she has co-edited a special-themed issue on *A New Approach to Literature: Corpus Linguistics* in the *International Journal of English Studies* (2012).

**Experience:** She has been a Lecturer of English at the University of Murcia since 2001. She teaches in different areas of English Philology such as Philology and Culture of English-speaking Countries, Business English, Legal English, etc. She was previously a Lecturer of English in the Department of Intensive Methods of Teaching Foreign Languages at the Yerevan Pedagogical University and a Lecturer in Phonetics and Oral Practice at the Yerevan Linguistics University.

# Design of software error detection system based on SPARC V8 and research on the key technology

**Chunmei Huang<sup>1\*</sup>, Chunmao Jiang<sup>1</sup>, MingCheng Qu<sup>2</sup>**

<sup>1</sup>*School of Computer Science Technology and Information Engineering, Harbin Normal University, Harbin, Heilongjiang 150025*

<sup>2</sup>*School of Computer Science and Technology, Harbin Institute of Technology*

*Received 12 June 2014, www.tsi.lv*

---

## Abstract

This paper analyses the key issues confronted when SIHFT is implemented on the SPARC V8 platform, gives the algorithm to solve the problem and the corresponding technology solutions. Software error detection technology system was designed based on SPARC V8, and software signature control flow error detection technology was implemented, the system is based on the architecture of SPARC V8, uses software signature control flow error detection technology and copy instruction error detection technology as the prototype, it's a software system which detects transient faults induced by space radiation and was developed through research, analysis and transformation, with availability, modifiability, portability, maintainability, readability, scalability and other features. The error detection coverage rate of software error detection technology suitable for target platform was tested through simulation experiments. The result data of experiments conducted in the emulator TSIM shows that on the basis of given average performance overhead, the system had high error detection coverage rate when brought in register injected fault and memory injection fault. This proved the SIHFT technology is feasible and effective.

*Keywords:* SIHFT, architecture, SPARC V8

---

## 1 Introduction

SIHFT (Software Implemented Hardware Fault Tolerance) technology is a Software fault-tolerant technology, which is developed by the Stanford university centre for reliability calculation for ARGOS projects, through the method of software to detect and correct hardware transient fault caused by radiation. It is comprised of a software implemented EDAC technology, copy instruction error detection technology (EDDI), Control Flow Checking by Software Signatures (CFCSS) and Watchdog timer. Due to the majority of storage has implemented hardware EDAC at present, while watchdog timer belongs to the category of hardware research, so in this article, we are dedicated to research two core technologies which are Control Flow Checking By Software Signatures(CFCSS) and copy instruction error detection (EDDI).

SIHFT is still in the preparatory stage in domestic, has not yet entered the stage of practical development [1, 2]. In foreign countries, space research institutions and academia carried on the thorough research on this question and some research results were obtained [3, 4]. The United States ARGOS (the Advanced Research and Global Observations Satellite) Satellite did an in-orbit experiment about the main method of the commercial device resist radiation and the SIHFT error detection coverage reached more than 99% [5].

Bound by design technology and production technology and other aspects, there is a large gap between the domestic research on microprocessor and application level with foreign. Generally speaking, the microprocessor performance is low and relies on import leading to cannot satisfy the spacecraft to achieve more control autonomously and data processing requirements. What's more, in the field of high technology our countries are blockaded by foreign countries. Therefore, it is very difficult to obtain radiation-hardened hardware from abroad, so the software fault-tolerant technology has a special significance in the development of China's space industry. This paper analyses the key issues confronted when SIHFT is implemented on the SPARC V8 platform, gives the algorithm to solve the problem and the corresponding technology solutions. And on this basis, design a software error detection technology system based on SPARC V8.

## 2 Fundamental Concepts

### 2.1 SOFTWARE SIGNATURE CONTROL FLOW ERROR DETECTION TECHNOLOGY

Software signature control flow error detection technology is a kind of do not need a watchdog processor or other hardware accessory pure software control flow error detection technology, used to test the execution of a program control flow errors which caused by the

---

\* *Corresponding author* e-mail: hsdrose@126.com

radiation. The primary idea is to divide program into basic blocks, and next construct the program flow chart, for each basic block gives a unique signature in advance, which called a compile time signature, and compute the value which is the result of function  $f$  while the input is the signature of the father's basic block and the son's, and later save the value into the basic block of the son, this information is compiled into the program at compile time. When the program executed, the control flow for each run to a new node will generate a runtime signature  $G$  according to the information, which is compiled into the program at compile time. Subsequently compare the runtime signature  $G$  and compile time signature, if they are equal, there are no program control flow errors occurred; if unequal, there are the program control flow errors have taken place, then turn into fault handler. When multiple control flow inflow the same node, runtime adjusted signature  $D$  is required to test control flow detection.

Software signature control flow error detection technology used  $V = \{v1, v2, \dots, vn\}$  represents the basic block of a collection of nodes, used  $E = \{brij \mid brij \text{ is the edge from node } vi \text{ to another node } vj\}$  represent a collection of edges of control flow between the basic blocks, so a program can be expressed by a program flow chart map  $P = \{V, E\}$ . Node  $vj$  is included by the succeed collection  $suc(vi)$  of node  $vi$ , if and only if  $brij$  in  $E$ . Node  $vi$  is included by the precursor collection  $pred(vj)$  of node  $vj$ , if and only if  $brij$  in  $E$ . If during the program executes, the edge  $brij$  exist but  $brij$  is not in  $E$ , says edge  $brij$  is illegal. Illegal side shows that the control flow error happened. If one node  $v$  received more than two control flow, referred to as the fan-in node, which has more than one node in the  $pred(v)$ .

2.2 COPY INSTRUCTION ERROR DETECTION TECHNOLOGY

Copy instruction error detection technology refers to the repeated calculation by "copy" instructions to error detection. The basic idea is to use different registers and instruction of the variable "copy" assembly language source program. The original instruction of the assembly language source program is called main instructions; the "copy" instructions added to the source program is called shadow instruction. General registers and memory units are divided into two groups to correspond to the main instructions and shadow instructions. The homologous registers and memory units should always have the same value, if the value in a pair of register corresponding to the main instructions and the shadow instructions becomes not equal because of the instantaneous error, therefore, compare values in this two register can detect the errors. In the copy instruction error detection technology, introducing comparison order to compare the value of the corresponding register, if they are an unequal then call error handler.

3 System Design

Software error detection technology system to after modification of the GCC cross compiler generates a fixed part of registers the assembly language source as an input, and to the error detection function of an assembly language program as an output. As shown in figure 1, it is the top chart of the system data flow graph, depicts that the developed system data exchange relations with the surrounding environment.

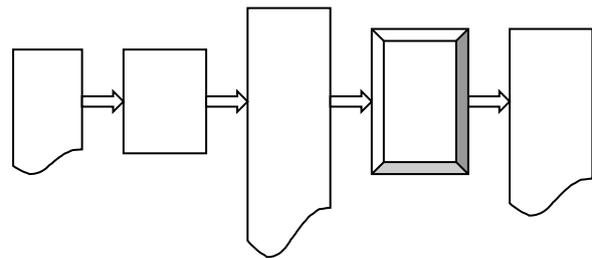


FIGURE 1 The system model of software error detection system

By analysing software signature control flow error detection technology and copy instruction error detection technology, the processing steps of the system should be, first of all, needed to divide the assembly language source program into basic blocks, based on basic block generate the program flow chart, and next according to the structure of the program flow chart in each basic block's head to add software signature check instructions. For each basic block, no stores basic block is divided at first, subsequently generates shadow instruction and constructs a dependency graph, finally carries on the instruction scheduling. Therefore, we will functions divided into basic blocks; software signature technology processing and copy instructions deal with three parts, as shown in figure 2.

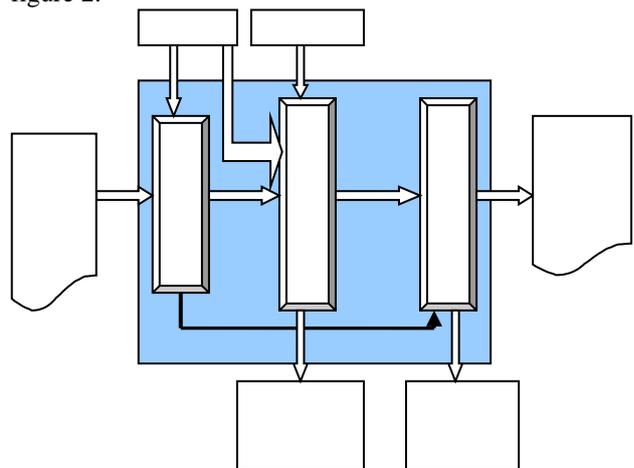


FIGURE 2 The function of software error detection system

System has three input files, which are the assembly language source program; instruction opcode grouped information file and register division method file. Considering the modifiability and scalability of the system, will be mainly related to the specific platform information about an input to the form of a configuration

file. Instruction opcode grouped information file storage system related instruction operation code and the information about the operation code group. Register division method file storage SPARC V8 registers the use method includes the corresponding relation of main registers and shadow registers; software signature control flow error detection technology used registers, the matching relation of status register and a certain register. After the assembly language source program divided into basic blocks, with basic block's information and the configuration file input to the software signature part processing, therefore, obtained the assembly language program, which has the function of control flow check. Finally, delivery the signature check instructions and assembly language source program of each basic block to copy instruction's parts processing, obtained the assembly language program which has the function of copy instructions check and control flow check.

After dividing the basic block, the data flow passed to copy instruction's error detection technology is the information of basic block. After software signature error detection technology processing, the data flow passed to copy instruction technology is with signature check instructions information of basic block. These two types of information transmitted is intended to enable the two technologies can be used independently; the system is easy to cut.

In summary, the software error detection technology based on SPARC V8 system architecture is shown in figure 3. System consists of three modules. Basic block partition module called label mapping, function return statement processing, determine the basic block entry and determine the basic block exports. Software signature technology processing function subdivided into control flow graph generation; add software signature check instructions and optimization. Copy instruction's technology processing called no store basic block division, produce the shadow's instructions, and construct a dependency graph and instruction scheduling. Command recognition module almost is called by all modules (not noted on the picture). Each module implements must follow high cohesion and low coupling principles. System according to the level gradually refined until the bottom of the module.

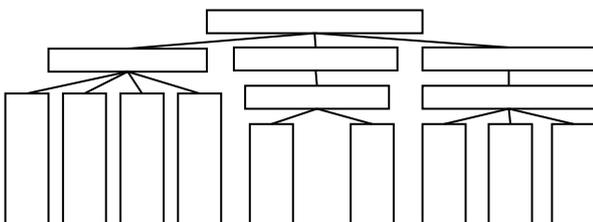


FIGURE 3 The structure of software error detection system

## 4 The key technology research

### 4.1 INSTRUCTION RECOGNITION

Function of the instruction recognition modules is given an instruction; identify the opcode of this instruction. In basic block division, shadow instruction generation and other process modules, all need to be able to correct recognition instruction, thus instruction recognition module is called by the upper most modules.

In order to improve the efficiency of instruction recognition, all the opcode saved into the hash table in alphabetical order, hash function: hash address = first address of the hash table + opcode field first character ASCII value - 97. During the recognition, first calculate the hash address according to the opcode field first character, second followed by matching opcode of the overflow table, if the overflow table has a prefix of the encoding, depending on the opcode field length matching recognition.

All the instruction opcode related to system and grouping information are stored in the instruction opcode grouping information files, the instructions grouped by category (data processing, data transmission and control flow), syntax format and the processing method of shadow instruction.

### 4.2 BASIC BLOCK DIVISION

Basic block is the largest collection of the statement's sequence for the program, which can be executed sequentially; it has only one entry statement and export. Basic block is an important concept in the control flow and data flow analysis to the compiler theory. Basic block division is the preamble step to construct a program flow graph. Program flow graph is the program's control flow semantic representation; each node represents a basic block, and edge represents control flow between basic blocks. The division of basic block in compiler theory and based on basic block division code optimization is carried out on the intermediate code, while the software error detection technology basic block division was conducted on the assembly code. The abstract intermediate code of a basic block division algorithm is applied to the assembly language on basic block division, and the handling of the specific issues related to an instruction set is needed for basic block division problem. The intermediate code of the basic block division algorithm from here, assembly language basic block division algorithm is given.

#### 4.2.1 Basic block division algorithm of assembly language program

---

Input: assembly language source program;  
Output: The information table of Basic block  
(1) Label mapping

Establish one-to-one relationship between the label and its line number.

(2) Sub functions return statement processing

Set up corresponding relationship between function return statements of assembly language program and the statement to jump.

(3) Determine the entry of basic block

After program entry, the line numbers in the assembly instructions to fill entrance set.

For the part of operation code, which is B/CALL/TA/FBA/CBA or B + A unconditional jump instruction, the label will have to jump map to line number, fill the assembly instruction line number which after the line number just mentioned to the entrance set .

For the function return statements, fill the line number followed by the line number of assembly instruction, to which is function return statement jumped to the entrance set.

For conditional transfer instruction, assembly instruction line number after this instruction fills to the entrance set. Then let the conditional transfer instruction label mapping line number, fill the assembly instruction line number followed by this line number to the entrance set.

(4) Determine the export of basic block

Sort the entrance set by the line number ascending,  $S_1, S_2, \dots, S_n$ . For  $S_i, i = 1, 2, \dots, n - 1$ , will be the first transfer instruction's line number after  $S_i$  into export sets, if there is no transfer instruction, will be the last one before entrance  $S_{i+1}$  assembly instruction line number fill in the exports set. For  $S_n$ , will be the first transfer instruction's line number after  $S_n$  into export sets

### 4.3 CFCSS INSTRUCTION GENERATION

#### 4.3.1 Generation algorithm of signature error detect instruction

Input: Program flow graph

Output: The information table is comprised of error detect instruction

Explanation: Algorithm is to define the composition of each node in the flow diagram error detect instructions. The error detect instructions are divided into instruction related to G as "G=G $\oplus$ dj", "br G $\neq$ sj error", "G=G $\oplus$ D" and related to D such as "D=0000" or "D=S $_i$  $\oplus$ S $_m$ ". There are all or part of the instructions included by one node. For one certain node, the position of G\_FLAG represents the composition of error detect instruction related to G has been defined; the position of D\_FLAG represents whether error detect instruction related to D has been defined in the error detect instructions. When it detects a situation that is beyond the scope of processing a message will be sent out and quit.

Algorithmic process:

(1) Initialization, count the data which the subsequent steps of the algorithm needed. Including, the

number of each node's precursors, if the number of precursor is greater than 1, mark for the fan-in node; the number of each node's successor and the number of fan-in node in the successor, mark whether each successor is fan-in node.

(2) Grants each node of the program flow graph a unique signature.

(3) Sort all the nodes by the number of fan-in node in successor descending.

(4) For each node i in the list after sorting

If (no precursor node) //head node

{ Assembly code is "mov GSR, Si"; location is G\_FLAG;

If there is no successor node, there is no instruction of "D=S $_i$  $\oplus$ S $_j$ " in the mark error detection instruction.

}

else

//both have precursor and successor node

{ if (successor node j is non-fan-in node)

{ calculate dj=S $_i$  $\oplus$  S $_j$ , there is instruction of "G=G $\oplus$ dj" and "br G $\neq$ sj error" in the mark error detection instruction, there is no instruction of "G=G $\oplus$ D", location is G\_FLAG.

If there is no fan-in node in the successor of node, there is no instruction of "D=S $_i$  $\oplus$ S $_j$ ", location is G\_FLAG.

}

else if(successor node j is fan-in node)

{ if(i's D\_FLAG has set)

{ if(there is a fan-in node in it's successor node G\_FLAG has not set)

tip exit;

else go to(4)(handle next node)

}

else

{ if(there is a D\_FLAG had positioned pioneer node ,which is node i's successor's fan-in ) tip exit;

else

for each fan-in node j: calculate dj=S $_i$  $\oplus$  S $_j$ , there is instruction of "G=G $\oplus$ dj", "br G $\neq$ sj error" and "G=G $\oplus$ D" in the mark error detection instruction, location is G\_FLAG.For i, instruction of "D=0000" in the mark error detection instruction ,location is D\_FLAG. For each non- precursor node of i, there is instruction "D=S $_i$  $\oplus$  S $_m$ " in the mark error detection instruction, location is D\_FLAG.

}

}

}

#### 4.3.2 The optimization of the signature error detect instructions

The function of optimization of the signature error detected instructions is to optimize the signature error detect instruction, to reduce the cost of add signature error detect instructions. According to software signature

error detection technology optimization algorithm, under the need in some basic blocks first do not conduct the compare of error detection. If there is an error of control flow, error will be according to the control flow being transferred to the back, until detected in once compared. The advantage is to reduce the comparison instruction, improves the application performance; the shortcoming is to accumulate the error backward, once found errors, code needs to be executed again whose length gets greater during restoration.

#### 4.4 INSTRUCTION GENERATION OF EDDI

Copy instruction error detection technology through the "copy" instruction to conduct repeated to calculate to detect transient faults. In this technique, general-purpose registers are divided to two groups, called primary registers group and shadow registers a group. The values of the primary register with the corresponding shadow register always are same. Under the SPARC V8 architecture, some registers do not have enough corresponding shadow registers, so that bringing shadow instruction generation method difference with the original technology method. Learned by the ideas of SWIFT algorithm, in the implementation, not have the copy on storage instructions and memory, which does not implement the memory test.

#### 4.5 INSTRUCTION SCHEDULING

Instruction scheduling is the method of a sort for main instruction and shadow instruction to improve the error detection coverage rate and reduce the execution time in the instruction error detection technology. In order to ensure the instruction sequence after scheduled and without scheduling instruction sequence on the semantic equivalence, instruction scheduling be based on instruction dependency relationship graph. In copy instruction error detection technology, the dependency relationship graph is instructions as the vertex and the dependency relationship as to the edge of the directed graph.

The dependency relationship between instructions is must satisfy the constraint conditions during instruction scheduling. Reference [6] is pointed out shortcoming of the definition [5] of the original dependency relationship, according to the principle of compilation techniques [7, 8], dependency relationship between instructions was redefined as: the instruction  $j$  after the instruction  $i$ , instructions  $j$  depend on the  $i$  when the instructions  $i$  and  $j$  is really relevant / anti-related or output-related. In the no storage basic block when the last instruction is store instructions or transfer instructions, not to participate in instruction scheduling. Thus, in no storage basic block, instruction which participated in scheduling could only happen in the register data-related. In establishing instruction dependency graph, only need to according to

the register data-related to establish dependency relationship.

In the copy command error detection technology, goal of scheduling instruction is to maximize the error detection coverage and minimize the execution time. Under the superscalar architecture, instruction scheduling can reduce the program execution time [9, 10]. Assembly line of SPARC V8 processor is five levels; it has no correlation [11] in every stage, so the instruction scheduling will not reduce the instruction execution time. Thus, the goal of instruction scheduling is how to maximize the error detection coverage. According to the instruction scheduling algorithm, on the premise of meet the dependency graph, required to execute commands to the  $i$ , number of primary instructions and shadow not equal. On the premise of meet the dependency graph, priority scheduling primary instruction, when perform to a certain instruction, the number of primary instructions and shadow are greatest, therefore, the error detection coverage is the largest. In no stored basic block does not participate in instruction scheduling is storage instruction located on the last of no stored basic block and transfer instruction and so on. Therefore, before the instruction scheduling, it first determines the boundaries of scheduling and then establishes a dependency graph, finally perform scheduling.

Instruction Scheduling Algorithm:

---

input: Instruction list of the no stored basic block

output: Instruction list after scheduling

Algorithmic process:

- (1) Determine the boundaries of scheduling
- (2) Establish dependency graph
- (3) Perform scheduling according to dependency graph:

The node which precursor is 0 arranged in two lists by the main command and shadow command.

while (number of main command in the list > 0 || number of shadow command > 0)

{if (number of main command in the list > 0)

{Select the line number the smallest primary instruction scheduling. In the original list to delete the node, added to the end of the list of instructions to be scheduled, number of main command in the list reduce one; Subsequent penetration of this command reduce one, add the nodes which penetration is zero to the end of the list of main and shadow command by main command and shadow command. Calculate the number of list command of main command and shadow command. }

else {Steps to deal with the main command list and shadow command list is the same. }

}

---

**5 Simulation and result analysis**

In our experiment, we use a simulator is tsim-eval-2.0.7a [12]. Experimental test code using the QuickSort, InsertSort, Fibonacci, MatrixMul these four commonly used test programs, using Gaisler Research compile BBC cross compiler developed by LEON2 and LEON3 to assembly file, and then use software error detection

technology system processing into assembly code with fault-tolerant code, using the BCC [13] compile a connection to generate the ELF file and execute in the TSIM.

The experiment process is to inject a fault into register and memory code runtime, count the experimental data. Limited space, Table 1 and Fig.4 only show the final experimental results.

TABLE 1 Result of fault injection

index	Fibonacci	QuiciSort	InsertSort	MatrixMul	The average detected coverage of SIHFT	Source average accuracy
PC test	88.0%	82.1%	86.7%	92.1%	87.2%	54.5%
NPC test	75.9%	76.3%	88.6%	70.8%	77.9%	45.0%
R test	98.0%	96.1%	90.4%	90.0%	93.6%	75.7%
FP test	76.0%	72.9%	99.0%	92.1%	85.0%	28.1%
MEM test	82.4%	83.2%	75.7%	78.0%	79.8%	37.5%

Table of Real-time data contrast

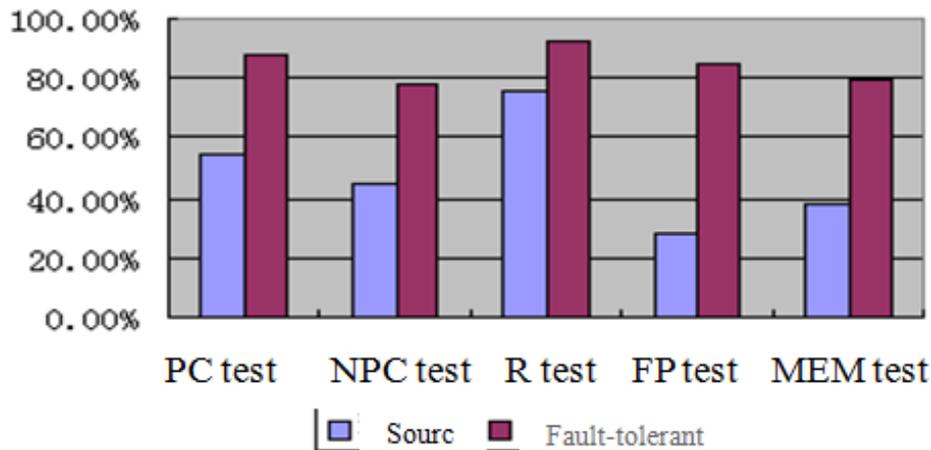


FIGURE 4 Comparison chart of source code and SIHFT error detection rate

According to calculation of 32 general-purpose registers/ PC/ NPC and four status register, statistics of Fibonacci,

etc. four test cases used r register, through calculation, error detection coverage over the last as shown in table 2.

TABLE 2 Result of fault injection

Test procedure	Source Code incorrect results rate	SIHFT undetected rate	SIHFT detected rate
Fibonacci	10.3%	4.8%	95.2%
QuickSort	11.0%	4.0%	96.0%
InsertSort	10.4%	5.3%	94.7%
MatrixMul	11.4%	4.8%	95.2%
average value	10.8%	4.7%	95.3%

It can be seen from table 2 source code register failure not detection rate was 10.8% (system crash + wrong result+ Infinite loop) and after joining fault-tolerant code, register failure not detection rate fell to 4.7%, namely the register error detection coverage rate is 95.3%.

In the results from the experiments on the simulators, TSIM data indicate that assuming the given performance overhead, register to inject faults, error detection coverage rate is 95.3%; Memory Injects faults, error detection coverage of 79.8%, this suggesting that SIHFT technique is effective and feasible.

**6 Conclusions**

This paper, according to the actual requirements of project, analysed the key problems that are realized on SPARC V8 platform to SIHFT technology, and gave the algorithm to solve the problems and the corresponding technical solution, on this basis, designed a software error detection technology system based on SPARC V8, realized the software signature control flow error detection techniques, tested the error detection coverage that is suitable for the target platform software error detection technology by the simulation test. The results

from the experiments on the simulators TSIM data show that, on the basis of the average performance overhead has been given, the coverage of error detection is higher when register and storage injection failure has been introduced, it shows that SIHFT technique is effective and feasible.

**References**

[1] Keyan Pan, Changlong Wang 1998 Radiation-hardened Technology *Onboard Digital Electronic Devices* (3) 67-8  
 [2] Qiongying Ren, Jinrong Cai, Guangxuan Luo 2012 *Single Particle Effects of Board Computer and Protection and Reinforcement of it's Software* Guizhou University (Natural Science Edition) 15(3) 178-80  
 [3] Oh N, Shirvani P P, McCluskey E J 2002 Control-flow checking by software signatures *IEEE Transactions on Reliability* 51(1) 111-22  
 [4] Oh N, Shirvani P P, McCluskey E J 2002 Error detection by duplicated instructions in super-scalar processors *IEEE Transactions on Reliability* 51(1) 63-75  
 [5] Shirvani P P, Saxena N, Oh N, Mitra S, Yu Shu-Yi, Huang Wei-Je, Fernandez-Gomez S, Toubia N A, McCluskey E J 2013 *Fault-Tolerance Projects at Stanford CRC* CRC Technical Report  
 [6] Zhenyuan Huang 2006 *Research and Implementation of an Onboard Computer Software Error Detection Technology* Harbin Institute of Technology Master Thesis 2006 50-7

**Acknowledgment**

This work was supported by Harbin reserve Talent project (2014RFQXJ073).

[7] Muchnick S S 2005 *Advanced Compiler Design and Implementation* Mogran Kaufman Kejia Zhao and Zhiyu Shen. Machinery Industry Press 195-9, 381-90  
 [8] Shen J P, Lipasti M H 2004 *Modern Processor Design - Basis of Superscalar Processor* Chengyi Zhang, Yu Deng, Lei Wang. Electronic Industry Press 44-7, 8586  
 [9] Boyin Lu, Baolin Yin 2001 A Scheduling Optimization Algorithm Based on DAG Graph Instruction *Computer Engineering and Applications* 2001(12) 121-4  
 [10] Shuxin Yang, Zhaoqing Zhang 2004 Global Instruction Scheduling Summary *Computer Engineering and Applications* 2004(21) 24  
 [11] *The SPARC Architecture Manual Version 8* SPARC International, Inc. Revision SAV080SI9308. 1992 1~57  
 [12] *TSIM2 Simulator User's Manual for Version 2.0.7* Gaisler Research AB. January 2007 6-11  
 [13] *BCC-Bare-C Cross-Compiler User's Manual. Version 1.0.29* Gaisler Research. February 2007 3-18

Authors	
	<p><b>Chunmei Huang</b></p> <p><b>Current position, grades:</b> Lecturer of informatics at Computer &amp; Informatics Department, Haerbin Normal University.  <b>University studies:</b> M.Sc. in Mathematics (2011) from Harbin Engineer University.  <b>Research interests:</b> different aspects of Cloud Computing and Embedded Computing.</p>
	<p><b>Chunmao Jiang</b></p> <p><b>Current position, grades:</b> professor of informatics at Computer &amp; Informatics Department, Haerbin Normal University  <b>University studies:</b> he received his M.Sc. in Mathematics (2004) and PhD in Information Sciences (2013) from Harbin industry University.  <b>Research interests:</b> different aspects of Cloud Computing and Distributed Systems.</p>
	<p><b>Mingcheng Qu, born in 1980</b></p> <p><b>Current position, grades:</b> Ph.D. in school of computer science and technology of Harbin Institute of Technology.  <b>Research interests:</b> embedded computing and P2P etc.</p>

# Image reconstruction of concrete based on Filtered Backprojection method using ultrasonic time of flight data

Honghui Fan<sup>1, 2\*</sup>, Hongjin Zhu<sup>1, 3</sup>, Qingbang Han<sup>4</sup>

<sup>1</sup>School of Computer Engineering, Jiangsu University of Technology, Changzhou 213001, Jiangsu, China

<sup>2</sup>Sichuan Province Key Laboratory of Bridge non-destructive testing and engineering computing, Zigong 643000, Sichuan, China

<sup>3</sup>Key Laboratory of Cloud Computing & Intelligent Information Processing of Changzhou City, Changzhou 213001, Jiangsu, China

<sup>4</sup>College of IOT Engineering, Hohai University, Jiangsu 213022, China

Received 12 June 2014, www.tsi.lv

## Abstract

This research aims to recognize the defect of concrete materials using an ultrasonic computed tomography imaging technique. Filtered Backprojection method was used to reconstruct concrete images in this paper. Ultrasonic time of flight data was measured to reconstruct computer tomography images. 306 data paths were obtained in total by manual scanning for one computer tomography image. We examined the effect of the interpolation data as the density of time of flight data has a considerable effect on image quality. The feasibility of concrete reconstruction system and time of flight data interpolation were examined in detail using numerical and concrete phantoms.

*Keywords:* Image reconstruction, Time of flight, Filtered Backprojection, Interpolation, Concrete

## 1 Introduction

Concrete is currently the most widely used construction material. Its huge popularity is the result of a number of well-known advantages, such as low cost, general availability, and wide applicability. Testing and quality checkup is important at different stages during the life of a structure [1, 2]. The traditional method of evaluating the quality of concrete in civil structures is to test specimens cast simultaneously for compressive, flexural, and tensile strengths; these methods have several disadvantages such as the absence of immediate result prediction [3].

Computer tomography seems to provide a better alternative for quality inspection since it provides the visual difference as reflected in the radiation attenuation profile, radiation attenuation distribution, detail position and detail dimension [4, 5]. Ultrasonic time of flight (TOF) computed tomography was applied to evaluate wooden pillars quality by Tomikawa et al [6]. In the recent research reports, Filtered Backprojection (FBP) algorithm was used for image reconstruction [7-9]. The FBP algorithm was modified to reconstruct computed tomography images from incomplete TOF profiles of wood by Yanagida [10]. FBP algorithm has been extended to fan beam data acquisition geometry in some report [11-13], and has been widely used in industrial computer tomography [14].

In ultrasonic TOF computer tomography, the spatial distribution of sound velocity is estimated [15].

Ultrasonic TOF data testing of concrete is based on ultrasonic velocity in concrete method to provide information on the uniformity of concrete, cavities, cracks and defects [16-18]. The date of TOF in concrete depends on its density and its elastic properties, which in turn are related to the quality and the compressive strength of the concrete. It is therefore possible to obtain information about the properties of components by ultrasonic velocity [19].

FPB algorithm is an image reconstruction method for ultrasonic TOF computer tomography. It was modified to reconstruct computer tomography images from the incomplete time of flight profiles of wood by Fan and of concrete by Suryono et al [20, 21]. Although the quality of the object can be determined, it is difficult to accurately find the position of the holes in reconstruction images [22, 23]. In this paper, we proposed an approach to concrete inspection by ultrasonic TOF computer tomography on the basis of the FBP algorithm. Furthermore, the interpolation of TOF data was applied to enhance image quality in FBP imaging process. The effects of image quality and the number of interpolations in the TOF data were examined in detail with concrete phantom.

## 2 Image reconstruction using FBP algorithm

FBP algorithm for parallel beam projection data with and without attenuation has been well established [24]. Without attenuation, the FBP algorithm has been

\* *Corresponding author* e-mail: fanhonghui@hotmail.com

extended to fan-beam data acquisition geometry, and has been widely used in computer tomography.

The FBP turns to Fourier theory to approach the problem of finding the linear attenuation coefficient at various points in the cross-section of an object [25]. A fundamental result linking Fourier transform to cross-sectional images of an object is the Fourier Slice Theorem [26], and this paper only concerns parallel beam projection data. The Fourier Slice Theorem for the parallel beam projection data is given here. The same justifications can be made for fan beam and cone beam projection data.

When the data  $I_o$  obtained from the observation is expressed as a function by the line integral for route  $s$  physical weight distribution  $f(x, y)$  of the measurement section.

$$I_o = \int_s f(x, y) ds. \tag{1}$$

The arrangement is shown in Figure 1,  $P(r, \theta)$  is the projection data obtained at position  $r$ , and observed from direction  $\theta$ .  $r$ - $s$  coordinate system that is rotated by an angle  $\theta$  from the  $x$ - $y$  coordinate system.

$P(r, \theta)$  could be represented by the Equation (2). Here,  $\delta$  is delta function of Dirac, and the Equation (2) is called Radon transform.

$$P(r, \theta) = \int_{-\infty}^{\infty} f(r \cos \theta - s \sin \theta, r \sin \theta + s \cos \theta) ds, \tag{2}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - r) dx dy$$

$$F(\mu, \nu) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp\{-j2\pi(\mu x + \nu y)\} dx dy, \tag{3}$$

$$F(\rho \cos \theta, \rho \sin \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp\{-j2\pi\rho(x \cos \theta + y \sin \theta)\} dx dy, \tag{4}$$

$$F(\rho \cos \theta, \rho \sin \theta) = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(r \cos \theta - s \sin \theta, r \sin \theta + s \cos \theta) ds \right] \exp\{-j2\pi\rho r\} dr, \tag{5}$$

$$= \int_{-\infty}^{\infty} p(r, \theta) \exp\{-j2\pi\rho r\} dr$$

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\mu, \nu) \exp\{j2\pi(\mu x + \nu y)\} dx dy. \tag{6}$$

Equation (5) indicates that Fourier transforms of projection data  $P(r, \theta)$  from the angle  $\theta$  is identical to the cross-section cut at angle  $\theta$  spectra of the two-dimensional Fourier transform of a physical quantity distribution  $f(x, y)$ . Thus, two dimensional Fourier spectrum is obtained by collecting the projection data from all directions, and restored physical quantity distribution  $f(x, y)$  from a set of projection data  $P(r, \theta)$  by working the inverse operator. Equation (6) expresses the method of conversion Equation (5) to two dimensional inverse Fourier transform.

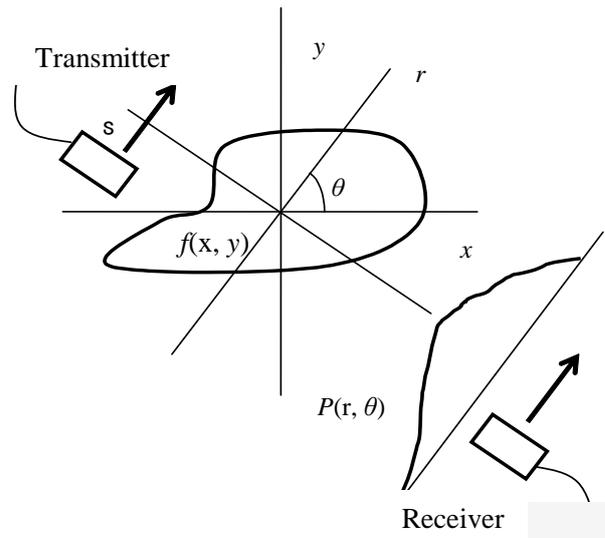


FIGURE 1 Object  $f(x, y)$  and its projection  $P(r, \theta)$

Based on Equation (2),  $f(x, y)$  from  $P(r, \theta)$  could be calculated easily by using two-dimensional Fourier transform and Equation (3).

Here,  $\mu = \rho \cos \theta$ ,  $\nu = \rho \sin \theta$  make a polar coordinate conversion. And using Equation (2), Equation (5) could get

However, because the calculation amount is large, the FBP method is often used in this process. Polar coordinates transform the Equation (6) by using the symmetry of the Fourier transform. Equation (7) is composed as follows.

$$f(x, y) = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} G(\rho) |\rho| \exp(j2\pi\rho r) d\rho \right] d\theta. \tag{7}$$

Here,

$$G(\rho) = \int_{-\infty}^{\infty} p(r, \theta) \exp(-j2\pi\rho r) dr. \tag{8}$$

The system is intended to apply to the contact type transmission time computer tomography image reproduction principle using the FBP method. Physical quantity to be measured in our system since the propagation time, the projection data are represented as follows.

$$TOF(i, j) = \int_s \frac{1}{c(x, y)} ds. \tag{9}$$

Here,  $(i, j)$  is measurement point number,  $TOF(i, j)$  is time of flight data when measurement point  $i$  be used as transmission point, and measurement point  $j$  be used as reception point. In addition,  $s$  is the shortest path of time flight,  $c(x, y)$  shows a sound speed distribution in the measured cross-section.

In Equation (1),  $I_0$  replaced by  $TOF(i, j)$ ,  $f(x, y)$  replaced by  $1/c(x, y)$ , TOF could be considered as a projection. Therefore, the image obtained by the reconstruction is distributed of the inverse of the sound velocity  $c$  (slowness).

In the Equation (10) above, the terms inside the square brackets (the operation indicated by the inner integral) represent a filtering operation and evaluate the filtered projections. The operation being performed by the outer integral evaluates the back projections, which basically represents a smearing of the filtered projections back onto the object and then finding the mean over all the angles.

Given that  $n$  rays pass through a measuring section,  $li$  is the length of ray  $i$  (distance between transmission and reception transducer),  $t_i$  is the time that ray travels along  $li$ . From Randon transform,

$$\tau_i = \int_{L_i} \frac{1}{V_j(x, y)} dl = \int_{L_i} f_j(x, y) dl, \tag{10}$$

where  $V_j(x, y)$  is the velocity of cell  $j$ ,  $f_j(x, y)$  is the slowness of cell  $j$ . It is assumed that the cell is small enough, so  $f_j(x, y)$  of each cell can be considered as constant. Equation (11) can be written as progression form,

$$\zeta_j = \sum_{j=1}^m a_{ij} f_j, \tag{11}$$

where  $a_{ij}$  is the length of ray  $i$  in cell  $j$ . In view of mathematics Equation (12) is a linear equation group.

$$\begin{cases} \tau_1 = a_{11}f_1 + a_{12}f_2 + \dots + a_{1m}f_m \\ \tau_2 = a_{21}f_1 + a_{22}f_2 + \dots + a_{2m}f_m \\ \vdots \\ \tau_n = a_{n1}f_1 + a_{n2}f_2 + \dots + a_{nm}f_m \end{cases}. \tag{12}$$

### 3 Experimental Methods

#### 3.1 EXPERIMENTAL SYSTEM

The experimental system was shown in Figure 2. A couple of bolt-clamped Langevin-type transducers (BLT) of 68 kHz centre frequency were used as an ultrasound transmitter and a receiver.

The 36 measurement points were placed on the circumference of the test sample and labelled 0 to 35. Every measurement point transmits and receives ultrasonic signals in our system.

36 measuring points were placed 10 degrees ( $\theta_c=10^\circ$ ) apart on the circumference of the test sample and labelled with numbers from 0 to 35. The burst wave was transmitted from the  $l$ -th point and received at the  $m$ -th point with a gap number  $n$ , that is,  $m=(l+n)_{\text{mod}36}$ . When gap angle  $\theta_c=10^\circ$ , the gap of numbers between the transmitter and receiver,  $n$ , was kept for  $\theta_c/10=1$  such as (transmitter 0, receiver 1), (1, 2), ... (l, m), ... (35, 0( $= (35+1)_{\text{mod}36}$ )).

TOF for a gap number  $n$  was described as  $t_l^{(n)}$ , where  $l$  is the number of measuring point. TOF of gap angle  $\theta_c=10^\circ$  can be described as

$$t_l^{(n)} = t_l^{(i)} \quad (l = 0 \dots 35). \tag{13}$$

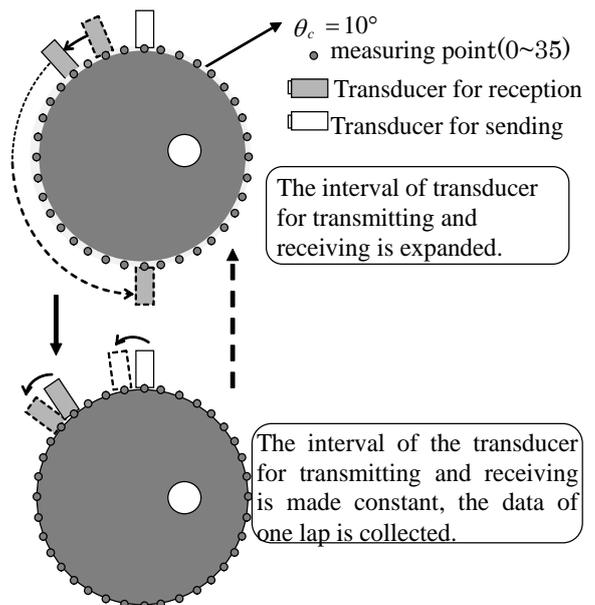


FIGURE 2 Measurement procedures

When wave data were received, there were some noises of low-frequency and high-frequency. To remove the noises from the measurement data, a band pass filter of 10–100 kHz was used for all TOF data. After the filtering, the amplitudes of the wave data were calculated. To avoid error detection by high-frequency noise, 20% of the maximum amplitude value of each wave data was used as threshold levels for TOF detection (Figure 3).

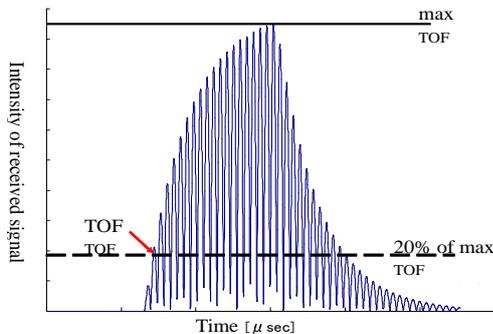


FIGURE 3 The TOF determination

### 3.2 TOF DATA INTERPOLATION

The wavelength of ultrasound of 68 kHz was about 20 mm. 36 measurement points were placed on the circumference. The distance of the interval was about 17mm. All measurement paths were 306. It takes about two hours to obtain 306 TOF for one computer tomography image. So, the number of data that can be measured is thought that 306 were near the upper bound timely and spatially. However, the 306 data was not enough to obtain a clear computer tomography image. To raise the pixel level, angle interpolation for fan beam and distance interpolation for parallel beam were used in the imaging process. The TOF data were obtained measured with the fan - beam geometry of non-equal and coarse intervals. The obtained TOF profiles are converted into the small and regularly-interval data by the spline interpolation.

In the case of  $\theta_c=10^\circ$ , the measurement path, such as (0, 1), (0, 3), (0, 5) ... (0,35) was interpolated, and TOF data with a denser fan beam geometry was obtained. So one measurement point would be passed by more fan-beam, such as (transmitter 0, receiver 1), (0, 2), (0, 3) ... (0, 34), (0, 35). The number of paths was 35 with one measurement point, the TOF data number became 612 {=(number of measurement point 36 × measurement path from one measurement point 35 - duplicate path of diameter route 36) ÷exchange 2} after angle interpolation.

The second step interpolation was used for obtaining the parallel beam with equal intervals of measurement path distance. After angle interpolation, such as (0, 1), (35, 2), (34, 3) ... (19, 18), the number of measurement path with same horizontal was 19. Using distance interpolation we could get 39 (=19 × 2+1) path for one horizontal parallel beam, so total interpolation TOF data of 1404 (=39 × 36) paths were obtained. After these

interpolations, one 64 × 64 pixel computer tomography image could be clearly reconstructed.

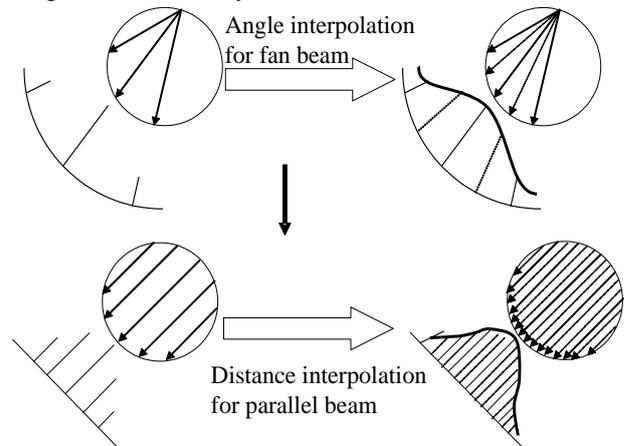


FIGURE 4 TOF data interpolation by fan beam geometry and parallel beam geometry

The process of reconstruction image based on FBP method in our system is shown in Figure 5. The position of measurement points and shape data were first calculated according to coordinates of measurement points and measurement object. According to the interval of transducer for sending and reception, all ultrasonic velocity data were made up of 9 groups. Interpolation TOF data was got using fan beam and parallel beam geometry under measurement TOF beam profile, and pixel value was designated on that wave path. A computer tomography image could be reconstructed based on FBP algorithm.

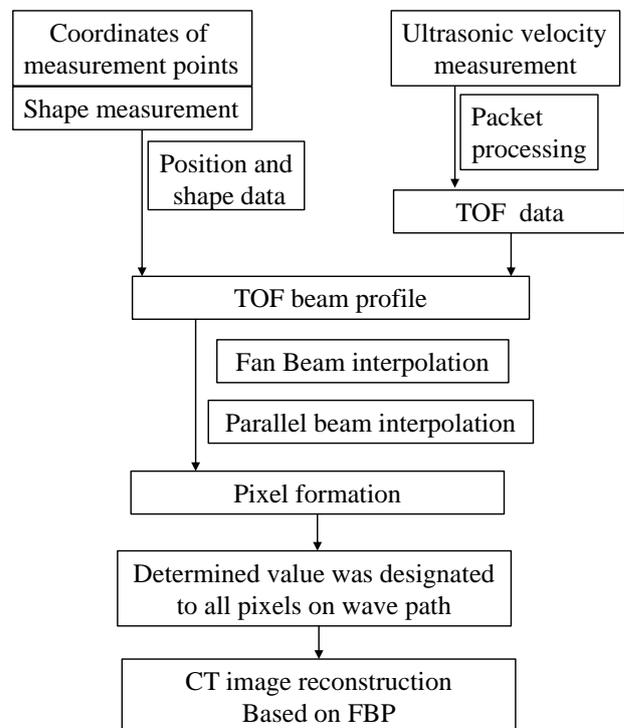


FIGURE 5 Flowchart for reconstruction image based on FBP

4 Results

4.1 NUMERIAL PHANTOM

A numerical phantom containing a circle shaped defect was assumed which was composed of 64 x 64 square pixels. The acoustic velocities were 5000 m/s for normal part, the acoustic velocities 2500 m/s for defect part. Parameters of numerical concrete phantom were shown in Table 1.

TABLE 1 Parameters of concrete

Property	Normal part 5000 [m/s]	Defect part 2500 [m/s]
Poisson's ratio	0.3	0.3
Young's modulus [N/mm <sup>2</sup> ]	1.075×10 <sup>10</sup>	2.688×10 <sup>9</sup>
Density [kg/m <sup>3</sup> ]	430	430

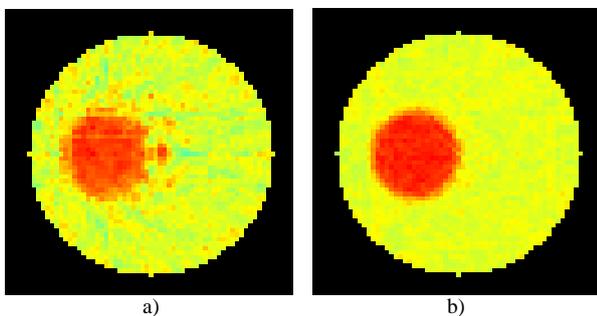
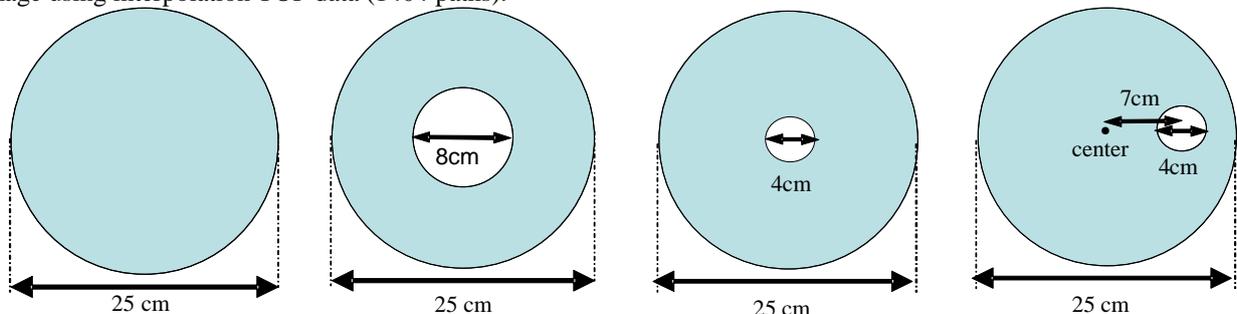


FIGURE 6 Reconstruction images of numerical phantom

The reconstruction images based on FBP method are shown in Figure 6. Figure 6a is the reconstruction image which uses 306 TOF data, Figure 6b is the reconstruction image using interpolation TOF data (1404 paths).



A B C D

FIGURE 7 Inspection object

From the reconstructed images which without interpolation, some artefacts were observed in the images and no clear defects were observed. When angle and distance interpolation TOF data was used, the better-quality computer tomography images could be reconstructed.

4.2 CONCRET PHANTOM

Four concrete test specimens (A test concrete phantom, B test concrete phantom, C test concrete phantom and D test concrete phantom) were produced for the verification of the proposed experimental system. The test specimens were composed of mortar and Styrofoam. A test concrete phantom was consisted entirely of mortar, and was without defect. B, C and D test concrete phantoms were consisted entirely of mortar and polystyrene foam. As shown in Figure 7, B and C test concrete phantoms had a defect which was set in the centre, and the diameter of the defect was 8 cm and 4 cm respectively. All of the test specimen diameters were 25 cm.

The images in Figure 8 were the reconstruction results of the four test concrete phantoms. Test concrete phantoms could be reconstructed based on FBP algorithm. The different defects in the three test concrete phantoms could be found in the reconstructed images. The defect position and size of B, C and D test concrete phantoms were observed, but it was possible that some artefacts were recognized. The results demonstrated the effectiveness of the proposed algorithm.

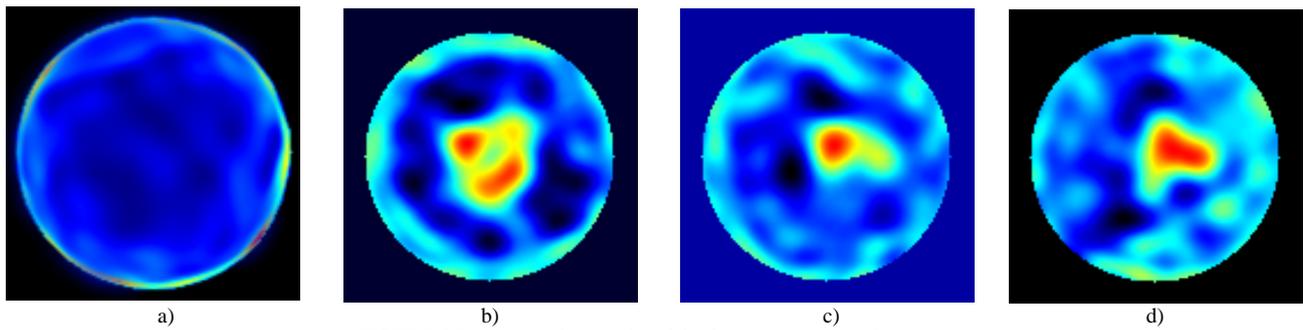


FIGURE 8 Reconstruction results of the four test concrete phantoms

## 5 Discussions

When the FBP method was used for the ultrasonic computer tomography of concrete, 306 TOF data were not sufficient to obtain clear images. By applying the FBP method to ultrasonic computer tomography, we should obtain more TOF data or use an interpolation in the imaging process. Defects with diameters ranging from 4cm to 8cm were recognized by visual observation. However, the defects could not be found clearly in the reconstructed images of the concrete phantom. By comparing the reconstruction images with numerical phantom and concrete phantom, we concluded that visually, the defect was clearly reconstructed in the numerical phantom, and the defects could not be found clearly in the reconstructed images of concrete phantoms. The reason was considered that the sound propagation path was a straight line and it intends to reconstruct the concrete section image in proportion to the distribution of the equivalent sound velocity. The anisotropic property of

concrete and Volatility of ultrasonic were not considered in our system.

## Acknowledgements

The authors are very thankful to Atati Laboratory and Tamura-Yanagida Laboratory from Yamagata University of Japan for providing experimental data. This work was supported by China National Natural Science Fund of China (61302124 and 11274091), Natural Science Fund of Jiangsu Province (BK20130235), Natural Science Foundation of the Higher Education Institutions of Jiangsu Province (13KJB520006), Program of six talent tops of Jiangsu Province (DZXX-031), Sichuan Province Key Laboratory of Bridge non-destructive testing and engineering computing (2013QYJ05), and was partially supported by the Key Laboratory of Cloud Computing & Intelligent Information Processing of Changzhou City under Grant No. CM20123004.

## References

- [1] Mehta P K 2002 *Concrete International* **24**(7) 23-8
- [2] Meyer C 2009 *Cement and Concrete Composites* **31** 601-5
- [3] Maiti S C, Agarwal K 2009 *The Indian concrete journal* September 20-7
- [4] Chauhan M S 2013 *International Research of Sustainable Science and Engineering* **1**(1) 1-3
- [5] Chai H K, Momoki S, Kobayashi Y, Aggelis D G, Shiotani T 2011 *NDT and E International* **44** 206-15
- [6] Tomikawa Y, Iwase Y, Arita K, Yamada H 1986 *IEEE Trans. Ultrason.Ferroelectr. Freq. Control* **33** 354
- [7] Jiangsheng Y, Gengsheng L, Zhengrong L 2005 *Inverse Problems* **21**(5) 1801
- [8] Joshua D E, David G P, Bruce R W, Joseph O S, Jeffrey F W 2011 *Medical Physics* **38**(3) 1444-58
- [9] James B, Susu Y, Justin Roper, William G, Fangfang Y 2014 *Medical Physics* **41**(1) 010701
- [10] Hirotsuka Y, Yasutaka T, Kwang M K, Jun J J 2007 *Japanese Journal of Applied Physics* **46**(8) 5321-5
- [11] Xiaochuan P, Dan X, Yu Z, Lifeng Y 2004 *Physics in Medicine and Biology* **49**(18) 4349-69
- [12] Jongduk B, Norbert J P 2010 *Medical Physics* **37**(5) 2074-81
- [13] Adam W, Frederic N 2008 *Physics in Medicine and Biology* **53**(10) 2471-93
- [14] Suryono, Kusminarto, Bayu S, Aris S 2011 *International Journal of Civil and Environmental Engineering* **11**(5) 17-22
- [15] Honghui F, Hirotsuka Y, Yasutaka T, Shuqiang G, Tadashi S, Tasuhisa T 2010 *Japanese Journal of Applied Physics* **49** 07HC12
- [16] Wunderlich C, Niederleithinger E 2013 *Nondestructive Testing of Materials and Structures* **6** 227-32
- [17] Lee K M, Kim Y H, Bae D B 2004 *Cement and Concrete Research* **34** 631-40
- [18] Mullapudi T R, Di G, Ashraf A 2013 *Magazine of Concrete Research* **65**(18) 1081-91
- [19] Adam F, Jirf J, Igor P 2012 *Radio Engineering* **21**(1) 533-44
- [20] Honghui F, Hongjin Zang, Yao W 2013 *Review on Computers and Software* **8**(1) 17-20
- [21] Alwin S, Antariksa S, Purnama S 2013 *Journal of Basic and Applied Scientific Research* **3**(8) 418-22
- [22] Schickert M 2005 *Materials and Structures* **38** 807-15
- [23] Yanli C 2010 *International Forum on Information Technology and Applications* **1** 392-6
- [24] Hongli S, Shuqian L 2013 *Bio Medical Engineering Online*: <http://www.biomedical-engineering-online.com/content/12/1/50>
- [25] Wang K, Anastasio M A 2012 *Physics in Medicine and Biology* **57**(23) N493
- [26] Chiou S F, Shih S L 1998 *Image and Vision Computing* **16**(9-10) 689-701

Authors	
	<p><b>Honghui Fan, born in October, 1980, Shandong Province</b></p> <p><b>Current position, grades:</b> School of Computer Engineering, Jiangsu University of Technology, Changzhou, Jiangsu, China, Associate professor.</p> <p><b>University studies:</b> Systems and Information Engineering</p> <p><b>Publications:</b> 11 publications</p> <p><b>Experience:</b> He has a M.Sc. And Ph.D. from the Yamgata University of Japan in 2008 and 2011, respectively. He is a member of China Computer Federation. In 2011 he joined Jiangsu University of Technology. His research interests include ultrasound imaging, image restoration, digital image, and signal processing in biomedical engineering.</p>
	<p><b>Hongjin Zhu, born in May, 1981, Jilin Province</b></p> <p><b>Current position, grades:</b> School of Computer Engineering, Jiangsu University of Technology, Changzhou, Jiangsu, China, Lecturer.</p> <p><b>University studies:</b> Systems and Information Engineering</p> <p><b>Publications:</b> 10 publications</p> <p><b>Experience:</b> She has an M.Sc. And Ph.D. from the Yamgata University of Japan in 2007 and 2010, respectively. She was employed as a special researcher in the Department of Engineering, Yamagata University of Japan in 2010. In 2011 she joined Jiangsu University of Technology. Her research interests include image processing, computer vision, pattern recognition and evolutionary computation.</p>
	<p><b>Qingbang Han</b></p> <p><b>Current position, grades:</b> College of IOT Engineering, Hohai University, Jiangsu, China, Professor.</p> <p><b>University studies:</b> Acoustic and information processing technology</p> <p><b>Publications:</b> 25 publications</p> <p><b>Experience:</b> His research interests include signal processing, NDT, Acoustic wave propagation characteristics and Power ultrasonic.</p>

# A line segment detection algorithm based on statistical analyses of quantified directions in digital image

Liang Jia, Nigang Sun\*

School of Information Science & Engineering, Chang Zhou University, 213164, China

Received 12 June 2014, www.tsi.lv

---

## Abstract

Line segment detection is a typical image processing problem with constantly evolving solutions. Following the line segment detect (LSD) by Grompone von Gioi, two branches of algorithms merged. The first branch aimed to improve its speed at the cost of lower accuracy; the second applied in the opposite way. We investigated the philosophies of these methods, and attempt to develop a line segment detection algorithm based on statistical analyses of quantified directions (LSDSA) to achieve better accuracy and faster speed. We utilize a statistical approach estimating the distributions of pixels with direction values approximating the direction changes when traversing along the edges given by any edge detector. It efficiently reduces the dimension of the input data, and incurs limited increasing in computation time for validation process. The simulation results show that the proposed algorithm achieves better performance compared to the existing typical LSD algorithms. The experiment using industrial data in noisy cases also exhibits excellent performance.

*Keywords:* Line segment detection, Hough transformation, Image processing, Pattern recognition

---

## 1 Introduction

Straight lines, such as straight roads, horizons and the walls, are basic visual elements in the world. They are represented by line segments in digital images. As mobile devices and digital cameras became popular, processing images are serviced as a common daily task for many people, and the number of digital images has increased heavily.

As a basis of the image processing algorithms, line segment detection are useful for various high-level image processing tasks such as crack detection in materials [1], robot-navigation [2] and many others [3, 5, 6, 4]. Roughly, there are three sets of typical algorithms for line segment detection [7]:

(1) Algorithms based on geometric duality [9, 11, 12, 10] such as Hough transformation (HT) [8] and Gaussian kernel-based Hough Transform (KHT) [13]. They usually suffer from the expensive computational costs of implementing geometric duality and the low detection accuracy. Although KHT made a great improvement of the voting procedure introduced by HT, it totally depends on the pre-processing procedure composed of algorithms proposed in [14] and [15] to provide the input data.

(2) Algorithms based on the analysis of the gradient orientations. Following the typical LSD of non-parameter-turning features [16], Akinlar proposed a line segment detector based on edge drawing algorithm (EDLines) [7]. EDLines is about 10 times faster than LSD, while its accuracy just approximates LSD. Yang proposed a line segment detector using two-orthogonal

direction image scanning (TODIS) [22], which achieves better detection accuracy compared to LSD, but it consumes about 8 times of the computational time than LSD.

(3) Algorithms using line geometrical properties and the relative positions of pixels. They extract the line segments by traversing along the edge pixels given by the edge detectors. As an example, the algorithm proposed in [20] tries to find the blurred lines in a grey level image and its prototype is reported in [19]. Although the algorithm detects segment accurately, it also generates lots of positive false [7, 17, 22].

Generally, algorithms faster than LSD such as EDLines suffer from lower accuracies; algorithms with higher accuracies than LSD such as TODIS consume more computation resources. In this paper, we attempt to explore a method that could preserve higher accuracy and higher speed, namely, faster than TODIS and more accurate than EDLines

Inspired by the strategy introduced in [18] and direction value processing method proposed in [23], we develop an algorithm called LSDSA maintaining statistical records about direction values found in steps of the traverse and it dynamically decides whether the current traverse should continue or cease in each step based on the records. Once LSDSA finds the traverse leads to a line whose direction values differ a lot from the traversed line segment, the traverse terminates and the line segment is stored. After all found segments are validated, the distorted ones are split to shorter segments based on a statistical computation of the records. Finally,

---

\* *Corresponding author* e-mail: ngsun@cczu.edu.cn

LSDSA checks the possibility of linking the segments of the similar direction values by appropriately extending the segments.

Basically, LSDSA needs two parameters: sample size and the minimal length of a line segment. Since there is no record initially, we sample a number of pixels as the first parameter. The minimal length filters the found segments based on their lengths. Actually, we can combine these two parameters into one parameter and automatically determine it by the dimensions of the image space if necessary. The experiment shows that LSDSA runs at least 2 times faster than TOIDS and its accuracy is higher than EDLines.

The rest of this paper is organized as follows. Section 2 introduces the work related to LSD. Section 3 presents LSDSA. The experimental results are reported in in Section 4. Finally, we conclude the paper in Section 5.

## 2 Background

In this section, we focus on two issues related to LSDSA, namely, inner border tracing [23, 24] and foot-of-normal method [27].

### 2.1 INNER BORDER TRACING

The border tracing is used to find the inner border of a region in a binary image. Typical border tracing methods includes versions for 4-connectivity and 8-connectivity. They label the positions in the neighbourhoods of the different connectivity's by using *direction values* tied to the directions in the image plane. The direction values of 8-connectivity as shown in Fig. 1 reflect more directions than 4-connectivity.

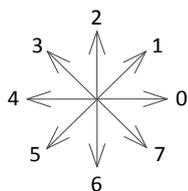


FIGURE 1 Direction values of 8-connectivity

The border tracing algorithm of 8-connectivity is shown in a diagram named activity diagram [25, 26] which satisfies the standards of Unified Modelling Language (UML) as shown in Fig. 2. The mod operations associated with the estimation of odevity of the *dir* yield values lying in a fixed range when *dir* can only be one of the values shown in Fig. 1. Hence there is a mapping between the input and output values of *dir*. The mapping can be represented by a matrix, i.e., a look-up table and the computations of the mod operations which actually are replaced by simple searches are accordingly reduced to  $O(1)$ .

The anti-clockwise search for the non-zero pixel is implemented by updating the variable *dir* after each check of a pixel in the neighbourhood, and stops once a non-zero pixel is found. The search will lead the centre of

the searching to move to the found pixel and repeat. A continuous series of searches is called *tracing*. The tracing always starts at the first non-zero pixel in the upper left corner of the region and ends at the same pixel, no matter what shape of the checked region is used. Non-zero pixels of the inner border may be record more than once if the region is not closed, e.g., a one-pixel-width curve.

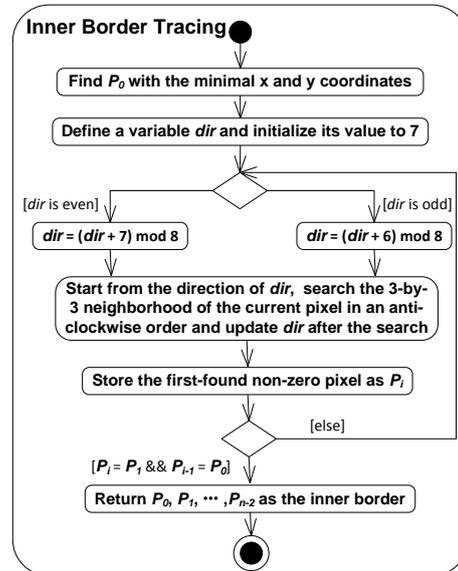


FIGURE 2 General structure of inner border tracing

### 2.2 INNER BORDER TRACING

The algorithms based on Hough transformation (HT) usually have high computation cost. Their most expensive step is the procedure called *voting* to change the values of points (cells) in the parameter space according to pixel coordinates *s* in the image foreground. Its computational complexity is  $O(m \cdot n^2)$  where *m* denotes the degree of the parameter space discretization and *n* denotes one dimension of an image.

The foot-of-normal algorithm could reduce the computational complexity of the voting to  $O(n^2)$  based on the fact that one line can only have one intersection with the normal which crosses the origin of the coordinate system. The intersections just are defined as votes as shown in Fig. 3.

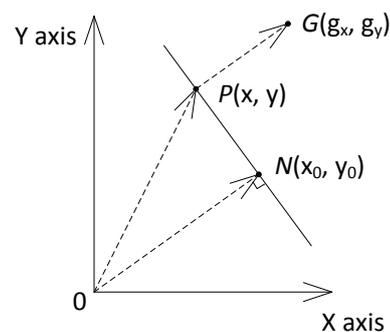


FIGURE 3 Foot-of-normal method

The origin is a fixed reference point. Assuming P is a point on the line and line  $\overline{ON}$  is the normal of line  $\overline{PN}$ , from the facts that vector  $\overline{ON}$  is perpendicular to vector  $\overline{PN}$  and  $\overline{ON}$  parallels to gradient vector  $(g_x, g_y)$  given by a Sobel operator, we get

$$g_x / g_y = y_0 / x_0 \text{ and } (x-x_0)x_0 + (y-y_0)y_0 = 0. \quad (1)$$

Solving the above formulae for  $x_0$  and  $y_0$ , we obtain the formula of the voting point,

$$x_0 = v \cdot g_x, \quad y_0 = v \cdot g_y, \text{ where } v = \frac{x \cdot g_x + y \cdot g_y}{g_x^2 + g_y^2} \quad (2)$$

Davies [27] analysed the line estimation error and found the image should be subdivided to reduce the error. The basic steps of the foot-of-normal method consist of subdividing the image, computing  $(x_0, y_0)$  in each sub-image whose origin is its centre instead of the upper-left corner, and finally making a vote at  $(x_0, y_0)$ . Fig. 4 shows the voting results of a real-word image. The original image is divided to 20-by-20 and 50-by-50 sub-images indicated by white lines. The squares in the resulting images denote the voting points. Obviously, the more the image is subdivided, the more lines can be detected.

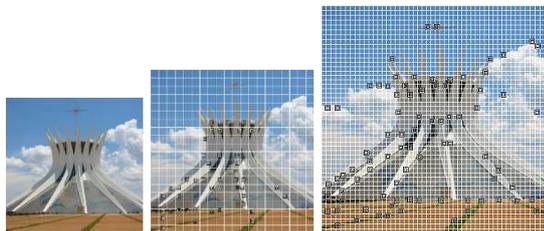


FIGURE 4 Results of the foot-of-normal method

### 3 The Proposed Algorithm

The basic idea of LSDSA is to dynamically link the pixels in the foreground of an image by using a statistical approach with low computational complexity. We develop a novel approach to represent the geometrical properties of line segments. Here, 8 infinite directions could be represented by direction values defined in section 2.1. The number of directions is determined by the resolution of the input image and the size of the neighbourhood employed. Among the neighbourhood of small sizes, a 3-by-3 neighbourhood reflects an adequate range of directions.

The general structure of LSDSA is shown in Fig. 5. It requires a pre-processing to generate the binary edge image based on the input image. The pre-processing is denoted by a solid rectangle labelled by the text “Detect Edge”. Any edge detector can be employed in the pre-processing.

In Fig. 5, the largest dotted rectangle denotes the body of LSDSA. In this rectangle, three subroutines are

designed to (1) find roughly straight-line segments, (2) find and break distorted line segments, and (3) link adjacent line segments with similar directions. They are represented by rectangles marked Subroutine 1, Subroutine 2 and Subroutine 3 in Fig. 5 and the sub steps of Subroutine 2 and Subroutine 3 are shown in their rectangles respectively.

When the detection of line segments is completed, the detected segments can be directly returned or processed by a post-processing to generate the global straight lines. As shown in Fig. 5, the two possible choices are denoted by two branches below the rectangle of the proposed algorithm. By combining the foot-of-normal method and the inverse HT algorithm [28, 29], the global lines can be obtained as the final result.

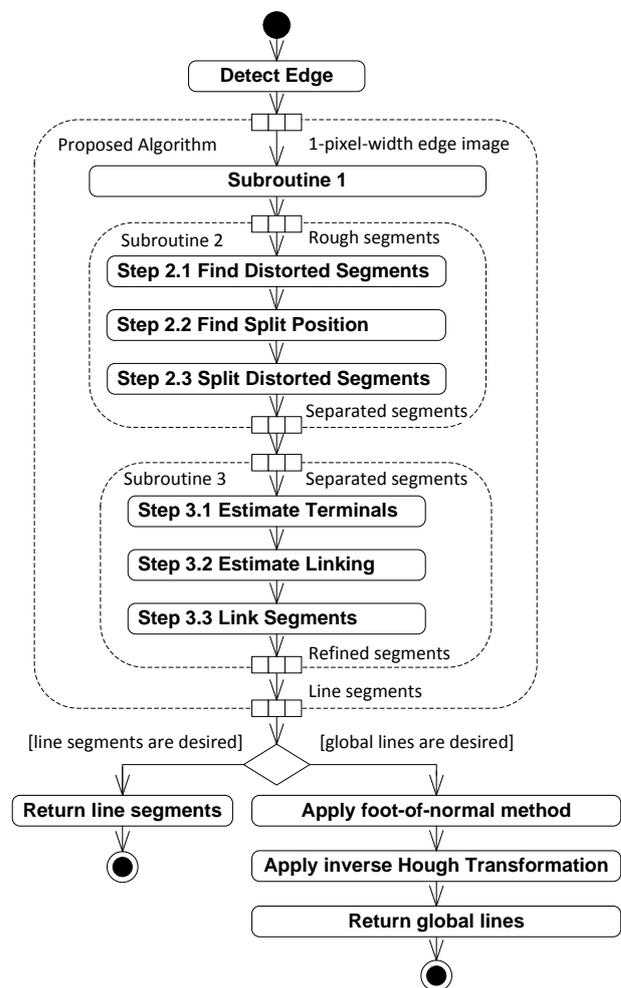


FIGURE 5 General procedure

#### 3.1 FINDING ROUGHLY STRAIGHT LINES

Subroutine 1 performs a scan of the edge image and conditionally tracks each continuous curve (edge). It is used to rapidly find roughly straight-line segments without using any voting process. The following

subroutines after Subroutine 1 do not need to scan the whole image, and the size of the data is reduced to segments found by Subroutine 1. The relatively sophisticated processing can be integrated to the following subroutines without drastically decreasing the speed of the algorithm. The details of Subroutine 1 are shown in Fig. 6.

In Fig. 6,  $m$  and  $minLength$  represent the sample size and the acceptable minimal length of the found line segment respectively, which are configured manually as input arguments. Variable  $toleranceVal$  is employed to distinguish the line segments of difference directions, e.g., if  $toleranceVal$  is set to 2, then direction values 0 and 1 are envisaged as representing the same direction because their difference is less than  $toleranceVal$ . The difference between two direction values is defined to be the smaller number of sectors between the direction values shown in Fig. 1. Namely, for direction values 0 and 3, the number of sectors can be anti-clockwise counted as 3 or clockwise counted as 5, and the difference is 3. The difference between two direction values such as  $dir1$  and  $dir2$  can also be expressed as

$$\text{if } |dir1 - dir2| > 4 \text{ then } difference = 8 - |dir1 - dir2|; \quad (3)$$

$$\text{else } difference = |dir1 - dir2|$$

When a non-zero pixel is found during Subroutine 1 scan, the scan is temporarily paused and the track operation is inserted. The track operation first tries to locate the right end of the curve. To locate the right end, the subroutine only checks the positions coinciding with direction values of 6, 7 and 0 of Fig. 1 in the neighbourhood of a non-zero pixel, moves the searching centre to it and repeat. The last-found pixel is envisaged as the right end of the line segment and the sampling starts at it. The searching procedure may split a line segment into several short segments, for example, if we apply the right-end searching to the right segment in Fig.7, it will stop before the pixel labelled 5 is reached because this pixel is in the direction 5 instead of 6, 7 or 0. This drawback may be fixed by merging segments in Subroutine 3.

$Segments$ ,  $Directions$  and  $Occurrences$  in Fig.6 respectively are the set of found line segments, the set of the direction values and the set of statistical information of direction values associated with each segment. Variables  $tempSegments$ ,  $tempDirections$  and  $tempOccurrences$  contain the temporary data of  $Segments$ ,  $Directions$  and  $Occurrences$  respectively. The data of three temporary variables can be arbitrarily overwritten.

To illustrate these steps clearly, we depicts a simple binary image whose foreground contains two line segments in Fig. 7. Table 1 illustrates the result of its sampling procedure. The sampling procedure is denoted by the solid rectangle marked "Sample  $m$  pixels" in Fig. 6.

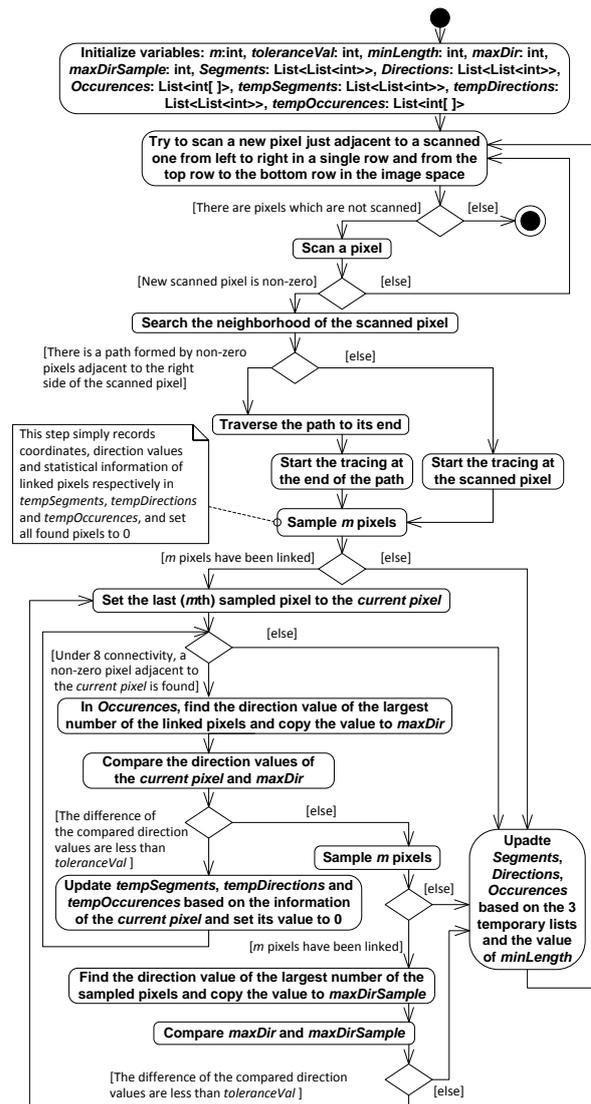


FIGURE 6 Details of Subroutine 1

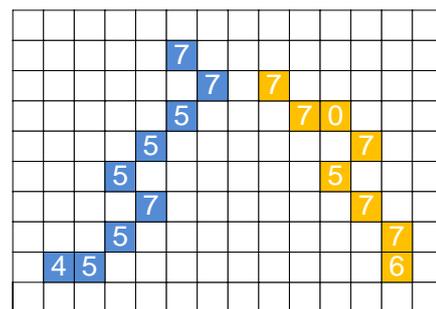


FIGURE 7 Representation of a binary image by lattices

TABLE 1 Values of Segments, Directions and Occurrences

Member	Value								
	0	1	2	3	4	5	6	7	8
$Segments[0]$	22	39	54	69	84	101	116	131	130
$Segments[1]$	41	58	59	76	91	108	125	141	∅
$Directions[0]$	7	7	5	5	5	7	5	5	4
$Directions[1]$	7	7	0	7	5	7	7	6	∅
$Occurrences[0]$	0	0	0	0	1	5	0	3	5
$Occurrences[1]$	1	0	0	0	0	1	1	5	7

### 3.2 DETECTING AND BREAKING DISTORTED LINES

In order to detect and break distorted lines, we launch three steps in Subroutine 2: (1) detect distorted segments; (2) determine positions of split; (3) splits distorted segments.

We detect distorted segments based on *Occurrences*. We find that the distribution of the direction values of a roughly straight line segments is similar with a uni-modal Gaussian distribution. Any multi-modal distribution implies the corresponding segment is distorted. Once the multi-modal is found, the difference between the direction values of the largest and the second largest modals is estimated. As shown in Fig. 8, we first detect the direction value of the second largest number of pixels which corresponds to a modal in the distribution in *Occurrences*, and then estimate the ratio and the difference of the maximal direction value and the found value which are respectively denoted by the variables *maxDir1* and *maxDir2*. Only when the difference exceeds *toleranceVal* and the ratio is larger than the parameter *thresholdRatio* whose value is experimentally set to 0.8, the corresponding segment is confirmed distorted.

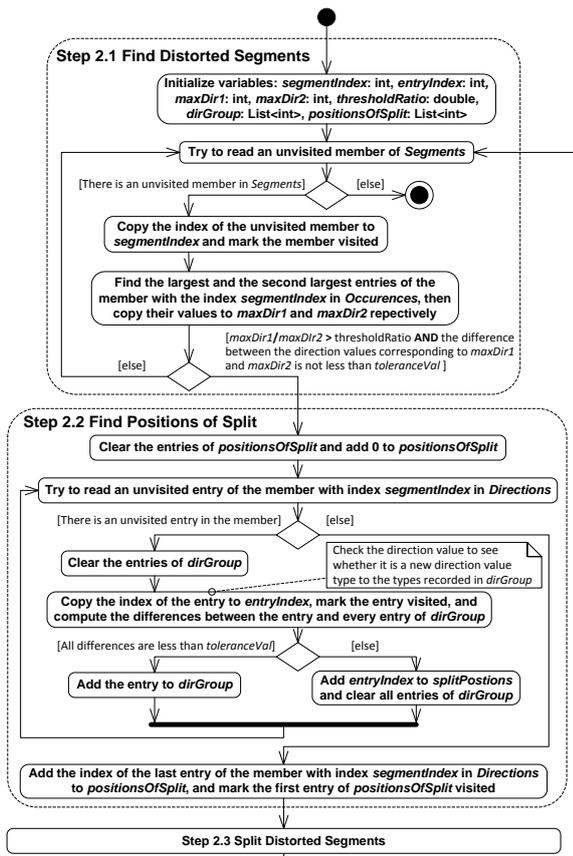


FIGURE 8 Details of the two steps of Subroutine 2

Next, in Step 2.2 we attempt to find the pixels where the main direction of a segment drastically changes. The strategy is assuming the segment is formed by several parts associated with specific direction value groups

represented by *dirGroup* in Fig. 8. For each group, the differences among all elements are less than *toleranceVal*. Initially, *dirGroup* is empty and then the direction value of the first pixel encountered in the checking is added to it. The direction values of the following pixels are compared with the value in group. If the difference is adequate, the pixel is taken into the part, otherwise it is marked as a possible position in the segment to split and recorded by *positionsOfSplit* in Fig. 8.

Step 2.3 shown in Fig. 9 tries to split the segment according to *positionsOfSplit*. The lengths of different parts of a segment are compared with *minlength*. Only a segment with parts of lengths exceeding *minLength* is split and causes *Segments*, *Directions*, and *Occurrences* to be updated. The updating procedure is denoted by four solid rectangles just above the thick horizontal bar in Fig. 9.

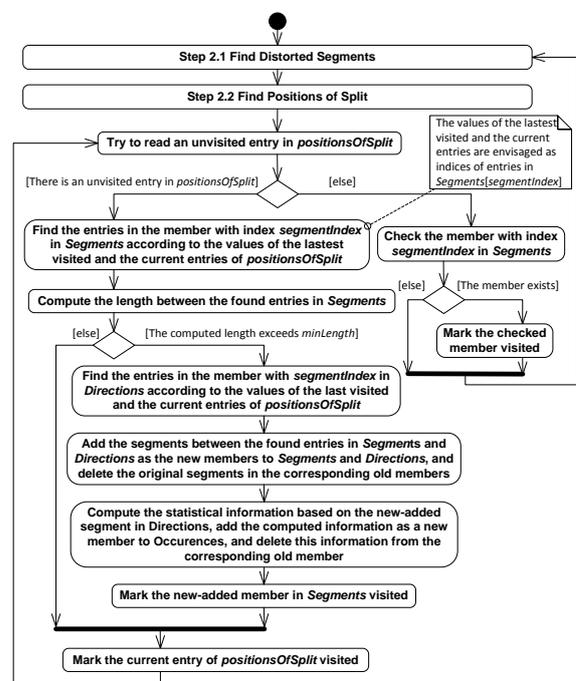


FIGURE 9 Details of the last step of Subroutine 2

### 3.3 LINKING ADJACENT LINES WITH SIMILAR DIRECTIONS

In Subroutine 3, we attempt to merge segments by appropriately extending the segments and comparing the directions with other segments found adjacent to the extensions. It composes three steps: Step 3.1 estimate terminals, Step 3.2 estimate linking and Step 3.3 link segments as shown in Fig. 5.

Step 3.1 locates the geometrical ends of segments for the subsequent extending operation. Although Subroutine 1 starts the traces from the right ends of segments, the first saved pixels may not be the right end because Subroutine 2 split segments and some of the first saved pixels of these segments are not the right end. The details

of finding geometrical ends are shown in Fig. 10. The ends can be easily found by estimating the directions of segments, i.e., if the maximal direction value is one of 5, 6, 7 and 0 ( $1 \leq \text{maxDirection} \leq 4$ ), the first recorded pixel is the geometrical right terminal; otherwise the first one is the left terminal.

After identifying all ends, we will estimate the path for extending in Step 3.2. Since the values of all entries in *Directions* range from 0 to 7 and these entries can reflect the geometrical shapes of segments basically, the entries may be envisaged as strings of letters 0 to 7 and their patterns can thus be found using approximating string matching [21]. Here we consider a simple but efficient strategy. We observe that the distribution of direction values around the middle point of a segment always follow a certain pattern. Therefore, we construct a path by repeatedly copying the middle direction values as shown in Fig. 11.

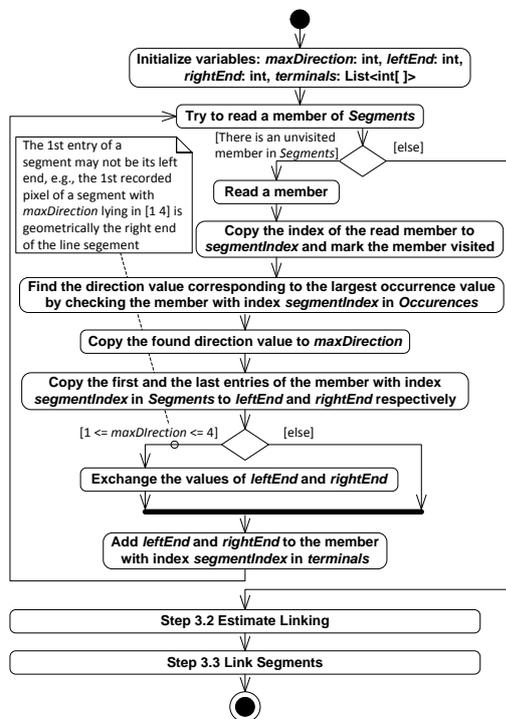


FIGURE 10 Details of Step 3.1 in Subroutine 3

Since Subroutine 1 traverses from one end of a line segment to the other, the recorded direction values reflects the pattern along the direction approximated by maximal direction value. Hence, the found path is only appropriate for one end, and inverses for the other. For instance, two red lattices in Fig. 12 indicate two ends; the colourful lattices between the ends denote the body of a line segment and the grey pixels form the two paths. The lattices of the segment except the paths are marked by their direction values obtained by tracing from the geometrical right terminal. The yellow pixel lies in the middle of the segment. The path adjacent to the left terminal, denoted by the variable *leftPath* in Fig. 11, is obtained directly by copying the direction values of

lattices around the middle lattice. The path adjacent to the right terminal is obtained by reversing the direction values of the left path consisting of direction values 5, 3, 5, 4 and 4; the right path contains the corresponding inverse direction values 1, 7, 1, 0 and 0. The inverse path is denoted by *rightPath* in Fig. 11.

When moving along a path, we could check the neighbourhood of the moving centre to find other segments but it is computationally expensive. We employ an alternative approach by checking the ends of segments whose identities lying in a range with the identity of the current segment as centre and parameter *detectRadius* in Fig. 11 as a radius. This is because the difference between identities of two segments in *Segments* partially reflects their geometrical distance in image space. The *detectRadius* is set to 20 in our experiment.

After the connection information of all segments is collected by Step 3.2, Step 3.3 showed in Fig. 12 checks each member of *Connections* to perform the merging. If the statuses of ends indicate valid linking, it will simultaneously merge the associated segments with the current one and checks the statuses of the merged segments to see whether they can be further merged with other ones. This iterative procedure will stop until the statuses of both ends are null. For instance, segment 1, 3 and 4 will be merged together according to *Connections*.

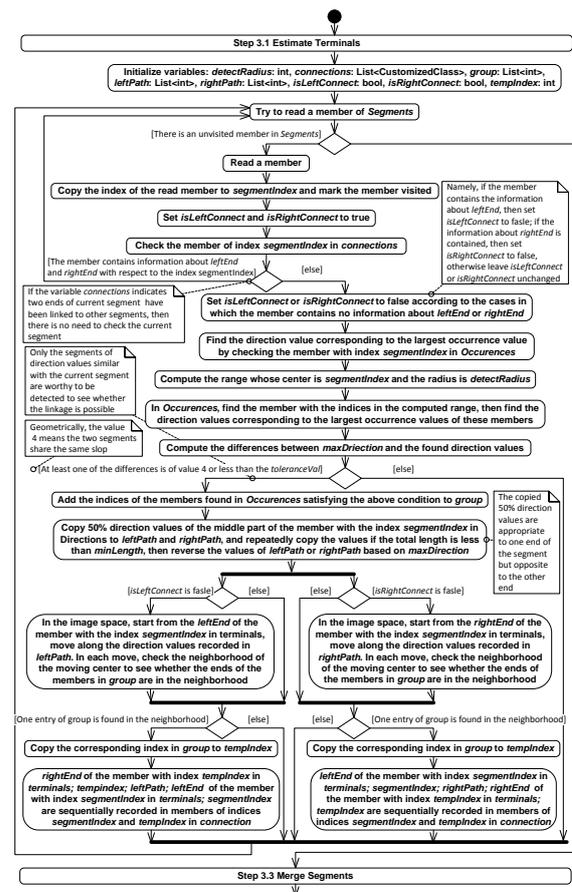


FIGURE 11 Details of Step 3.2 in Subroutine 3

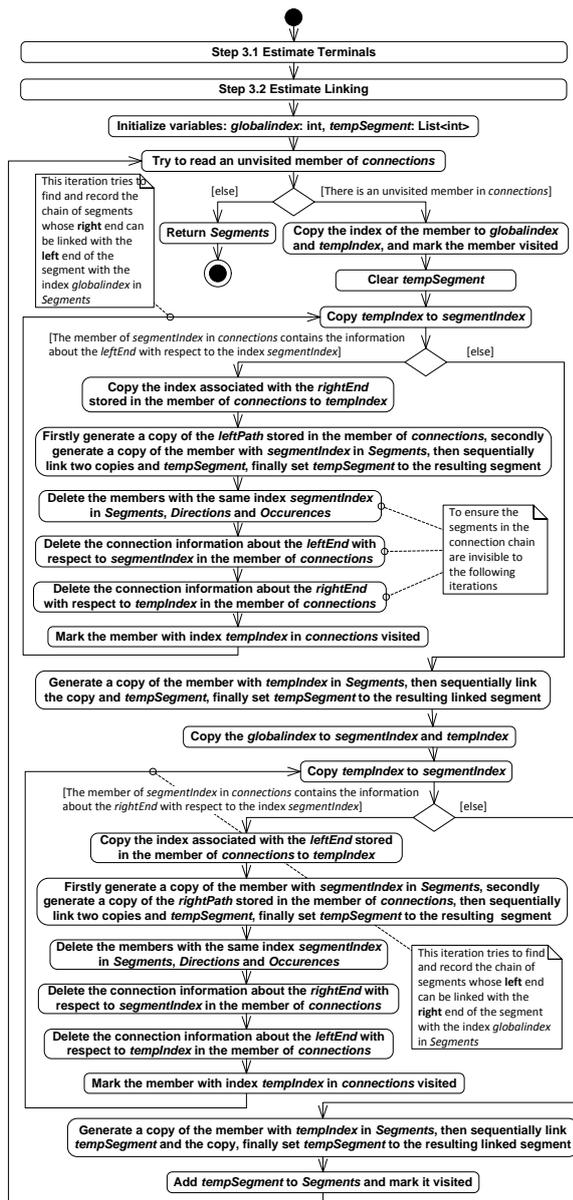


FIGURE 12 Details of Step 3.3 in Subroutine 3

### 4 Experimental Results and Discussion

In order to evaluate the performance of LSDSA, we test KHT [13], EDLines [7], TODIS [22] and LSDSA using two corpora of artificial and the real-world images. We developed an application package [30] for edge detection using C#. In general, LSDSA achieves performance similar to TODIS, which is a better result than KHT and EDLines. However, the computation cost of LSDSA is obviously lower than TODIS.

#### 4.1 CASE COMPARISONS

Figure 13 lists some samples of line detections using LSDSA and KHT. We use colourful lines to display the segment results of LSDSA to distinguish the distinctive

parts of a continuous curve. For example, the circle on the top of the building in the sub image labelled as H10 of Fig. 15 is indicated in H12. At the first glance, the indicated circle may seem to be a false detection. Actually, the detected circle in H12 is denoted by two kinds of colours (blue upper and purple lower). Thus, two roughly straight line segments are detected by the algorithm instead of a circle. The values of parameters  $m$  and  $minLength$  are set to 4 and 30 respectively according to the resolution of the test images.

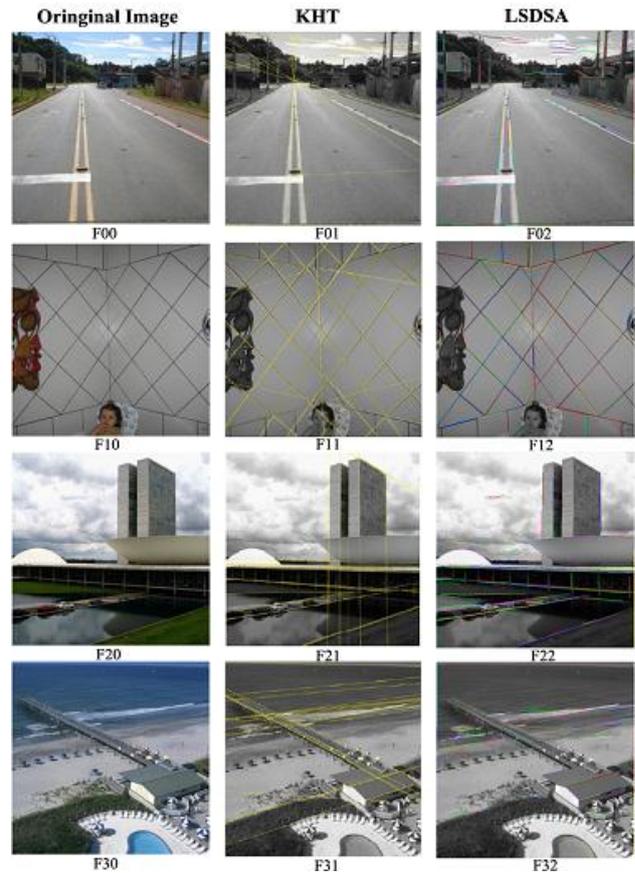


FIGURE 13 LSDSA compared with KHT

Fig. 14 compares the lines detected by LSDSA and EDLines. The parameters  $m$  and  $minLength$  of LSDSA are set to 3 and 8 respectively. All line segments detected by EDLines are marked by solid lines in the second column. Although the difference between LSDSA and EDLines is not as large as the difference between KHT and LSDSA, LSDSA can discover more significant line segments that are ignored by EDLines, such as the seams on the face of the building in G10 and the edges of the front doors of the house in G20. We also note that LSDSA can detect a line segment about the lady's right arm in G32, which is the border between the highlight and the lowlight areas of the arm surface, and its curvature continuously changes. However, EDLines fails to detect it.

Although TODIS is the slower than EDLines, KHT and LSD, it exhibits the best accuracy. Generally, LSDSA achieves accuracy as good as TODIS as shown in

Fig. 15. Almost all line segments detected by TODIS in H01 are indicated by LSDSA in H02 as well, except the broken vertical line on the left. Similar results are shown in H11 and H12. LSDSA successfully detects the traces in the right corner of H10 and the edges of the circle on the top of the building. Note, the detected edges of the circle are indicated by two colours (the blue upper/lower curves). The two curves approximate two roughly straight line segments. It says that LSDSA can detect roughly straight line segments, and further example can be found in G32 of Fig. 14. The detection of LSDSA is clearly more accurate than TODIS inasmuch as the edges of windows in the top of the skyscraper are detected in H22 but not in H21. The edges of the white circular lights on the top of the hallway are indicated in H32 and ignored by TODIS in H31. It suggests that LSDSA can detect the roughly straight line segments correctly.

when the current pixel is non-zero and set all traced pixels to 0.

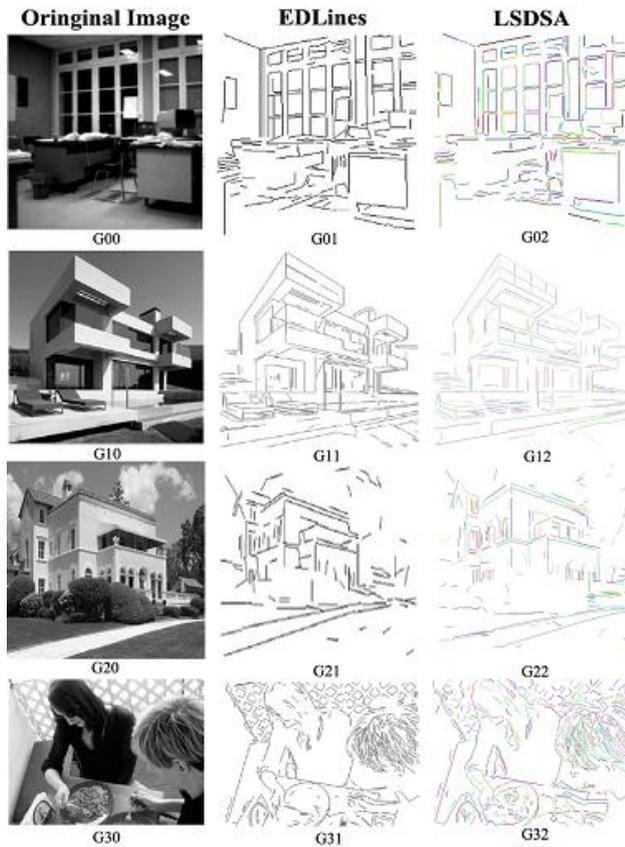


FIGURE 14 LSDSA compared with EDLines

4.2 COMPUTATIONAL COMPLEXITY

The computation cost of LSDSA is determined by Subroutine 1 for finding roughly straight lines. The time complexity of this part is  $O(ls)$  where  $l$  is the largest length of a line segment and  $s$  denotes the number of segments. Unlike TODIS with  $O(n^2)$  where  $n$  represents one dimension of the image space, LSDSA consume less time obviously when original images are large. This is because the first subroutine only performs the tracing

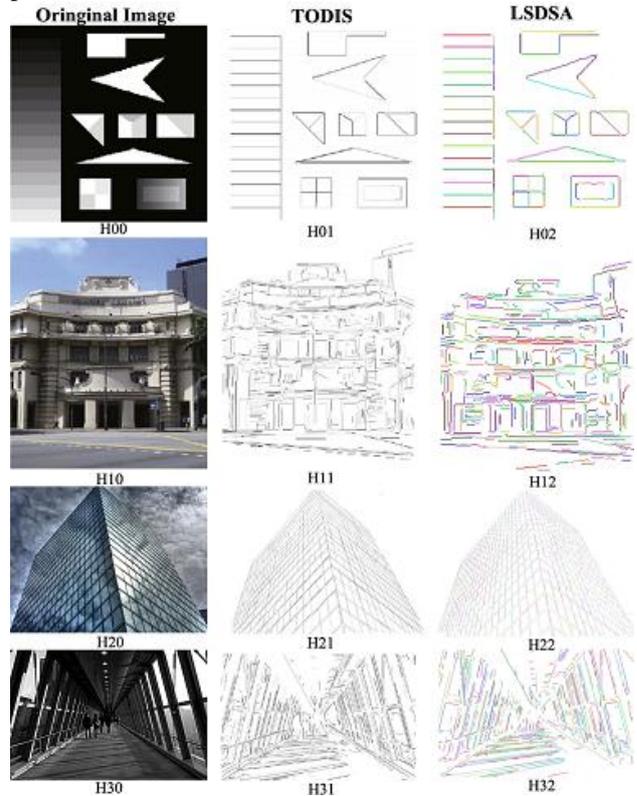


FIGURE 15 LSDSA compared with TODIS

Furthermore, we find a close relation between  $l$  and  $s$ . If the length of a line segment  $l$  is very large, then the number of segments  $s$  must decrease. Since the longer the segments are, the more space in the image they will occupy, and less space are left to the rest segments. It suggests that fewer segments can be released in this concise space. Conversely, if there are many segments, then their lengths tend to be short. Therefore, a balance exists between  $l$  and  $s$ .

The complexity of Subroutine 2 is  $\text{Max}(O(s), O(ld))$  where  $d$  is the number of the distorted line segments found in this subroutine by performing a searching of complexity  $O(s)$ . Since  $d$  can't exceed  $s$ , the worst case of the second subroutine is  $O(ls)$  as same as the first subroutine.

The third subroutine is of complexity  $O(s)$ .

Theoretically, the time complexity of the proposed algorithm is  $O(ls)$ , while the complexity of TODIS is at least  $O(n^2)$  according to the analysis of the BU-Scan procedure [22], which contributes only a half of the computational cost of TODIS. Practically, we compare the time consumed by KHT, EDLines, LSDSA and TODIS under different conditions. The time consumed by algorithms do not consist the pre-processing as edge detection and post-processing as inverse Hough Transformation.

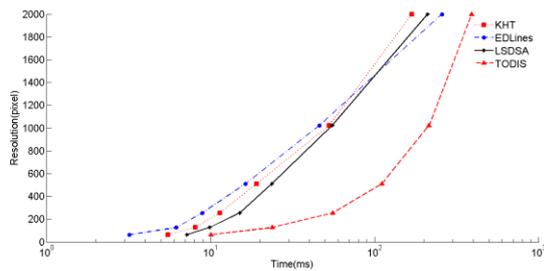


FIGURE 16 Time consumptions with respect to few lines

Fig. 16 shows time spent on processing the image containing few lines when the resolution ranges from 0-by-0 to 2000-by-2000 in pixel. Note the unit of x-axis is logarithm of the time, hence the crooked curve indicates a roughly straight line in a plane with non-logarithm axes. In Fig. 16, although TODIS consumes more time than other algorithms, its curve implies the ratio of resolution and time is linear (even the slop is very large). KHT, EDLines and LSDSA all share a similar pattern when resolution is lower than 1200 pixel in Fig. 16, but their curves lead to different destinations when the resolution reaches 2000 pixel. This implies their consumed time may be quite different when the input data becomes very large. In the level of 2000 pixel, LSDSA consumes less time than EDLines and TODIS. This illustrates the capability of LSDSA to reduce the dimension of input data.

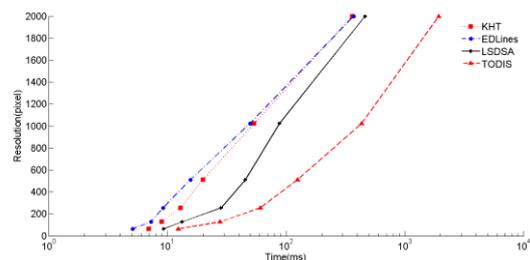


FIGURE 17 Time consumptions with respect to moderate lines

Fig. 17 shows the time consumed to process image moderate lines when resolution is changing, As in Fig. 16, TODIS apparently consumes more time than other algorithms, and EDLines and KHT share a similar pattern. The curve of LSDSA follows the pattern when resolution is higher than 1000 pixel, but still ends behind EDLines.

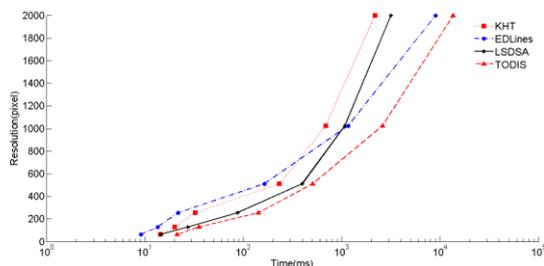


FIGURE 18 Time consumptions with respect to many lines

Fig. 18 illustrates the time when the processed image contains lots of lines. All algorithms even TODIS follows a similar pattern when resolution is lower than 600 pixel. As the resolution increases, EDLines and TODIS are exceeded by KHT and LSDSA. In level of 2000 pixel, there is a large gap between the group of algorithms with capability of reducing input dimension, i.e., KHT and LADSA, and the group of algorithm with no such capability, i.e., EDLines and TODIS.

Generally, when the processed image is small, there are little difference among the time consumed by KHT, LSDSA, EDLines and TODIS. As the resolution increases, the number of lines contained by image becomes an important factor affecting the consumed time. If lines are few, the difference may still remain small even when resolution increases. If lines are many, there will be an obvious gap between the time consumed by algorithms with or without capability of reducing input dimension. Hence, the feature of LSDSA illustrated by Fig. 16 to Fig. 18 is the capability of fast processing images with high resolution and complicated content.

### 5 Conclusions and Future Work

In our proposed method, LSDSA employs a statistical tracing strategy to reduce the dimension of the input data and distinguish the distorted segments by analysing the distributions of direction values which are approximate quantified values of geometrical directions in image space. LSDSA collects statistical data of quantified directions, so it is able to achieve higher speed under the condition of good accuracy. We report the experiment results of LSDSA using test images and industrial images, and compare its performance with typical LSD algorithms such as KHT, EDLines and TODIS. It indicates that the accuracy of LSDSA is clearly better than KHT and EDLines, and it is as good as TODIS. But LSDSA consumes much lower computation cost than TODIS does.

In order to further increase the accuracy of LSDSA, we plan to explore more refined direction values and large neighbourhood in the following research. At the same time, we will investigate how to simplify its processing procedure and proposed algorithm.

### Acknowledgments

This work is supported by National Natural Science Foundation of China under Grant 61103172. The authors would like to thank Geyong Min of Bradford University in UK and Shiren Ye of Changzhou University in China for their comments and insightful suggestions.

## References

- [1] Mahadevan S, Casasent D P 2001 Detection of triple junction parameters in microscope images *Proc. SPIE* **4387** 204-14
- [2] Kahn P, Kitchen L, Rieseman E M 1990 A fast line finder for vision-guided robot navigation *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(11) 1098-102
- [3] Tupin F, Maitre H, Mangin J F, Nicolas J M, Pechersky E 1998 Detection of linear features in SAR images: application to the road network extraction *IEEE Trans. Geosci. Remote Sens.* **36**(2) 434-53
- [4] Zhu Y, Carragher B, Kriegman D J, Milligan R A, Potter C S, et al 2001 Automated identification of filaments in cryo-electron microscopy images *Journal of Structural Biology* **135**(3) 302-12
- [5] Yu X, Lai H C, Liu S X F, Leong H W 2005 A gridding Hough transform for detecting the straight lines in sports videos *In Proc. Int. Conf. on Multimedia and Expo., Amsterdam, The Netherlands, 6-8 Jul. 2005* 45-8
- [6] Zheng Y, Li H, Doerman D 2005 A parallel-line detection algorithm based on HMM decoding *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(5) 777-92
- [7] Akinlar C, Topal C 2011 EDLines: A real-time line segment detector with a false detection control *Pattern Recognition Letters* **32**(13) 1633-42
- [8] Hough P V C 1962 *Method and Means for Recognizing Complex Patterns* U.S. Patent No. 3069654
- [9] Aguado E, Montiel A S, Nixon M S 2000 On the intimate relationship between the principle of duality and the Hough transform *Proc. Roy. Soc. A* **456**(1995) 503-26
- [10] Wright M, Fitzgibbon A, Giblin P J, Fisher R B 1996 Convex Hulls, Occluding Contours, Aspect Graphs and the Hough Transform *Image Vision Comput.* **14**(8) 627-34
- [11] Bhattacharya A, Rosenfeld A, Weiss I 2003 Geometric and Algebraic Properties of Point-to-line Mappings *Pattern Recogn.* **36**(2) 483-503
- [12] Bhattacharya A, Rosenfeld A, Weiss I 2002 Point-to-line Mappings as Hough Transforms *Pattern Recogn. Lett.* **23**(14) 1705-10
- [13] Fernandes L A F, Oliveira M M 2008 Real-time Line Detection Through an Improved Hough Transform Voting Scheme *Pattern Recognition* **41**(1) 299-314
- [14] Pope A R, Lowe D G 1994 Vista: A software environment for computer vision research *In Proceedings of Computer Vision and Pattern Recognition, Seattle, WA, USA* 768-72
- [15] Lowe D G 1987 Three-dimensional object recognition from single two-dimensional images *Artificial Intelligence* **31** 355-95
- [16] Burns J B, Hanson A R, Riseman E M 1986 Extracting Straight Lines *IEEE Trans. Pattern Anal. Mach. Intell.* **8**(4) 425-55
- [17] von Gioi R G, Jakubowicz J, Morel J M, Randall G 2010 LSD: A fast line segment detector with a false detection control *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(4) 722-32
- [18] Etemadi A 1992 Robust Segmentation of Edge Data *In Proc. Int. Conf. on Image Processing and Its Applications, Maastricht, the Netherlands, 7-9 Apr 1992* 311-4
- [19] Debled-Renessona I, Feschet F, Rouyer-Degli J 2006 Optimal Blurred Segments Decomposition of Noisy Shapes in Linear Time *Computers & Graphics* **30**(1) 30-6
- [20] Kerautret B, Even P 2009 Blurred Segments in Grey Level Images for Interactive Line Extraction *In Proc. Int. Conf. on Combinatorial Image Analysis, Playa del Carmen, Mexico, 24-27 Nov. 2009* 176-86
- [21] Skiena S S 2008 *Set and String Problems In The Algorithm Design Manual*, 2nd ed. NY: Springer, Ch. 18, sec. 4, 631-6
- [22] Yang K, Ge S S, He H 2011 Robust Line Detection Using Two-orthogonal Direction Image Scanning *Computer Vision and Image Understanding* **115**(8) 1207-22
- [23] Shapiro V 2006 Accuracy of the Straight Line Hough Transform: The non-voting approach *Computer Vision and Image Understanding* **103**(1) 1-21
- [24] Sonka M, Hlavac V, Boyle R 2008 *Image Processing, Analysis, and Machine Vision* 3rd ed. CT, USA: Cengage Learning
- [25] Jacobson I, Booch G, Rumbaugh J 2005 *Unified Modelling Language User Guide* 2nd ed. MA, USA: Addison-Wesley
- [26] Rumbaugh J, Jacobson I, Booch G 2010 *Unified Modelling Language Reference Manual* 2nd ed. MA, USA: Addison-Wesley
- [27] Davies E R 2005 *Machine Vision: Theory, Algorithms, Practicalities* 3rd ed. CA, USA: Morgan Kaufmann
- [28] Anastasios L, Kesidis A L, Papamarkos N 1999 On the inverse Hough transform *IEEE Trans. Pattern Anal. Mach. Intell.* **21**(12) 1329-43
- [29] Kenzie D S, Protheroe S R 1990 Curve description using the inverse Hough transform *Pattern Recogn.* **23**(3-4) 283-90
- [30] Jia L, Sun Y, Wang M, Gu Y 2011 A Research on Implementation of Image Scattergram by Using C# *In Proc. Int. Conf. on System Design and Data Processing, Tai Yuan, China, 26-28 Feb. 2011* 353-5

## Authors

	<p><b>Liang Jia</b></p> <p><b>Current position, grades:</b> is a faculty member at Changzhou university of Jiangsu province in China.</p> <p><b>University studies:</b> MS degree in computer science from the Nanjing university of Science and Technology in 2009 and the BS degree in computer science from Beifang university of Nationalities in 2004.</p> <p><b>Scientific interests:</b> image processing, computer vision, service-oriented software development.</p>
	<p><b>Nigang Sun</b></p> <p><b>Current position, grades:</b> joined East China University of Science and Technology in 2007 and Changzhou University in 2010. Now he is an associate professor in the department of the Computer Science and Technology.</p> <p><b>University studies:</b> Ph.D. degree in information security, University of Chinese Academy of Sciences, 2007.</p> <p><b>Scientific interests:</b> formal methods, system modelling and analysis, information security.</p>

# A model-based assurance case construction approach for system control software

**Dajian Zhang\*, Minyan Lu, Nan Wu**

*School of Reliability and System Engineering, Beihang University, Beijing, P. R. China*

*Received 12 June 2014, www.tsi.lv*

---

## Abstract

As the massive damage caused by the failures of system control software becomes increasingly prominent, people pay more attention to the construction of assurance case to demonstrate the dependability level of system control software. In this paper, a new assurance case construction approach for system control software is proposed. Based on the metamodel of modular GSN, we give the basic procedure and tree structure deductive algorithm of the approach, and verify our work using Brake Control software used in an aircraft. The results show that the approach can develop assurance case effectively and efficiently.

*Keywords:* Software, Dependability, assurance case, GSN, modularization

---

## 1 Introduction

With the wide deployment of software in critical control systems whose potential failure may cause huge damage, the dependability of the control software has become a major factor for proper system operation. Therefore, it is of great importance to study on the dependability assurance of this software. Demonstrating the expected dependability properties of this software to provide sufficient confidence for potential users is a key issue [1]. Traditional software development and certification are generally based on a prescriptive standards, such as DO-178B [2], IEC 15608 [3], ISO/IEC 15408 [4-6]. However, these certification approaches have some deficiencies such as unclear rationale underlying some process activities, lack of organization between evidence, highly prescribed technical activities [7-8]. Therefore, a new goal-based assurance case approach is proposed [9]. Through a clear argument structure and flexible and effective organization of evidence, it can demonstrate the system meet its original requirements in an explicit and structured way. This approach can overcome the deficiencies of traditional approaches and receives growing attention [10].

Assurance case is originally used in the safety area [11], and gradually extended to other dependability area [12-14]. It is defined as “a documented body of evidence to provide a compelling justification that the systems performing a specific task satisfies specified critical properties in a specific environment” [15]. An assurance case generally consists of three elements: claim, argument, and evidence. How to represent the structure of argument efficiently and concisely is a key problem in assurance case research area. Many approaches are proposed [16-17]. Goal Structuring Notation (GSN) is a

popular one among these approaches [18]. It combines rich graphical notations with modularization thinking to present argument in an intuitive, explicit way. It can clearly exhibit the logical relationship between product-oriented and process-oriented evidence by establishing the argument structure model and can be used in qualitative or quantitative analysis to achieve the evaluation of software dependability level [19]. However, the rich elements in GSN also bring confusion in its usage. Because of a lack of guidance on how to use this powerful tool systematically and unambiguously, the modular elements are often abused or misused by developers of assurance case. Therefore, the result of the argument is strongly influenced by subjective factors, which leads to the low effectiveness and the reduction of confidence placed in the argument conclusion.

This paper proposes a structured development approach for modular GSN in order to guide the construction of assurance case for control software. Based on the analysis of the GSN modular argument elements, an extended GSN metamodel is proposed. According to the metamodel, we give the progression algorithm for constructing the core structure of assurance case and the standard procedure of the argument construction. We illustrate our contributions by application to a Brake Control software system. The results show that our approach can provide explicit guidance and help to standardize the development process of assurance case. It also reduces subjectivity and ambiguity in the process, and improves the effectiveness.

---

\* *Corresponding author* e-mail: greatdijz@163.com

## 2 Model-Based Assurance Case Construction Approach

### 2.1 GSN MODEL MODULARIZATION

On the basis of in-depth study of GSN basic concepts, we propose a modular GSN meta-model, as is shown in figure 1-4.

Figure 1 shows the correlation of macroscopic concepts in assurance case. As is shown above, every

assurance case has an assurance subject. Assurance case varies from safety assurance, security assurance, reliability assurance, etc. Every assurance subject has a central assurance objective, this assurance objective is proposed in the form of claim, and is the top-level objective of assurance case. An assurance case consists of three basic parts: claim, argument and evidence. Assurance case is presented in the form of structured argument and GSN is one of these arguments.

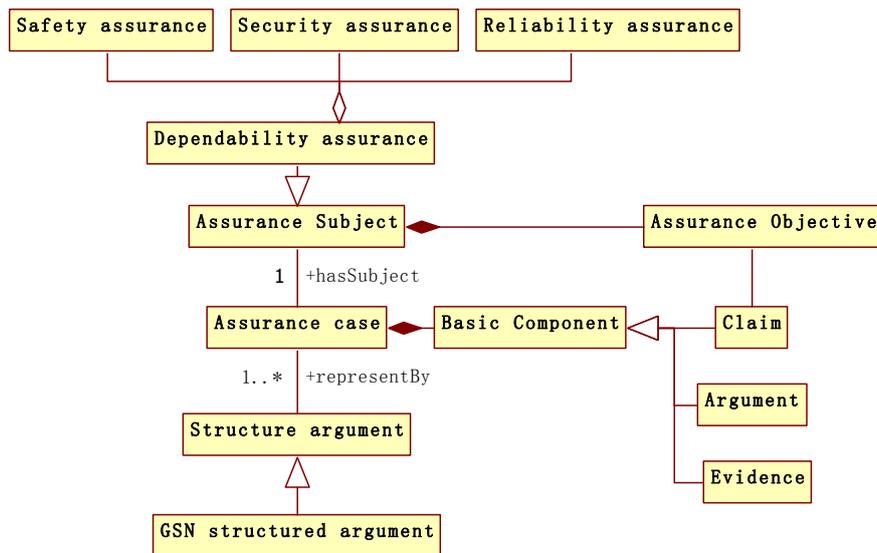


FIGURE 1 Macroscopic concept of assurance case

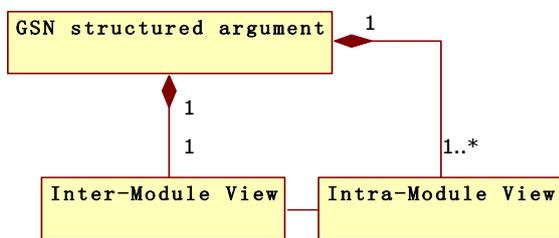


FIGURE 2 Macroscopic composition of GSN structured argument

Figure 2 shows the macro composition of GSN structured argument. GSN structured argument can be treated as two abstract granularities: the macroscopic and abstract argument between modules, and specific inter-module argument. In general, for relatively simple software systems, conducting fine-granularity argument using basic GSN nodes can meet the requirement. However, when we are arguing a complex large scale system, adopting basic argument structure will make the argument structure too big and complicated to manage, especially when a system consists of many modules. Introducing coordination mechanism of two-level abstract of module view and inter-module view can deal with this problem. Module view displays the relationship between modules, it shows the overall argument structure in a higher level of abstraction. Every module in the module view represents a specific argument structure, and an inter-module view in parallel.

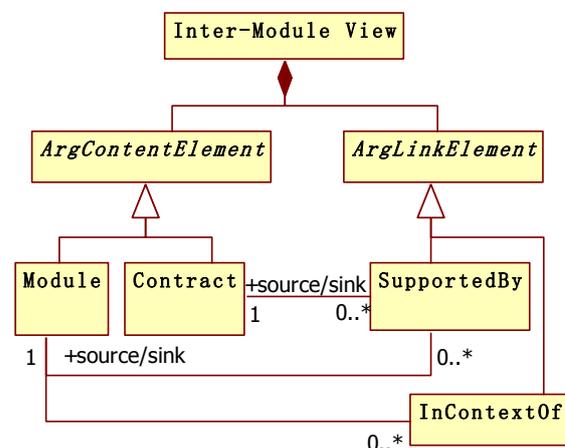


FIGURE 3 Basic concepts and correlation of module view

Figure 3 shows basic concepts and correlation of module view. Module and contract are basic elements in module view, module is a high-level abstraction of specific argument structure, a contract shows the relationship between modules and defines how a goal in the module is supported by the argument in another module. Modules and contracts are connected by “supportedby” and “incontexof” elements to build the macro view of the argument.

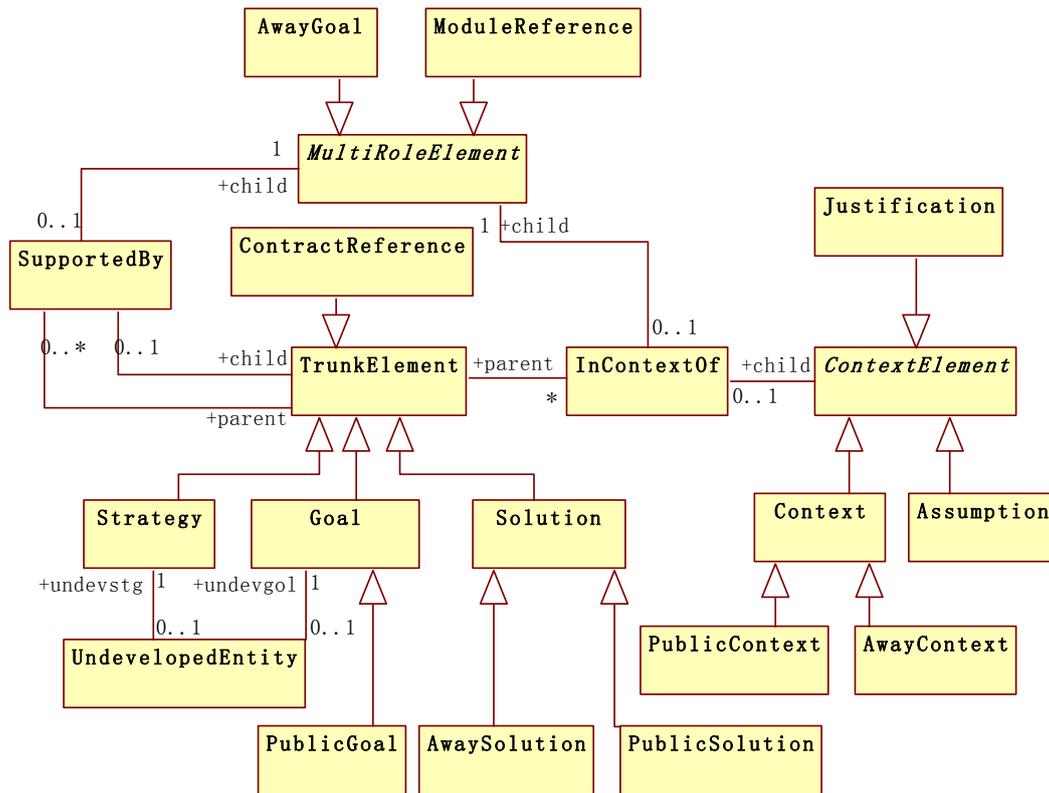


FIGURE 4 Concept and correlation of argument view

Figure 4 shows the concept and correlation of specific argument view. Elements in specific argument view can be divided into standard GSN elements and extended modularization elements. In standard GSN elements, goal, strategy and solution forms the backbone element of argument, undeveloped bodies provides effective support for the development phased in argument through the attachment to Goal and Strategy; context, assumption and justification are ancillary elements providing background information for the argument, “supportedby” and “incontextof” elements are connectors in the argument, “public elements” (public goal, public solution, public context) and “away elements” (Away solution, away the context and away goal) correspond each other, and together they provide a general mechanism for the share of inter-module argument elements. “Public elements” are open external interface of the argument, indicating these elements can be referenced by other argument modules. When referenced by other argument modules, they must be presented in the form of “away elements”. Module reference elements and contract reference elements corresponds module elements and contract

elements, respectively. They can be considered as the projection of module and contract element in the specific argument view. The roles different elements play in the argument are different, contract reference element can be considered as a backbone element, “modulereference element” is more special. Similar with “away elements”, they come in a wide variety of roles in the argument, they can be referred to as the backbone elements in argument and reasoning, and they can provide background information for argument and reasoning, so in this paper we will define these as “MultiRoleElement”.

## 2.2 MODEL-BASED GSN ARGUMENT CONSTRUCTION APPROACH

Based on the meta-model given above, we define a Model-Based GSN Argument Construction Approach. Because the modular GSN argument includes two levels: macro and micro argument, we adopt the principle of building the argument from macro to micro, and then back to macro.

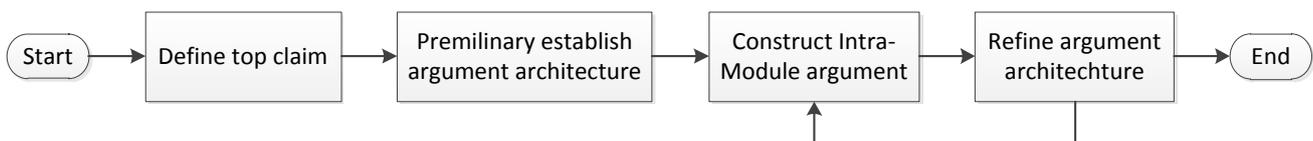


FIGURE 5 Basic procedure of modular GSN Construction Approach

Figure 5 shows the basic procedure of modular GSN construction approach. Firstly, define the top-level claim according to our argument subject, which is the ultimate goal of our argument. Then establish the preliminary overall argument structure according to top-level claim. We can establish the argument structure from different angles according to the features of top-level claim and evidence acquired. Through the investigation and analysis of the literature on assurance case, we found some typical argument structures, such as functional arguments, workflow process arguments, life cycle arguments, system structure arguments and risk mitigation arguments, etc. Argument structure is shown in the form of module view. This is only the preliminary established macroscopic module view structures, but the specific interface between modules is difficult to specify at the moment, specific argument structure refinement is needed to establish step by step. Next, we need refine each module one by one. According to the meta-model given in the figure above and the concepts related to instantiation, we can gradually establish argument using top-down approach. Along with the unfolding of argument structure, modular elements are added to the specific structure of argument. The introduction of modular elements means that the shared interface between this module and other modules are changed. In this case, we need to backtrace the overall argument, change the module view accordingly, which is step 4. This process might involve two cases: 1. the share interface are changed between this module and an existing module, this case is likely to cause changing relations between the two modules in the module view. 2. This module has an interaction with a module excluded in the module view, in this case, we need to create a new module in the module view to reflect this change.

Establishing the specific argument view, namely step 3, is the focus of the modular GSN argument construction. The introduction of modular elements greatly enhanced the expression ability of GSN, however, the abuse of modular elements can lead to the unclear role of elements, bringing chaos in the argument view structure and further influence the macro module view. These will bring difficulties to users to understand and communicate on assurance case. This paper defines a modular development approach using formal methods. This approach can clearly demonstrate the role and the timing of use of each modular element in constructing the argument, eliminating the ambiguity on understanding, and implements a systematic construction of argument.

First, we will define some primitives to describe the process, through the combination of these primitives we can describe the establishment process of the argument. The process describing primitives are as followed:

#### Declare primitives:

Declare(SetInstance, ConceptName)

- SetInstance {instanceName<sub>1</sub>, instanceName<sub>2</sub>, ..., instanceName<sub>n</sub>}, instanceName<sub>i</sub> is the name of the instance;

- ConceptName is the name of the concept;  
This primitive declares the instance set "SetInstance" of the concept "ConceptName"

#### Relationship defining primitives:

Define(SetInstance, LinkType, SourceInstance)

- SetInstance {instanceName<sub>1</sub>, instanceName<sub>2</sub>, ..., instanceName<sub>n</sub>}, instanceName<sub>i</sub> is the name of the instance;
- LinkType is enumeration type variable, enum{Supportedby, InContextof};
- SourceInstance is the name of the instance;

This primitive defines the relationship between instance set "SetInstance" and instance source "SourceInstance". If SetInstance  $\neq \emptyset$ , then this primitive defines the LinkType argument relationship from SourceInstance to SetInstance; if SetInstance =  $\emptyset$ , then this primitive does not define any relationship.

#### Judge primitives:

ifContextNeeded(InstanceName)

Determine if instance "InstanceName" needs background information

ifAssumptionNeeded(InstanceName)

Determine if instance "InstanceName" needs assumption information

ifJustificationNeeded(InstanceName)

Determine if instance "InstanceName" needs judge information

ifDecomposeNeed (InstanceName)

Determine if instance "InstanceName" needs to be decomposed

Based on the primitives above, we put forward a GSN modular constructing process as followed:

BEGIN

Declare({topGoal}, Goal)

Step\_DefineContextInfo(topGoal)

Step\_DecomposeGoal(topGoal)

Step\_DecomposeGoal(aGoal::Goal)

ifNeedDecompose(aGoal)

Delare({aStrategy}, Strategy)

Step\_DefineContextInfo(aStrategy)

Define({aStrategy}, Supportedby, aGoal)

Declare({subGoal<sub>1</sub>, subGoal<sub>2</sub>, ..., subGoal<sub>n</sub>}, Goal)

Define({subGoal<sub>1</sub>, subGoal<sub>2</sub>, ..., subGoal<sub>n</sub>},

Supportedby, aGoal)

for each subGoal<sub>i</sub>,  $i \in \{1, \dots, n\}$

Step\_DecomposeGoal(subGoal<sub>i</sub>)

Declare({awayGolname<sub>1</sub>, awayGolname<sub>2</sub>, ..., awayGolname<sub>n</sub>}, AwayGoal)

Define({awayGolname<sub>1</sub>, awayGolname<sub>2</sub>, ..., awayGolname<sub>n</sub>}, SolutionBy, aGoal)

RefreshModView({awayGolname<sub>1</sub>, awayGolname<sub>2</sub>, ..., awayGolname<sub>n</sub>}, SolutionBy)

Declare({modname<sub>1</sub>, modname<sub>2</sub>, ..., modname<sub>n</sub>}, Module)

```

    Define({modname1, modname2, ..., modnamen},
    SolutionBy, aGoal)
    RefreshModView({modname1,
    modname2, ..., modnamen}, SolutionBy)
    Declare({contrname1, contrname2, ..., contrnamen},
    Contract)
    Define({contrname1, contrname2, ...,
    contrnamen}, SolutionBy, aGoal)
    RefreshModView({contrname1, contrname2, ...,
    contrnamen}, SolutionBy)

else
    Step_GiveSolution(aGoal)
EXIT

Step_GiveSolution(aGoal::Goal)
    Declare({solname1, solname2, ..., solnamen},
    Solution)
    Define({solname1, solname2, ..., solnamen},
    SolutionBy, aGoal)

Step_DefineContextInfo(aElement::TrunkElement)
    ifContextNeeded(aElement)
    Declare({contname1, contname2, ..., contnamen},
    Context)
    Define({contname1, contname2, ..., contnamen},
    ContextBy, aElement)
    Declare({awaycontname1, awaycontname2, ...,
    awaycontnamen}, AwayContext)
    Define({awaycontname1, awaycontname2, ...,
    awaycontnamen}, ContextBy, aElement)
    RefreshModView({awaycontname1,
    awaycontname2, ..., awaycontnamen}, ContextBy)
    Declare({awayGolname1, awayGolname1, ...,
    awayGolname1}, AwayGoal)
    Define({awayGolname1, awayGolname1, ...,
    awayGolname1}, ContextBy, aElement)
    RefreshModView({awayGolname1, awayGolname1,
    ..., awayGolname1}, ContextBy)
    ifAssumptionNeeded(aElement)
    Declare({assumname1, assumname2, ...,
    assumnamen}, Assumption)
    Define({assumname1, assumname2, ..., assumnamen},
    ContextBy, aElement)
    ifJustificationNeeded(aElement)
    Declare({justname1, justname2, ..., justnamen},
    Justification)

```

```

    Define({justname1, justname2, ..., justnamen},
    ContextBy, aElement)
    Declare({awayGolname1, awayGolname2, ...,
    awayGolnamen}, AwayGoal)
    Define({awayGolname1, awayGolname2, ...,
    awayGolnamen}, ContextBy, aElement)
    RefreshModView({awayGolname1, awayGolname2,
    ..., awayGolnamen}, ContextBy)

```

### 3 Application

This chapter takes brake system software on airplane for example to illustrate to construction of modular GSN argument. This software is the core control software of the landing gear brake control system on the airplane, which collects information like wheel speed sensor signal and braking instruction signal, and realizes the function of braking, skid resistance, and ground protection. It is the key software to ensure the safety of taking-off and landing of the plane, so it must have a high-level of safety and reliability.

#### (1) Define the top-level claim

The theme of this example is to ensure the dependability of braking software, we will consider this matter from two angles: safety and dependability. Due to space limitations, this section only demonstrates the argument of the safety. Therefore, we set the top-level claim as “the braking software on the airplane is safe”.

#### (2) Preliminary establish the structure of argument

According to the top-level claim established above, we can preliminary establish the structure of argument. The nature of software safety is “the ability of not causing an accident of the software”, concerns about the safety of the software are derived from system accidents, system accidents are caused by system hazards. Therefore, to analyse software safety, we must look from the system level. We must consider the role of the software as a component in accident of the system, and the contributions they make to system hazards. Analysis of these contributing factors, proposal of safety requirements in related software, eliminating or retarding the danger caused by software in the system, these are the keys to ensure software safety. Therefore, in this example, we preliminarily established the three-tier argument structure of “top-level claim-system hazard-safety function”. Argument structure is shown in figure 6.

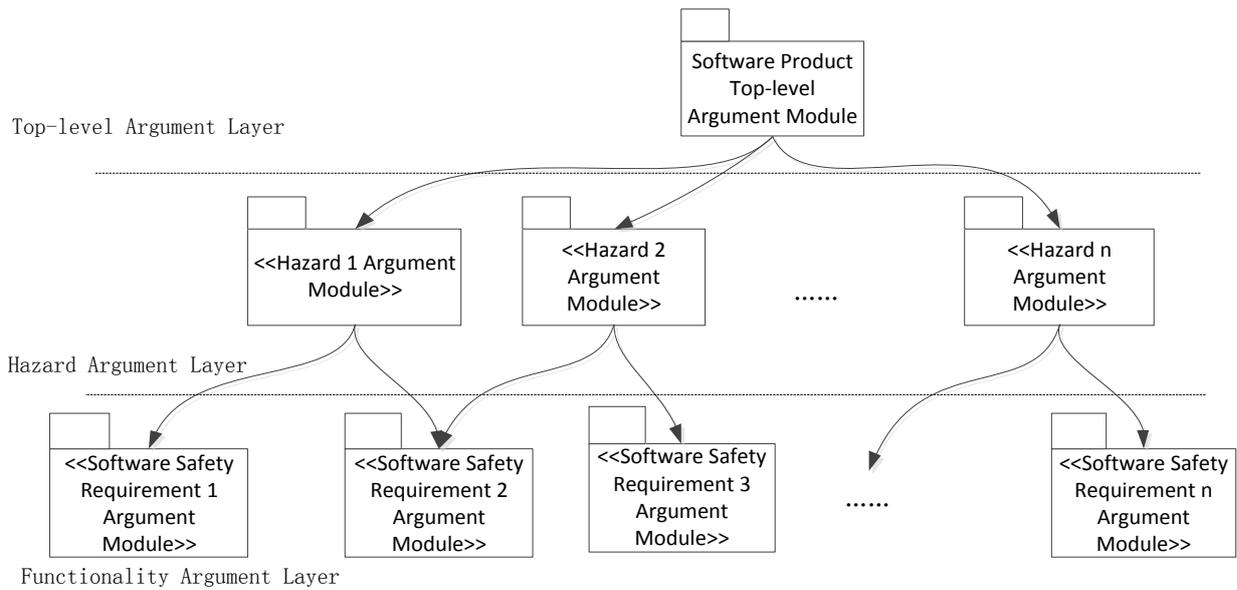


FIGURE 6 Safety argument structure of braking system software

(3) Build concrete argument and elaborate argument view

After the initial argument structure, we need to further refine each module. According to the build process presented in the above section, we can systematic develop internal argument of each module. Take the top-level argument module in software products for example:

First, declare top-level goal `Declare({topGoal}, Goal)`

We have declared a goal instance `topGoal=Goal("The final implementation of ABS software meets the software safety demands")`, in which `topGoal` is the ultimate goal of important argument in this module.

Then, define background information of the goal `Step_DefineContextInfo(topGoal)`. Realization of software product defines and explains through requirements, design documentation, and source code. Therefore, we define the background information of `topGoal` as `Declare({contname1, contname2, contname3}, Context)`, in which `contname1=Requirements of ABS software`, `contname2=Design Specification of ABS software`, `contname3=Source Code of ABS software`.

After finishing the goal information definition, we can determine whether the goal can be further decomposed. According to the macro module structure, we further decompose top-level goal adopting the way of risk-orienting. Therefore, we will enter the iterative process of goal decomposition `Step_DecomposeGoal(topGoal)`. Define argument strategy as `Delare({aStrategy}, Strategy)`, `aStrategy={"Argument over hazard introduced`

by ABS software"}.

Then, define relevant background information for the strategy. We defined two context elements, `Declare({contname1, contname2}, Context)`, `contname1={"Hazard list of ABS software"}`, `contname2={"Hazard Identification method of ABS software"}`. We also declared an Assumption element `Declare({assumname1}, Assumption)`, Assumption information is given in the argument strategy `assumption1={"Hazards are independent and can be argueded respectively"}`.

Based on the information in argument strategy background, we listed three system hazards to be respectively argued. According to the macro argument structure, each hazard should be argued in a separate module, and we must use `awayGoal` instances to show interface between top-level module and hazard argument module. We will declare three, like `Declare({awayGolname1, awayGolname2, awayGolname3}, AwayGoal)`, `awayGolname1={"Hazard 'Airplane can't decelerate by braking function' is managed adequately"}`, `awayGolname2={"Hazard "Tire blowout" is managed adequately"}`, `awayGolname3={"Hazard 'sideslipping and off tracking' is managed adequately"}`. After finishing the declare of `awayGoal` element, we didn't declare other modular modules. By the modular building approach above, decomposition process of the goal is over.

Graphical results of modular argument are shown in figure 7.

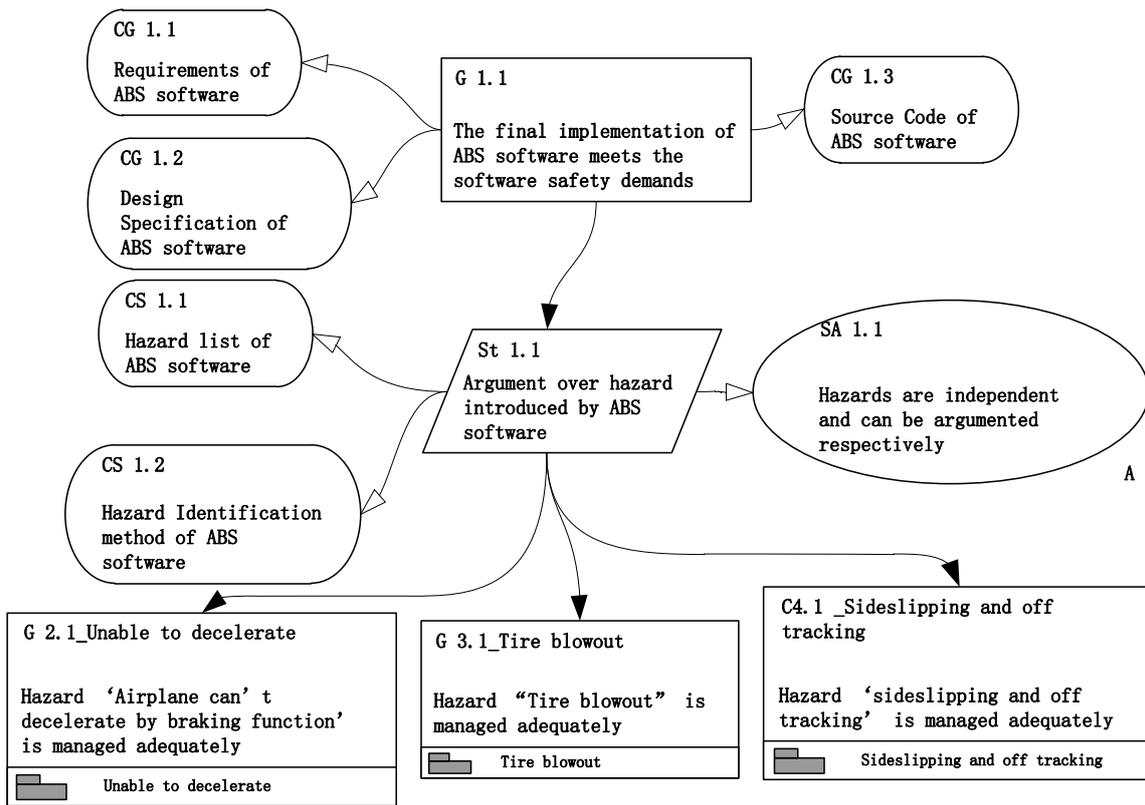


FIGURE 7 Results of modular argument

According to the top-level argument module, the updated module view is shown in figure 8:

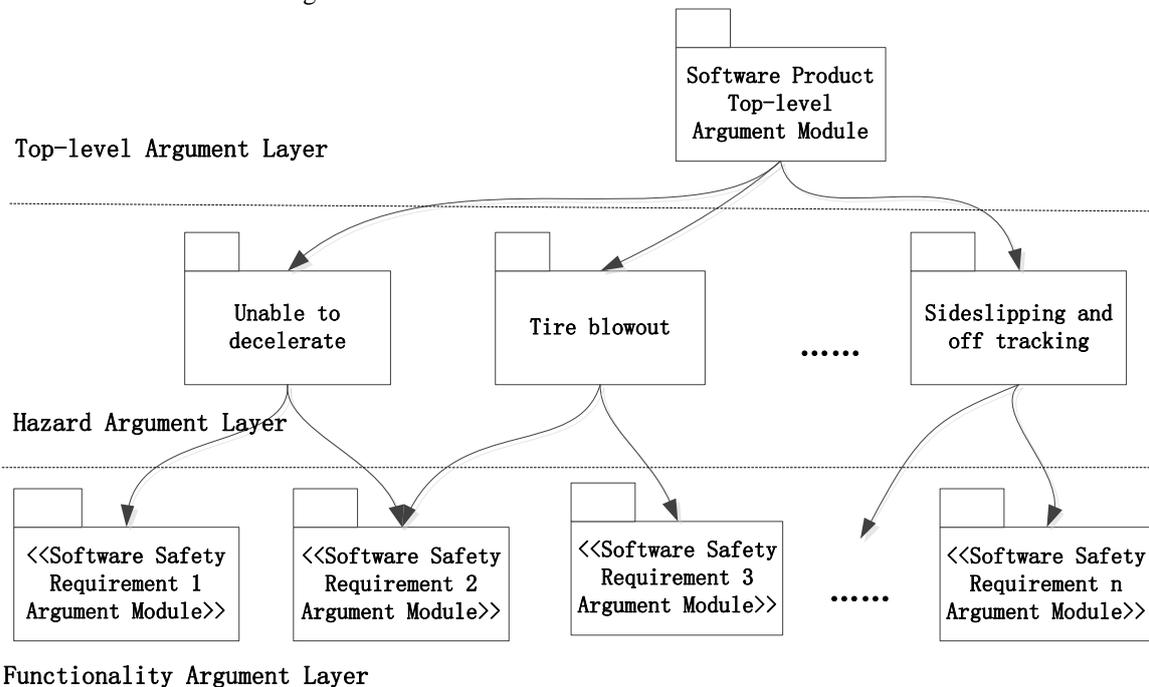


FIGURE 8 view module

#### 4 Conclusion

This paper has presented a systematic construction approach for modular software assurance case. On the basis of in-depth analysis of the argument modelling

technique, GSN, we extract the basic concepts and terminology of modularization, and summarize the relationships between concepts and constraints that must be met. A new modular GSN metamodel is proposed, which is a comprehensive, reusable description for the

internal logic of modular GSN, and can be seen as the basis of a normative argument construction process for modular assurance case. According to the metamodel, we give the progression algorithm for constructing the core structure of assurance case and present the implementation process of modular GSN argument construction in a “macro-micro-macro” iterative way. Our approach can help the assurance case developers gradually extract and analyse the argument elements, and provides a modelling process guidance for assurance case developer. We apply our approach in an ABS software system, which is a typical control software, to examine

the feasibility and effectiveness. Results show that this approach can greatly enhance the development efficiency of dependability assurance case, and can improve systematicness, comprehensiveness and scientificity of the assurance case itself, thus providing strong support for ensuring the software product can reach its desired dependability level.

At present, our approach does not include the concept of GSN pattern and its related elements, and there is also lack of tool supporting. These will be important directions of our future work.

## References

- [1] Jackson D, Thomas M, Millett L I 2007 *Software for Dependable Systems: Sufficient Evidence?* National Research Council
- [2] DO-178B *Software Considerations in Airborne Systems and Equipment Certification*
- [3] IEC 61508 *Functional Safety of Electrical/Electronic/Programmable Electronic Safety-Related Systems*
- [4] ISO 15408 *Information technology — Security techniques — Evaluation criteria for IT security — Part 1: Introduction and general model*
- [5] ISO 15408 *Information technology — Security techniques — Evaluation criteria for IT security — Part 2: Security functional requirements*
- [6] ISO 15408 *Information technology — Security techniques — Evaluation criteria for IT security — Part 3: Security assurance requirements*
- [7] Bloomfield R, Bishop P 2010 Safety and Assurance Cases: Past, Present and Possible Future – an Adelard Perspective *Proceedings of the Eighteenth Safety-Critical Systems Symposium*
- [8] Habli I, Hawkins R, Kelly T 2010 Software safety: relating software assurance and software integrity *International Journal of Critical Computer-Based Systems* 1(4) 364-83
- [9] Bishop P, Bloomfield R, Guerra S 2004 The future of goal-based assurance cases *Proceedings of Workshop on Assurance Cases of 2004 International Conference on Dependable Systems and Networks*
- [10] IEEE P15026-2 *IEEE Standard for Systems and software engineering - Systems and software assurance - Part 2: Assurance case*
- [11] Kelly T P 1998 *Arguing Safety – A Systematic Approach to Managing Safety Cases* Department of Computer Science University of York
- [12] Miller A, Gupta R 2008 *Assurance Cases for Reliability: Reducing Risks to Strengthen ROI for SCADA Systems Recent Advances in Reliability and Quality in Design* Ed H Pham Springer: London pp 465-89
- [13] He Y, Johnson C W 2012 Generic security cases for information system security in healthcare systems *7th IET International Conference on System Safety, incorporating the Cyber Security Conference*
- [14] Nakazawa J, Matsuno Y, Tokuda H 2011 Evaluating degree of systems' dependability with semi-structured assurance case *Proceedings of the 13th European Workshop on Dependable Computing(Pisa, Italy)ACM* pp 111-2
- [15] Ankrum T S, Kromholz A H 2005 Structured assurance cases: three common standards *Proceedings of Ninth IEEE International Symposium on High-Assurance Systems Engineering (HASE 2005) IEEE*
- [16] Bishop P, Bloomfield R 1998 A methodology for safety case development *Safety-Critical Systems Symposium (SAFECOMP)*
- [17] Cyra L, Górski J 2011 Support for argument structures review and assessment *Reliability Engineering & System Safety* 96(1) 26-37
- [18] Kelly T, Weaver R 2004 The Goal Structuring Notation – A Safety Argument Notation *Proceedings of Dependable Systems and Networks*
- [19] Langari Z, Maibaum T 2013 Safety cases: A review of challenges *Proceedings of 1st International Workshop on Assurance Cases for Software-Intensive Systems (ASSURE)*

## Authors

	<p><b>Dajian Zhang, born in 1982, China</b></p> <p><b>Current position, grades:</b> Ph.D. candidate  <b>University studies:</b> Beihang University  <b>Scientific interest:</b> software safety, software reliability  <b>Publications:</b> 2  <b>Experience:</b> He received his Bachelor degree in School of Information and Computing Science from Beijing Jiaotong University and now he is a Ph.D. candidate in School of Reliability and System Engineering of Beihang university.</p>
	<p><b>Minyan Lu, born in 1963, China</b></p> <p><b>Current position, grades:</b> Professor  <b>University studies:</b> Beihang University  <b>Scientific interest:</b> software dependability engineering, reliability engineering  <b>Experience:</b> She received her doctor degree in School of Reliability and System Engineering from Beihang University and now she is a professor in School of Reliability and System Engineering of Beihang university.</p>
	<p><b>Nan Wu, born in 1990, China</b></p> <p><b>Current position, grades:</b> postgraduate  <b>University studies:</b> Beihang University  <b>Scientific interest:</b> software reliability engineering  <b>Publications:</b> 1  <b>Experience:</b> He is a postgraduate in School of Reliability and System Engineering of Beihang university.</p>

# A noise estimation approach by assembling fast edge detection and block based methods

**Xiao-Yuan Qu\***

*School of Information Engineering, Yulin University, 719000, Yulin, China*

*Received 3 January 2014, www.tsi.lv*

---

## Abstract

Noise estimation is one of the most important research topics in image processing. Aishy and Eric had proposed a variance estimation method used in Gaussian white-noise, in which, a measure was provided to determine the homogeneous blocks and an analyser was used in calculating the homogeneities. The approach should present two shortcomings corresponding to structures and textures. One is that the blocks with edge textures should be considered as intensity-homogeneous blocks that could have an effect on estimation accuracy. The other is that some special blocks with high variance but low homogeneity could result in over estimation. In order to avoid the two shortcomings, in this paper we have proposed an improved noise estimation approach by combining fast edge detection and block based methods. The blocks hold continuous points were firstly excluded rejected by using fast edge detection method. The experimental results indicated that our method can avoid over estimation effectively in special conditions and can obtain more accurate results than the Aishy and Eric's method did.

*Keywords:* Guess Noise, noise variance estimation, edge detection, image processing

---

## 1 Introduction

Noise is described as the random variation of brightness or colour information in image. Denoise plays the most important role in image processing due to the performance should easily be corrupted by noise. In many denoising methods [1], the noise was used as a parameter but it was unknown in fact, so noise estimation approaches could have great effects on the qualities of denoising methods. Noise estimation has two mainly approaches, Filter-based [2] and Block-based methods [3]. The main idea of Filter-based methods is to calculate the noise variation from the high frequencies information, which are separated from the noise images and include textures and structures, so Filter-based methods can perform well for high noise levels but they are tending to overestimate in the images with large textures or details. Masoud [4] have proposed a Filter-based method that was suit for images with abundant textures, however, the method has a huge time complexity due to many times of flitting processing. The idea of the Block-based method is to calculate the noise variation of the intensity-homogeneities determined from the blocks obtained by dividing a noise image into many blocks. Block-based methods were tending to overestimate when noise levels are low [5].

Aishy and Eric [6] have proposed a novel method which fulfil well for both high noise level and low noise level, and also perform well for images with abundant textures. This method firstly determines the intensity-homogeneous blocks in the noisy image using an effective structure analyser and then estimates the noise

from these blocks. The approach should present two shortcomings. One is that the blocks with textures should be considered as intensity-homogeneous blocks that could have an effect on estimation accuracy. The other is that some special blocks have high variance but low homogeneity could result in over estimation. The two problems were due to that the non-homogeneous blocks recognized as homogeneous ones, and the problems could be solved by excluding the blocks with strong textures by some edge detection approaches.

In order to avoid the two shortcomings mentioned above, we have proposed a modified method to improve the performances of the method in [7] by combining fast edge detection and block based methods. The blocks with continuous points were firstly excluded by using fast edge detection method and then the noise variance was estimated from these intensity-homogeneous blocks. The experimental results indicated that our method can avoid over estimation effectively in special conditions and can obtain more accurate results than the Aishy and Eric's method did.

This work is partially based on the Assembling Fast Edge detection and Block Based Methods as an example and Supported by Natural Science Basic Research Plan in Shaanxi Province of China (2013JM8005).

## 2 Detection method

### 2.1 AISHY AND ERIC METHOD

Given an adding Gaussian image with unknown noise variance  $\sigma_n^2$ , there exists the following model:

---

\* *Corresponding author* e-mail: 282307708@qq.com

$$I(x, y) = S(x, y) + \eta(x, y), \tag{1}$$

where the  $I(x,y)$  is the noise image,  $S(x,y)$  is the original image and  $\eta(x,y)$  is adding Gaussian noise. The purpose of the noise estimation is to obtain the evaluative parameter for noise variance  $\sigma_n^2$ . A intensity-homogeneous block centred by pixel  $I(i,j)$  and with the size of  $S \times S$ , was presented in Equation (2).

$$B_{kl} = \{I(i, j) \mid (i, j) \in S_{kl}\} \\ S_{kl} = \{(i, j) \mid |i-k| < (S-1)/2, |j-l| < (S-1)/2\} \tag{2}$$

The variance and the mean of sub-block  $B_{kl}$  are defined as  $\sigma_{B_{kl}}^2$  and  $\mu_{B_{kl}}$ , respectively. The variance of an intensity-homogeneous block mainly presented as noise variance and satisfied with Equation (3).

$$\lim_{S \rightarrow \infty} \sigma_{kl}^2 = \sigma_n^2. \tag{3}$$

From the Equation (3), it can be found that the variance  $\sigma_n^2$  could be estimated accurately if there have enough intensity-homogeneous blocks. The homogeneity  $\zeta_{kl}$  of sub-block  $B_{kl}$  was obtained from a structure analyser. The analyser employed high throughput arithmetic operators with coefficients  $\{-1, \dots, W-1, \dots, -1\}$  to calculate the homogeneity from eight directions showed in Figure 1( with size of  $3 \times 3$  pixels), and the  $\zeta_{kl}$  was the sum of the absolute values of the eight templates and the convolutions of  $B_{kl}$ .

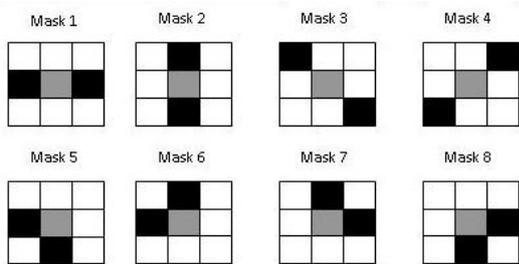


FIGURE 1 Directions of homogeneity analyser for  $3 \times 3$  blocks

The block based method is difficult to confirm the all intensity-homogeneous blocks. A self-adaptive method had proposed in [8] to select intensity-homogeneous blocks. Three blocks with minimal homogeneities were found firstly and then the referential variance  $\sigma_r^2$  was estimated as the mean of the variances calculated on the three sub-blocks mentioned above. Each block  $B_{kl}$  was included in the  $V$ , which was the aggregate of all intensity-homogeneous blocks, and must be satisfied with Equation (4), in where, the  $\sigma_i^2$  was the variance of the  $i$ -th block, PSNR was Peak Signal to Noise Ratio,  $t_{PSNR}$  was threshold.

$$\left| PSNR_{B_r} - PSNR_{B_{kl}} \right| < t_{PSNR} \\ PSNR_i = 10 \log_{10} \frac{255^2}{\sigma_i^2} \tag{4}$$

After the aggregate  $V$  was determined, the variance was estimated as following:

$$\sigma_e^2 = \frac{\sum_{(k,l) \in V} \sigma_{B_{kl}}^2}{m}, \tag{5}$$

where  $m$  was the number of blocks selected into  $V$ .

From above, we can get two conclusions:

- 1) When the sub-blocks  $B_{kl}$  with size greater than  $3 \times 3$ , there could be exist some special sub-blocks with very small homogeneities  $\zeta_{kl}$  and very high variance  $\sigma_{kl}^2$  as showed in Figure 2. If there was one special sub-block in the image, the  $\sigma_{kl}^2$  should be determined as the referential variance  $\sigma_r^2$ , so it should result in serious over estimation due to the  $\sigma_r^2$  was much higher than noise variance  $\sigma_n^2$ .
- 2) To determine that whether the sub-block  $B_{kl}$  was an intensity-homogeneous block or not, is corresponding to  $\sigma_{kl}^2$ , so the sub-blocks determined as intensity-homogeneous one should have structures and textures that should take bed effects on the calculating accuracy.

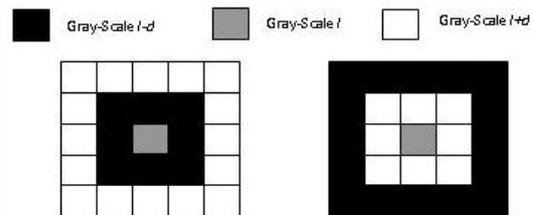


FIGURE 2 Two special sub-blocks with  $I-d \geq 0, I+d \leq 255$

### 2.2 OUR METHOD

The two problems mentioned above could be voided effectively by excluding the sub-blocks with continuous points, in this paper, a simple edge detection method was employed to accomplish the task. Firstly, the Sobel arithmetic operators [1] were used in calculating the gradient of image as showed in Equation (6).

$$G = \sqrt{G_x^2 + G_y^2} \\ G_x = I(x, y) * \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, G_y = I(x, y) * \begin{bmatrix} -1 & 0 & -1 \\ -2 & 0 & -2 \\ -1 & 0 & -1 \end{bmatrix} \tag{6}$$

A gradient threshold  $G_{th}$  was selected and used in detecting the boundary according to the following rules:

1. Given a settled threshold  $p$ , value of gradient  $G_{th}$  was determined by if the gradients which were higher than  $G_{th}$  took  $p\%$  of the all gradients.
2. A point had a gradient  $G$ , which was higher than  $G_{th}$  was confirmed as a boundary point.

So if a sub-block  $B_{kl}$  had  $N$  continuous points could be excluded as an intensity-homogeneous block. In our method, the approach mentioned in this section was used firstly as a filter to get rid of the sub-blocks, than the

method [1] was employed to estimate the noise variance of the remainders of sub-blocks.

### 3 Results and Discussions

In order to make comparisons between the two methods, a ratio [1] was employed and showed in equation (7):

$$\text{Estimation ratio} = \frac{\sigma_{\text{estimation}}}{\sigma_{\text{added}}}, \tag{7}$$

where  $\sigma_{\text{estimation}}$  was standard deviation of estimated noise,  $\sigma_{\text{added}}$  was the standard deviation of added noise. We had made tests on several images with the size of 512×512 and under different level of noise as showed in Figure 3. The image was decomposed into a series of sub-blocks and each of the blocks was not overlapped and with the size of 5×5. The parameters, thresholds  $t_{PSNR}$  and  $p$  were initialized with values of 3dB and 10, respectively. The parameter N was initialized to three; it means that if the number of continuous points detected in a sub-block was greater than  $N$ , the block was rejected to be an intensity-homogeneous block.

Tables 1 and 2 showed the estimation ratios (defined in Equation (6)) obtained from the eight images showed in Figure 3 between our method and Aishy method [1]. The results listed in the two tables were the mean values of 20 times noises variance estimations under the same experimental condition. From the results, we could find that our methods have better performances at many conditions. At the range of lower noise like  $\sigma_{\text{added}} \leq 3$ , our method obtained the same performances as method Aishy method [1]. Under this condition, the sub-blocks were hardly determined to be homogeneous because of variance of the block which had continuous points was much higher than the noise variance. At the range of middle noise like  $10 < \sigma_{\text{added}} < 30$ , our method showed superiority because the self-adaptive method [1] had high probability to select the sub-blocks which hold continuous points.

At high noise range of  $\sigma_{\text{added}} \geq 30$ , the two method Showed great performances when  $\sigma_{\text{added}}$  was at 30, 40 and 50. But our method had superiority on the images with abundant edge textures (see results listed in Table 2 of the images Mandrill and Hyderabad).

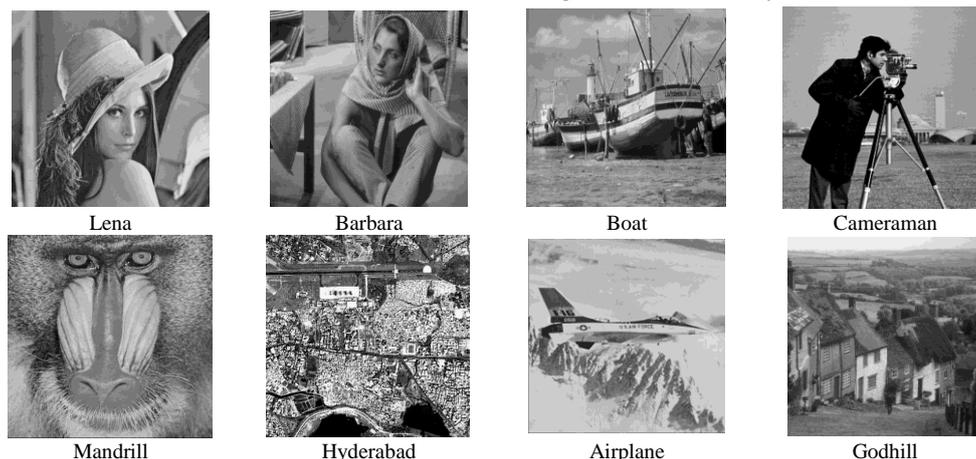


FIGURE 3 Eight images with size of 512X512

TABLE 1 The results obtained by our method and Aishy [1] on the images of Lean, Barbara, Boat and Cameraman

$\sigma_{\text{added}}$	Lena		Barbara		Boat		Cameraman	
	Method [1]	Ours'	Method [1]					
0.5	3.7740	3.7740	4.3330	4.3330	3.1730	3.1730	1.9430	1.9430
1	2.3235	2.3235	2.3785	2.3785	2.1215	2.1215	1.3940	1.3940
2	1.5660	1.5660	1.4890	1.4890	1.6495	1.6490	1.2063	1.2063
3	1.2860	1.2858	1.2892	1.2892	1.3873	1.3872	1.1003	1.1003
4	1.1596	1.1595	1.1951	1.1951	1.2993	1.2990	1.0609	1.0579
5	1.1429	1.1420	1.1281	1.1281	1.2389	1.2386	1.0312	1.0306
6	1.1139	1.1106	1.0847	1.0847	1.2039	1.2030	1.0407	1.0388
7	1.1031	1.1018	1.0656	1.0655	1.1416	1.1409	1.0034	1.0027
8	1.0489	1.0479	1.0761	1.0758	1.1176	1.1166	0.9989	0.9992
9	1.0646	1.0624	1.0792	1.0785	1.1056	1.1022	0.9969	0.9998
10	1.0294	1.0281	1.0361	1.0357	1.0937	1.0885	1.0128	1.0071
12	1.0539	1.0489	1.0251	1.0244	1.0895	1.0869	1.0023	0.9999
14	1.0188	1.0157	1.0291	1.0280	1.0493	1.0462	1.0072	1.0034
16	1.0128	1.0073	1.0147	1.0136	1.0713	1.0536	1.0075	1.0026
18	1.0164	1.0037	0.9882	0.9924	1.0363	1.0303	0.9817	0.9905
20	1.0062	1.0004	1.0191	1.0107	1.0184	1.0065	1.0101	1.0002
30	1.0005	0.9992	1.0107	0.9994	1.0199	0.9940	0.9836	0.9824
40	0.9906	0.9899	1.0009	0.9960	0.9999	0.9919	0.9817	0.9806
50	0.9875	0.9870	1.0266	0.9951	0.9870	0.9858	1.0041	0.9978

TABLE 2 Results obtained by our method and Aishy's on the images of Mandrill, Hyberabad, Airplane and Goldhill showed in Figure 3

$\sigma$ added	Mandrill		Hyberabad		Airplane		Godhill	
	Method [1]	Method						
0.5	5.5770	5.5770	0.9860	0.9860	1.1990	1.1990	1.5120	1.5120
1	3.1825	3.1825	0.9895	0.9895	1.1530	1.1530	1.3380	1.3380
2	2.5780	2.5780	0.9953	0.9953	1.1293	1.1293	1.1770	1.1770
4	1.8380	1.8380	0.9999	0.9999	1.0381	1.0379	1.1926	1.1924
5	1.8477	1.8476	0.9999	0.9999	1.0304	1.0301	1.1508	1.1401
6	1.7128	1.7125	0.9972	0.9972	1.0479	1.0471	1.1748	1.1738
7	1.6258	1.6254	1.0051	1.0051	1.0304	1.0295	1.1874	1.1814
8	1.5097	1.5093	1.0200	1.0200	1.0078	1.0042	1.1641	1.1607
9	1.5442	1.5436	1.0317	1.0317	1.0086	1.0071	1.1210	1.1128
10	1.4886	1.4878	1.0280	1.0280	1.0051	1.0002	1.1068	1.1019
12	1.3148	1.3140	1.0469	1.0469	0.9925	0.9954	1.0489	1.0437
14	1.2979	1.2962	1.0893	1.0891	0.9927	0.9952	1.0829	1.0719
16	1.2451	1.2422	1.1151	1.1103	0.9862	0.9919	1.0640	1.0476
18	1.2279	1.2228	1.0682	1.0679	1.0191	1.0072	1.0501	1.0378
20	1.1814	1.1753	1.1284	1.1080	0.9943	0.9965	1.0445	1.0223
30	1.1098	1.0967	1.1257	1.1198	1.0035	0.9969	1.0212	1.0022
40	1.0786	1.0585	1.1672	1.1590	0.9790	0.9753	1.0113	0.9970
50	1.0448	1.0238	1.1164	1.1016	0.9955	0.9885	0.9824	0.9816

#### 4 Conclusions

Our method was improved one based on the method [1], compared with which there were two superiorities of our method. One is that our method could get more accurate performances at lower and middle noise range. The other one is that our method could avoid over-estimations effectively, as we know that over-estimation could produce inestimable effects at later image processing. We

#### References

- [1] Amer A 2005 Fast and Relialbe Structure-Oriented Video Noise Estimation *IEEE Transactions On Circuits and Systems for Video Technology* **15**(1) 113-8
- [2] Yan C, Liu R 2010 A Game Theoretical Approach for Image Denoising *Proceedings of the 17th International Conference on Image Processing Hong Kong 2010* 1125-8
- [3] Donoho D L, Johnstone J M 1994 Ideal Spatial Adaptation by Wavelet Shrinkage *Biometrika* **81** 425-55

hope that our method could play an important role in the noise estimation research.

#### Acknowledgments

This work is partially supported by Natural Science Basic Research Plan in Shaanxi Province of China (2013JM8005). Thanks for the help.

- [4] Tai S C, Yang S M 2008 A Fast Method for Image Noise Estimation Using Laplacian Operator and Adaptive Edge Detection *Proceeding of the 2th International Symposium on Communications Control and Signal Processing Malta 2008* 1077-81
- [5] Zhu L, Xu P X 2006 Differential Video Noise Estimation *Proceeding of the 8th International Conference on Communication Technology Guangxi China 2006* **15**(1) 113-8
- [6] Hashemi M, Beheshti S 2010 Adaptive Noise Variance Estimation in *BayesShrink IEEE Signal Processing Letters* **17**(1) 12-5
- [7] Gonzalez R C, Woods R E 2002 Digital Image Processing Prentice-Hall Inc New Jersey USA

#### Authors



**Qu Xiao-yuan**, born on December 19, 1981, in Shannxi Yulin

**Current position, grades:** lecturer in Yulin University.

**University studies:** MS degree in Computer science from Xidian University in 2010.

**Scientific interest:** Data mining technology, Mass data processing technology.

# Supervised orthogonal tensor neighborhood preserving embedding for face recognition

**Jianjun Chen\***

*Yuanpei College of Shaoxing University, China*

*Received 1 March 2014, www.tsi.lv*

---

## Abstract

The deficiency of supervised discriminant information is the problem of Orthogonal Tensor Neighborhood Preserving Embedding (OTNPE) proposed recently for face recognition. So a dimension reduction algorithm called Supervised Orthogonal Tensor Neighborhood Preserving Embedding (SOTNPE) is proposed in the paper. On the basic of OTNPE, the algorithm achieves neighborhood reconstructions within the same class, preserving supervised class label information and neighborhood reconstruction information. Experiments on AR and YaleB face datasets show our proposed algorithm is efficient.

*Keywords:* face recognition, dimensionality reduction, orthogonalization, tensor neighborhood preserving embedding, neighborhood reconstruction

---

## 1 Introduction

Manifold learning is an effective way of machine learning in recent years, which discloses geometry structure features hidden in data and has been successfully applied to data mining. Typical manifold dimensionality reduction algorithms include Locally Linear Embedding (LLE) [1], Isometric Feature Mapping (ISOMAP) [2], Laplacian Eigenmaps (LE) [3], Locality Preserving Projection (LPP) [4] and Neighborhood Preserving Embedding (NPE) [5].

NPE is the approximation of local linear embedding, preserving local geometry structure and neighborhood relations. Due to power discriminant performance, NPE has attracted the attention of researchers and has been widely used in face recognition. Nowadays researches on NPE are divided into following three categories according to the processing way of data.

1) Vector based NPE [6-10]. These algorithms need to transform face image matrixes into vectors, which increase complexity of computing matrix to vector conversion.

2) Two-dimensional matrix based NPE [11-13]. However, the algorithms only are limited in the row or column, ignoring the spatial relationship of the image pixels.

3) Second-order and more order tensor based NPE [11-13]. These algorithms represent face image with second-order data, which not only preserve local information of the image pixel but also preserve the spatial structure of the image pixel [14-16]. On the basic of NPE, researchers proposed Tensor Neighborhood Preserving Embedding (TNPE) [15]. Liu et al [17] proposed Orthogonal Tensor Neighborhood Preserving Embedding (OTNPE). By orthogonalizing projections

matrixes, OTNPE has more ability for preserving local geometry structure and neighbor relations and has been successfully applied to facial expression recognition. However, if the sample image data is not smooth and compact in manifold embedded, the discriminant performance of OTNPE is not satisfactory.

Usually supervised information based on class label strengthens the between-class separability of samples, containing discriminant information. Inspired by OTNPE problem, a dimensionality reduction algorithm named Supervised Orthogonal Tensor Neighborhood Preserving Embedding (SOTNPE) for face recognition is proposed in the paper. The algorithm firstly regards multi-dimensional face image as a multi-order tensor data, and then achieves approximately linear reconstruction of samples within the same class and gets projections. Projected data not only preserves the geometric structure and local neighborhood information but also preserves the between-class separability of samples. Experiments on AR and YaleB show that our algorithm is efficient.

The organization of this paper is as follows. Related works is presented in Section 2. We discuss SOTNPE in Section 3. In Section 4, we present experiments for demonstrating the effectiveness of SOTNPE. Conclusions are drawn in Section 5.

## 2 Related Works

### 2.1 NEIGHBORHOOD PRESERVING EMBEDDING (NPE)

Given samples  $X = [x_1, \dots, x_n] \in R^{d \times n}$ , NPE attempts to search the projection matrix  $T$  to get  $Y = T^T X$ . There are some following steps for NPE:

---

\* *Corresponding author* e-mail: chengjianjun\_2012@163.com

1) Construct the adjacent graph G. Point sets of G consist of samples and common ways of selecting neighborhoods are  $k$ - nearest neighbors and  $\varepsilon$ -neighbor-hood.

2) Calculate reconstructive weights W. According to the adjacent graph G, each point can be reconstructed through the linear way with its k- neighborhoods. For  $x_i$  in X, the cost of reconstruction of  $x_i$  is described with following function [1]:

$$\begin{aligned} & \min_T \sum_i \left\| x_i - \sum_{j=1}^k w_{ij} x_j \right\|^2 \\ & s.t. \sum_{j=1}^k w_{ij} = 1 \\ & x_j \in O(x_i, k) \end{aligned} \quad (1)$$

where,  $O(x_i, k)$  denotes the k- neighborhoods set of  $x_i$ .

3) Projected data  $Y = T^T X$  satisfies Equation (1), we get:

$$\begin{aligned} & \min_T \sum_i \left\| y_i - \sum_{j=1}^k w_{ij} y_j \right\|^2 = \\ & \min_T \|Y(I-W)\|^2 = \\ & \min_T (Y(I-W)(I-W)^T Y^T) = \\ & \min_T (T^T X(I-W)(I-W)^T XT) = \\ & \min_T (T^T XMX^T) \end{aligned} \quad (2)$$

where  $M = (I-W)(I-W)^T$ , constrain conditions are introduced as follows:

$$\begin{aligned} & \sum_{i=1}^N y_i = 0 \\ & \frac{1}{N-1} \sum_{i=1}^N y_i^T y_i = I \end{aligned} \quad (3)$$

We replace  $Y = T^T X$  in Equation (2) and Equation (3). The optimization objective function of NPE is listed as follows:

$$\begin{aligned} & \min_T T^T XMX^T \\ & T^T XX^T T = I \end{aligned} \quad (4)$$

## 2.2 TENSOR NEIGHBORHOOD PRESERVING EMBEDDING (TNPE)

TNPE is the tensors extend of NPE. Given  $X = [x_1, \dots, x_n]$  in tensor space  $R^{I_1 \times I_2 \times \dots \times I_k}$ . The destination of TNPE is to search l projection matrixes  $U^i \in R^{m_i \times m_i}$  ( $m_i > m_i, i = 1 \dots l$ ) to preserve local

neighborhood reconstruction. The objective function of TPNE is described as follows:

$$\begin{aligned} & \min \sum_i \left\| x_i \times_1 U^1 \dots \times_k U^k - \sum_j M_{i,j} x_j \times_1 U^1 \dots \times_k U^k \right\|_F^2 \\ & s.t. \sum_i \left\| x_i \times_1 U^1 \dots \times_k U^k \right\|_F^2 = 1 \end{aligned} \quad (5)$$

## 2.3 ORTHOGONAL TENSOR NEIGHBORHOOD PRESERVING EMBEDDING (OTNPE)

On the basic of TNPE, orthogonal conditions of projections matrixes are added in TNPE. The objective function of OTNPE is described as follows:

$$\begin{aligned} & \min \sum_i \left\| x_i \times_1 U^1 \times_2 U^2 \dots \times_k U^k - \sum_j M_{i,j} x_j \times_1 U^1 \times_2 U^2 \dots \times_k U^k \right\|_F^2 \\ & s.t. \sum_i \left\| x_i \times_1 U^1 \times_2 U^2 \dots \times_k U^k \right\|_F^2 = 1 \\ & (U^i)^T U^i = I (i = 1, \dots, l) \end{aligned} \quad (6)$$

## 3 Supervised Orthogonal Tensor Neighborhood Preserving Embedding (SOTNPE)

### 3.1 THE OBJECTIVE FUNCTION

An i-dimensional image itself may be represented as a matrix or a second-order tensor. With tensor algebra used for the analysis of images, an i-dimensional image is regarded as a point of i-order tensor space. On the basic of Equation (6), we have introduced supervision discrimination information based on class label,

Given samples  $X = [x_1, \dots, x_n] = [\chi_1, \dots, \chi_k]$ , where  $\chi_i$  denotes samples of the i-th class. The objective function of SOTNPE is described as follows:

$$\begin{aligned} & \min \sum_i \left\| x_i \times_1 U^1 \times_2 U^2 \dots \times_k U^k - \sum_j M_{i,j} x_j \times_1 U^1 \times_2 U^2 \dots \times_k U^k \right\|_F^2 \\ & s.t. \sum_i \left\| x_i \times_1 U^1 \times_2 U^2 \dots \times_k U^k \right\|_F^2 = 1 \\ & (U^i)^T U^i = I (i = 1, 2, \dots, l) \\ & x_j \in \mathcal{X}_{label(x_i)} \end{aligned} \quad (7)$$

where  $label(x_i)$  denotes the class label of  $x_i$  and  $\mathcal{X}_{label(x_i)}$  denotes samples whose class label is as same as that of  $x_i$ .

### 3.2 ALGORITHM STEPS

**Input:** samples  $X = [x_1, \dots, x_n] \in R^{m_1 \times m_2}$ .

**Output:** projection matrixes  $U^1, U^2, \dots, U^l$ .

**Steps:**

1) On the basic of Equation (1), constrain condition are added, namely that  $y_i$  and  $y_j$  are in the same class.

$$\begin{aligned} & \min_{\mathbf{T}} \sum_i \left\| x_i - \sum_{j=1}^k w_{ij} x_j \right\|^2 \\ & s.t. \sum_{j=1}^k w_{ij} = 1 \\ & x_j \in \mathcal{X}_{label(x_i)} \end{aligned} \quad (8)$$

Calculate the reconstructive matrix  $W$  using the way in [18].

2) Set  $\sum_i U_0^i = I (i=1, \dots, l)$ .

3)

3.1) the number of iterations  $\tau=1, 2, \dots, t$

3.1.1) Calculate 1 projection matrixes in iterations  $m=1, 2, \dots, l$

3.1.1.1) Set  $\Psi_i^k = x_i \times_1 U_\tau^1 \times_2 U_\tau^2 \times \dots \times_l U_\tau^l (i=1, \dots, n)$

3.1.1.2) Set  $D_k = \sum_i \Psi_i^{(k)} (\Psi_i^{(k)})^T$  and

$$\begin{aligned} S_k &= \sum_i \sum_j M_{ij} \Psi_i^{(k)} (\Psi_j^{(k)})^T + \sum_i \sum_j M_{ij} \Psi_j^{(k)} (\Psi_i^{(k)})^T - \\ & \sum_i \left( \sum_j M_{ij} \Psi_j^{(k)} \sum_j M_{ij} (\Psi_j^{(k)})^T \right) \end{aligned}$$

3.1.1.3) If  $D_k$  is singular, then

$$D_k = D_k + \mu I_{mk} (\mu = 0.001).$$

3.1.1.4) Calculate  $u_i^k$  in  $(D_k - S_k)u_i^k = \lambda_l (D_k)^{-1} u_i^k$  using the generalized matrix solution.

3.1.1.5) loop of iteration  $p=2, \dots, m_k$

3.1.1.5.1) Set

$$U_k^{(E-1)} = [u_1^k, \dots, u_{p-1}^k], H_k^{(p-1)} = [U_k^{(p-1)}]^{-T} (D_k)^{-1} (D_k - S_k)$$

3.1.1.5.2) Set

$$\begin{aligned} \Phi_k^{(p-1)} &= \\ & (I_{mk} - (D_k)^{-1} U_k^{(p-1)} [H_k^{(p-1)}]^{-1} [U_k^{(p-1)}]^T) (D_k)^{-1} (D_k - S_k) \end{aligned}$$

3.1.1.5.3) Calculate  $u_p^k$  and Orthogonalized

$u_p^k = u_p^k / \|u_p^k\|$  in  $\Phi_k^{(p-1)} u_p^k = \lambda_l I_k u_p^k$  using the generalized matrix solution.

3.1.1.5.4) end the loop  $p$

3.1.1.6) Set  $U_\tau^k = [u_1^k, u_2^k, \dots, u_{m_k}^k]$ .

3.1.1.7) End the loop  $k$ .

3.1.2) If  $\tau > 1$  and  $\|U_\tau^k - U_{\tau-1}^k\| < \varepsilon$  ( $\varepsilon$  is error), then

end the loop  $\tau$ .

3.2) End the loop  $\tau$ .

4) Get projection matrixes  $U^i (i=1, 2, \dots, l)$  and

$$Y = X \times_1 U^1 \times_2 U^2 \dots \times_k U^l.$$

**4 Experiments**

**4.1 EXPERIMENTAL DATA**

In the experiment, AR and YaleB face datasets are used as experimental data and they are described as follows:

1) AR [18] consists of over 4000 face images of 126 individuals. For each individual, 26 pictures were taken in two sessions (separated by two weeks) and each section contains 13 images. These images include front view of faces with different expressions, illuminations and occlusions. Figure 1 shows a group of samples in AR.



FIGURE 1 A group of samples in AR

2) YaleB [19] consists of 2414 frontal-face images of 38 individuals. Face images were captured under various laboratory-controlled lighting conditions. Figure 2 shows a group of samples in YaleB.



FIGURE 2 A group of samples in YaleB

**4.2 EXPERIMENTAL SETTINGS**

TPCA [14], TLPP [15], TNPE [15] and OTNPE [17] are selected to compare with our algorithm. Parameter settings of them are listed in Table 1.

TABLE 1 Parameter settings of algorithm

Algorithms	Parameter settings
TPCA	no
TLPP	$\kappa=7$
TNPE	$\kappa=7, l=2$
OTNPE	$\kappa=7, l=2$
SOTNPE	$l=2$

Besides, we randomly select L images per class for training and the remaining for test. Nearest Neighbor algorithm for classification is adopted. All experiments were repeated 40 times and the average of recognition accuracy is gotten as experimental results.

**4.3 EXPERIMENTAL ANALYSIS**

In order to improve the computational efficiency, images of the AR are resized to  $30 \times 30$ . L is set to 6 and 10 and the recognition accuracy is gotten respectively. Figures 3 and 4 show experimental results on AR.

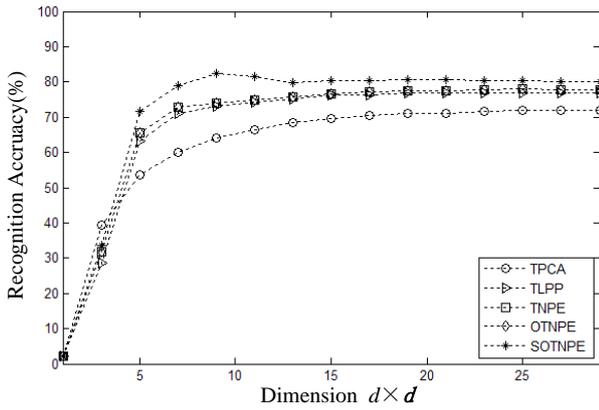


FIGURE 3 Recognition Accuracy VS. Dimension with L=6 on AR

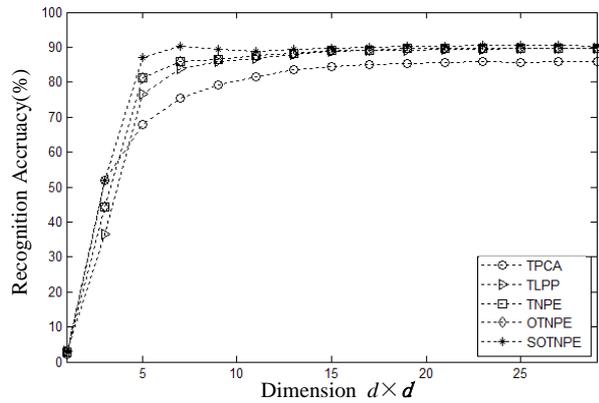


FIGURE 4 Recognition Accuracy VS. Dimension with L=12 on AR

From above Figures 3 and the following conclusions are drawn:

1) With increase in the dimension of the subspace, recognition accuracies of all algorithms improve rapidly. When the dimension exceeds a certain value, the recognition accuracy of TPCA, TLPP, TNPE and OTNPE become gradually stable while that of SOTNPE decreases slowly. This illustrates that SOTNPE gets the best performance in less dimension.

2) In contrast to other algorithms, advantages of SOTNPE decline when the number of training samples L is set to 12, which demonstrates that the overfitting problem exists in SOTNPE.

3) Similarly, images of the YaleB are resized to 30x30. Figures 5 and 6 show experimental results on YaleB.

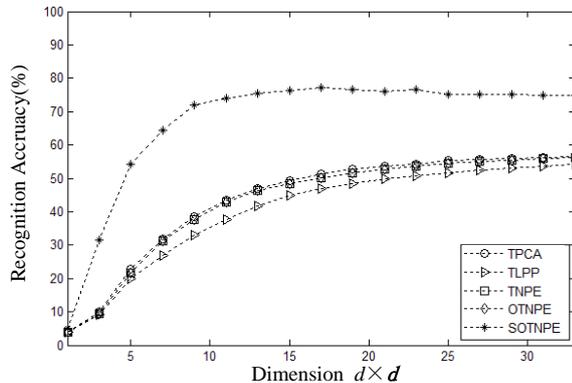


FIGURE 5 Recognition Accuracy VS. Dimension with L=10 on YaleB

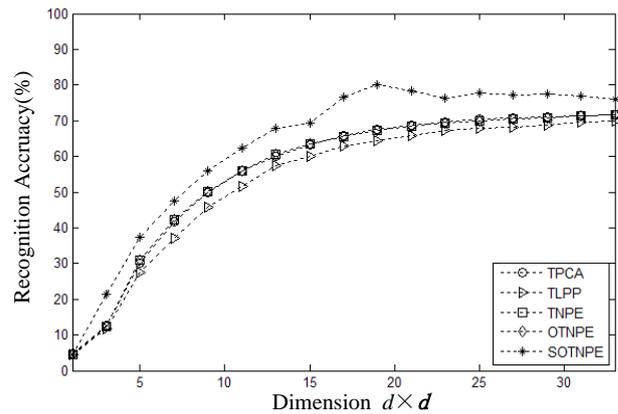


FIGURE 6 Recognition Accuracy VS. Dimension with L=20 on YaleB

We can draw following conclusions from Figure 5 and Figure 6:

1) In contrast to TPCA, TLPP, TNPE and OTNPE, SOTNPE is superior obviously to them. The reason is that SOTNPE not only captures the local nonlinear structure information but also contains discriminant information on YaleB.

2) Advantages of SOTNPE decline when L is set to 20, which also demonstrates that the overfitting problem exists in SOTNPE.

## 5 Conclusions

An algorithm called Supervised Orthogonal Tensor Neighborhood Preserving Embedding (SOTNPE) for dimensionality reduction is proposed in the paper. The algorithm achieves within-class reconstruction instead of reconstruction based on k-nearest neighborhoods of samples on the basis of OTNPE. In contrast to OTNPE, SOTNPE not only inherits the characteristics of OTNPE and fuses more supervision information based on class label, containing power discriminant information. The experiments in the AR and YaleB face database show that SOTNPE outperforms OTNPE obviously. However, like other supervised dimensionality reduction algorithms, the problem of overfitting remains in SOTNPE, how to fuse global unsupervised information is the next work.

## Acknowledgments

The work is support by NSF of Zhejiang province in China (LQ12F02007) and the reform project of the new century of Zhejiang province in China (YB2010092).

## References

- [1] Roweis ST, Saul LK 2000 Nonlinear dimensionality reduction by locally linear embedding *Science* **290**(5500) 2323-6
- [2] Tenenbaum J B de Silva V, Langford J C 2000 A global geometric framework for nonlinear dimensionality reduction *Science* **290**(5500) 2319-23
- [3] M Belkin P 2003 Niyogi Laplacian eigenmaps for dimensionality reduction and data representation *Neural Computation* **15**(6) 1373-96
- [4] He X F, Yan S C, Hu Y X, P Niyogi, H J Zhang 2005 Face recognition using Laplacianfaces *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**(3) 328-40
- [5] He X F, Cai D Y, Zhang S C 2005 Neighborhood Preserving Embedding *In Proc of the 10th IEEE International Conference on Computer Vision Washington IEEE Computer Society Press* 1208-13
- [6] Yong Wang, Yi Wu 2010 Complete neighborhood preserving embedding for face recognition *Pattern Recognition* **43**(1) 1008-15
- [7] Zhou W, Ahrary A, Kamata S 2012 Image Description with Local Patterns: An Application to Face Recognition *Transaction on Information and Systems* **95**(5) 1494-505
- [8] Chen Xi, Zhang Jiashu 2012 A novel maximum margin neighborhood preserving embedding for face recognition 2012 *Future Generation Computer Systems* **28**(5) 212-7
- [9] Kang-hua Hui, Chun-li Li, Zhang Lei 2012 Sparse neighbor representation for classification *Pattern Recognition Letters* **33**(5) 661-9
- [10] Gui-Fu Lu, Zhong Jin, Jian Zou 2012 Face recognition using discriminant sparsity neighborhood preserving embedding *Pattern Recognition* **31**(7) 119-27
- [11] Yong Wang, Yi Wu 2012 A two-dimensional Neighborhood Preserving Projection for appearance-based face recognition *Pattern Recognition* **45**(5) 1866-76
- [12] Yuan L, Mu Z C 2012 Ear recognition based on local information fusion *Pattern Recognition Letters* **33**(2) 182-90
- [13] Zhang D M, Fu M S, Luo B 2011 Image Recognition with Two-Dimensional Neighbourhood Preserving Embedding *Pattern Recognition and Artificial Intelligence* **24**(6) 810-5
- [14] He X, Cai D, Niyogi P 2005 Tensor subspace analysis *In Proceedings of the Neural Information Processing Systems*, 499-506
- [15] Dai G, Yeung D 2006 Tensor embedding methods *In Proceedings of the National Conference on Artificial Intelligence* 330-5
- [16] Wei Y T, Li H, Li L Q 2009 Tensor locality sensitive discriminant analysis and its complexity *International Journal of Wavelets Multiresolution and Information Processing* **11**(7) 865-80
- [17] S Liu Q Ruan 2011 Orthogonal Tensor Neighborhood Preserving Embedding for facial expression recognition *Pattern Recognition* **44**(7) 1497-513
- [18] Martinez A M, Kak A C 2001 PCA Versus LDA *IEEE Transactions on Pattern Analysis Machine Intelligence* **23**(2) 228-33
- [19] Lee K, Ho J, Kriegman D 2005 Acquiring linear subspaces for face recognition under variable lighting *IEEE Trans. IEEE Transactions on Pattern Analysis Machine Intelligence* **27**(5) 684-98

## Authors



**Jianjun Chen, born in May 1966, Zhejiang, China**

**Current position, grades:** associate professor of Shaoxing University, Shaoxing, China.

**University studies:** Bachelor's degree in physics science from Zhejiang University in 1987, the Master's degree from Taiyuan University of Technology in 2004.

**Scientific interest:** machine learning and image processing.

**Publications:** 11.

# Constraint-based sparsity preserving projections and its application on face recognition

**Jianjun Chen\***

*Yuanpei College of Shaoxing University, China*

*Received 1 March 2014, www.tsi.lv*

---

## Abstract

Aiming at the deficiency of supervise information in the process of sparse reconstruction in Sparsity Preserving Projections (SPP), a semi-supervised dimensionality reduction method named Constraint-based Sparsity Preserving Projections (CSPP) is proposed. CSPP attempts to make use of supervision information of must-link constraints and cannot-link constraints to adjust the sparse reconstructive matrix in the process of SPP. On one hand, CSPP obtains the high discriminative ability from supervised pairwise constraint information. On the other hand, CSPP has the strong robustness performance, which is inherited from the sparse representation of data. Experimental results on UMIST, YALE and AR face datasets show, in contrast to unsupervised SPP and existing semi-supervised dimensionality reduction method on sparse representation, our algorithm achieves increase in recognition accuracy based on the nearest neighbour classifier and promotes the performance of dimensionality reduction classification.

*Keywords:* semi-supervised dimensionality reduction, pairwise constraint, sparse representation, sparse reconstruction, sparsity preserving projections, face recognition

---

## 1 Introduction

Principal Component Analysis (PCA) [1] and Linear Discriminant Analysis (LDA) [2] fail to explore the essential structure of the data with non-linear distribution and how to select kernel and optimal kernel parameter in kernel version of them is still difficult. Representative Manifold learning algorithms [3-6] have been developed. Unfortunately, all of these algorithms are plagued by the out-of-sample problem. The solution for this problem is to apply a linearization procedure to construct explicit maps over new measurements. For example, Local Preserving Projections (LPP) [7] is a linearization version of LE; Neighbourhood Preserving Embedding (NPE) [8] is a linearization version of LLE; Isometric Projection (IsoProjection) [9] can be seen as a linearized Isomap; and Linear Local Tangent Spacen Alignment (LLTSA) [10] is a linearization of LTSA. But these algorithms fail to explore instinct geometry structure.

In recent years the study of sparse representation (SR) of signals has attracted many attentions. The purpose of the sparse representation is to optimize the most compact representation of a signal with linear combination of atoms in an over complete dictionary. SR has been successfully applied in many practical problems [11-15]. Researches [11] showed that classifier based on SR is exceptionally effective and achieves by far the best recognition rate on some face databases. Nowadays researches on dimensionality reduction based on SR have attached more and more attentions. Sparsity Preserving Projections (SPP) [16] is a representative algorithm. SPP

firstly constructs an adjacent weight matrix of the data set based on SR and then evaluate the low-dimensional embedding of the data to best preserve such weight matrix. SPP is proved to outperform PCA, LPP and NPE, and avoids the difficulty of parameter selection as in LPP and NPE. Although SPP is effective, SPP is sensitive to large variations in whole-pattern based feature extractors. On the base of SPP, researchers combine other dimensionality reduction algorithm to overcome the defect of SPP under semi-supervised dimensionality reduction frameworks. [17] proposed a sparse representation-based classifier (SRC) [11] oriented unsupervised dimensionality reduction algorithm which combines SRC and PCA in its objective function. [18] proposed discriminant sparse neighbourhood preserving embedding (DSNPE) by adding the discriminant information into sparse neighbourhood preserving embedding. DSNPE not only preserves the sparse reconstructive relationship of SNPE, but also sufficiently utilizes the global discriminant structures. Discriminative Sparsity Preserving Projection (DSPP) [19] attempts to maintain the prior low-dimensional representation constructed by the data points and the known class labels and, meanwhile, considers the complexity of  $f$  in the ambient space and the smoothness of  $f$  in preserving the sparse representation of data. However, there is a common defect in above semi-supervised dimensionality reduction algorithms based on SR, namely, these algorithms ignore making use of supervised information to guide sparsity reconstruction of samples. In general, there are different forms of supervision information or

---

\* *Corresponding author* e-mail: chengjianjun\_2012@163.com

prior knowledge, such as class label, pairwise constraint, and others. Class label may be strong information from the users and cost us many efforts. In contrast, it is more natural to specify which pairs of data points are similar or dissimilar [20]. As a kind of side information, pairwise constraint contain must-links where the pair of data points must be in a same class and cannot-links where the pair of data points must be in two different classes [21]. The utility of pairwise constraints has been demonstrated in many applications [20-30].

Inspired by SPP and pairwise constraints, Constraints-based Sparsity Preserving Projections (CSPP) is proposed in the paper. Different from above semi-supervised dimensionality reduction algorithms based on SPP, CSPP makes use of supervised pairwise constraint information guide and adjust sparse reconstructive weights with penalty item. Experimental results on Yale, UMIST and AR show, in contrast to DSNPE and DSPP, our algorithm is more efficient.

The rest of the paper is organized as follows: Section 2 reviews sparse representation, sparse reconstruction and SPP. Our CSPP is introduced in Section 3. In Section 4, CSPP is compared with some related works. The experimental results are presented and made analyses. Finally, some concluding remarks and future work are provided in Section 5.

**2 Related Background**

**2.1 SPARSE REPRESENTATION**

Given a set of training samples  $X = \{x_1, x_2, x_3, \dots, x_n\} \in R^{d \times n}$ , sparse representation seeks a sparse reconstructive weight vector  $s_i$  for each  $x_i$  through the following minimization problem:

$$\begin{aligned} \min_{s_i} \|s_i\|_0, \\ s.t. x_i = Xs_i \end{aligned} \tag{1}$$

where  $S_{ij}$  denotes the contribution of each  $x_j$  to reconstructing  $x_i$ .  $\|s_i\|_0$  is the pseudo- $\ell_0$  norm which is equal to the number of non-zero components in  $S$ . However, Equation (1) is NP-hard. The solution of  $\ell_0$  minimization problem is equal to the solution of  $\ell_1$  minimization problem as follows:

$$\begin{aligned} \min_{s_i} \|s_i\|_1 \\ s.t. x_i = Xs_i \end{aligned} \tag{2}$$

**2.2 SPARSE RECONSTRUCTION**

Sparse reconstruction seeks a sparse reconstructive weight vector  $s_i$  for each  $x_i$  through the following modified  $\ell_1$  minimization problem:

$$\begin{aligned} \min_{s_i} \|s_i\|_1 \\ s.t. x_i = Xs_i, \\ 1 = 1^T s_i \end{aligned} \tag{3}$$

where  $\|s_i\|_1$  denotes the  $\ell_1$  normal of  $s_i$ ,  $s_i = [s_{i1}, \dots, s_{i-1}, 0, s_{i+1}, \dots, s_{in}]^T \in R^n$  is a vector in which  $S_{ij}$  denotes the contribution of each  $x_j$  to reconstructing  $x_i$ , and  $1 \in R^n$  is a vector of all ones.

$$x_i = s_{i1}x_1 + \dots + s_{i-1}x_{i-1} + s_{i+1}x_{i+1} + \dots + s_{in}x_n \tag{4}$$

The sparse reconstruction matrix  $S = [s_1, s_2, \dots, s_n]^T$  is attained through calculating  $S_i$ .

**2.3 SPARSITY PRESERVING PROJECTIONS (SPP)**

SPP aims to preserve sparse reconstruction relation of high-dimensional data space to low-dimensional data space. Given the projection matrix  $T$ ,  $T^T X s_i$  is the projection point of  $x_i$  in high-dimensional data space. The objective function of SPP is as follows:

$$\begin{aligned} \min_T \sum_{i=1}^n \|T^T x_i - T^T X s_i\|^2 \\ s.t. T^T X X^T T = I \end{aligned} \tag{5}$$

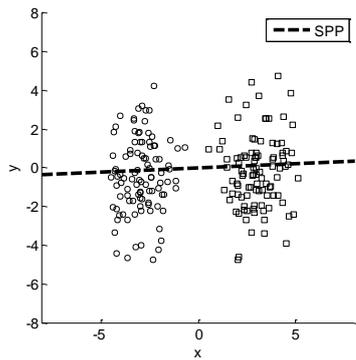
Equation (5) can be further transformed to

$$\begin{aligned} \max_T [T^T X (S + S^T - S^T S) X^T T] \\ s.t. T^T X X^T T = I \end{aligned} \tag{6}$$

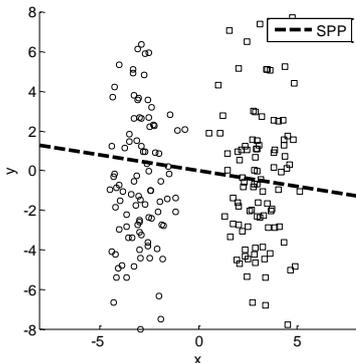
**3 Constraints-based Sparsity Preserving Projections (CSPP)**

**3.1 BASIC IDEA**

As described in section 1, in order to illustrate the problem existed in SPP, experiments of SPP on a two-dimensional dataset and the changed dataset are shown in Figure 1.



(a) SPP on a two-dimensional dataset



(b) SPP on the changed dataset

FIGURE 1 Experiments of SPP on a two-dimensional dataset and the changed dataset

Figure 1 depicts 2-dimensional 2-class examples. The circles and triangles denote the samples in positive and negative classes. The solid and dashed lines denote the 1-dimensional embedding spaces of SPP on the dataset. In contrast to the dataset in (a), the vertical scaling of the data is doubled in (b), which lead to change in the whole structure of the dataset. This change of scales affects SPP solutions, which illustrates a possible weakness of SPP arising from its unsupervised nature.

### 3.2 OBJECTIVE FUNCTION

According above analyses, it is important to introduce supervised information to guide sparse reconstruction in order to overcome the weakness of SPP. Given training samples  $X = \{x_1, x_2, x_3, \dots, x_n\} \in R^{d \times n}$ , containing a must-link (ML) set and a cannot-link (CL) set. According to Equation (3) and Equation (4), the sparse reconstructive weight matrix  $S = (s_{ij})_{n \times n}$  of samples is adjusted [13] proposed to make use of pairwise constraint supervised information to refine adjacency relations of samples with the weighted parameter way. Inspired by [13], the paper adjust the sparse reconstructive weight vector  $s_i$  of  $x_i \in X$  on the base of  $s_i$ . The adjustment of the adjusted sparse reconstructive coefficient  $\tilde{s}_{ij}$  of  $x_j$  to  $x_i$  is described as follows:

$$\tilde{s}_{ij} = \begin{cases} s_{ij} + \frac{\alpha \times n_M}{(n_M + n_C)} & \text{if } (x_i, x_j) \in ML \\ s_{ij} + \frac{\beta \times n_C}{(n_M + n_C)} & \text{if } (x_i, x_j) \in CL, \\ s_{ij} & \text{other} \end{cases} \quad (7)$$

where  $n_M$  denotes this size of *ML* and  $n_C$  denotes this size of *CL*.  $\tilde{s}_{ij}$  denotes the adjusted sparse reconstructive coefficient of  $x_j$  to  $x_i$ .  $\alpha$  and  $\beta$  denote adjustment parameters.

Equation (7) may be understood in such two sentences: if two samples are in the same class, the greater sparse reconstructive coefficient strengthens their relation as much as possible. If two samples are in two different classes, the less sparse reconstructive coefficient alienates their relation as much as possible.

According to Equation (6), with  $S = (\tilde{s}_{ij})_{n \times n}$  replacing  $S$ , the objective function is gotten as follows:

$$\max_T \frac{T^T X(S + S^T - S^T S) X^T T}{T^T X X^T T} \quad (8)$$

### 3.3 ALGORITHM STEPS

Input: face training sample  $X = \{x_i | x_i \in R^d\}_{i=1}^n$ .

Output: projection matrix  $T \in R^{d \times l} (l \leq d)$ .

Steps:

1) construct the sparse reconstructive matrix  $s$  using of Equation (3).

2) get the adjusted sparse reconstructive matrix  $s$  with pairwise constraint supervised information using Equation (7).

3) transform Equation (8) into the generalized matrix problem  $X(S + S^T - S^T S) X^T t_i = \lambda_i X X^T t_i, 1 \leq i \leq l$  and get the projection matrix  $T = [t_1, t_2, \dots, t_l]$ .

### 3.4 COMPUTATIONAL COMPLEXITY ANALYSES

Given samples  $X = \{x_1, x_2, x_3, \dots, x_n\} \in R^{d \times n}$ . CSPP contains main steps for solving  $s$  and the eigen-decomposition using Equation (8). According to Equation (7), solving  $S$  is the key part of solving the sparse reconstructive weight matrix  $s$ . The computational complexity of sparse learning is nearly that of solution of  $l_1$  norm minimization problems which is  $O(d^3)$  [31]. Therefore the computational complexity of solving  $S$  is  $O(d^3)$ . The eigen problem on a symmetric matrix can be efficiently computed by the singular value decomposition

(SVD), which is  $O(d^3)$ . Hence, the computational complexity of CSPP is  $O(d^3)$ .

4 Results and Analysis

4.1 EXPERIMENTAL DATASETS

Some following face datasets are selected in the experiment:

1) UMIST: This set contains 564 images of 20 individuals. Each face image is resized to  $112 \times 92$  pixels with 256 gray levels. The images are covering a range of poses from profile to frontal views. A group of faces in UMIST are shown in Figure 2.



FIGURE 2 A group of faces in UMIST

2) Yale: this database contains 165 face images of 15 individuals. There are 11 images per subject, and these 11 images are, respectively, under the following different facial expression or configuration: center-light, wearing glasses, happy, left-light, wearing no glasses, normal, right-light, sad, sleepy, surprised and wink. A group of faces in Yale are shown in Figure 3.



FIGURE 3 A group of faces in Yale

3) AR: this database consists of over 4000 face images of 126 individuals. For each individual, 26 pictures were taken in two sessions (separated by two weeks) and each section contains 13 images. These images include front view of faces with different expressions, illuminations and occlusions. A group of faces in AR are shown in Figure 4.



FIGURE 4 A group of faces in AR

4.2. EXPERIMENTAL SETTINGS

$L$  images are selected randomly from a group face and remains for test samples. Besides, pairwise constraints set with the size  $PC$  are created randomly from training samples. In order to eliminate the singular problem, training samples are projected into the PCA [1] subspace. The performance of the proposed algorithm is evaluated and compared with that of several methods using the Nearest Neighbour Classifier (NNC). As a baseline, the classification results of NNC directly used the raw data without dimensionality reduction is given. SPP, DSNPE and DSPP are also introduced for comparing with our algorithm. Parameters of various algorithms are set in Table 1.

TABLE 1 Parameter settings of various algorithms

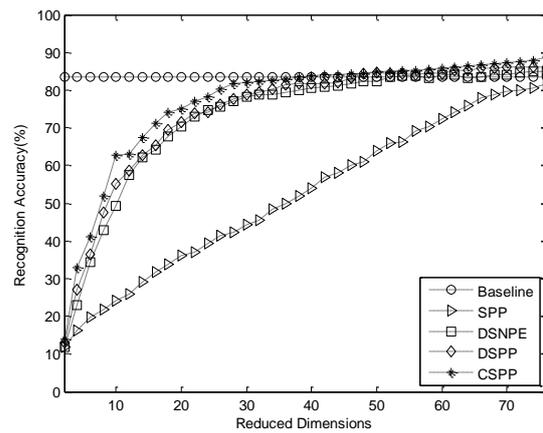
Algorithms name	Parameter settings
Baseline	no
SPP	no
DSNPE	$\gamma = 0.5$
DSPP	$\gamma_A = 0.001, \gamma_t = 1$
CSPP	$\alpha = 10, \beta = 30$

4.3 EXPERIMENTAL RESULTS

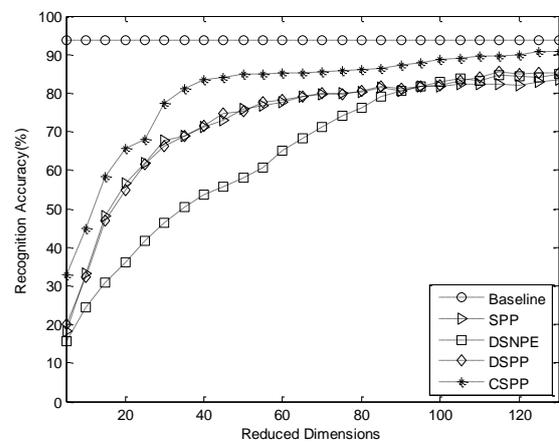
In order to verify efficiently the performance of our proposed algorithm, experiments and analyses are made under different reduced dimensions and pairwise constraints sets with the different size. All experiments are repeated twenty times and average recognition accuracy rates are gotten.

4.3.1 Effect of reduced dimension on the performance

Reduced dimensions are selected with the certain increment and corresponding average recognition accuracy are calculated. Concrete experimental results are shown in Figures 5–7.

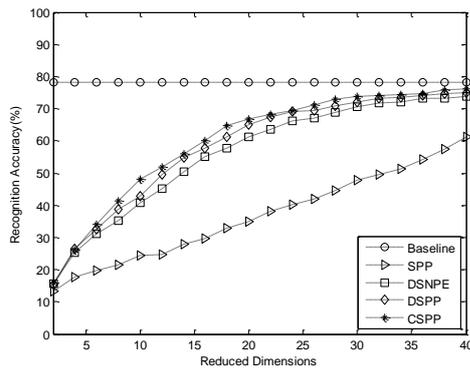


(a) L=4 and PC=800

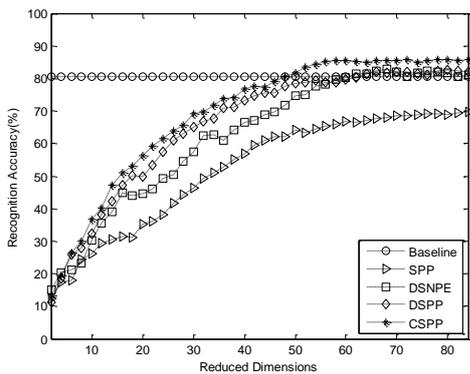


(b) L=8 and PC=1600

FIGURE 5 Recognition accuracy (%) VS. Reduced dimension on UMIST with L and PC

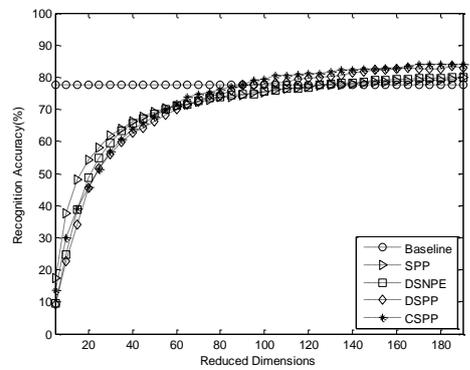


(a) L=3 and PC=200

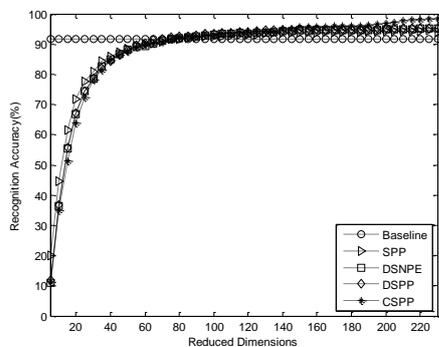


(b) L=6 and PC=400

FIGURE 6 Recognition accuracy (%) VS. Reduced dimension on Yale with L and PC



(a) L=5 and PC=2000



(b) L=10 and PC=4000

FIGURE 7 Recognition accuracy (%) VS. Reduced dimension on AR with L and PC

From Figures 5–7, the following conclusions are drawn:

1) With increase on reduced dimensions, the recognition accuracy of SPP and CSPP promote. CSPP is superior to SPP in face datasets with different character, which illustrates that the adjustment way on the sparse reconstructive weight matrix in Equation (7) is efficient.

2) Although DSNPE combines sparsity criterion and maximum margin criterion (MMC) together to project the input high-dimensional image into a low-dimensional feature vector, integrating both the robustness advantage of sparse representation and distinctiveness advantage of MMC, the recognition accuracy of CSPP still is higher than DSNPE. This show that, in contrast to the way of infusing sparse reconstruction information of SR and discriminative information of MMC, the way of making use of supervised pairwise constraints information to guide the adjustment on sparse reconstructive weight matrix is more efficient.

3) Owing to providing s an explicit feature mapping by fitting the prior low-dimensional representations which are generated randomly by using the labels of the labelled data points, DSPP has a high discriminative ability which is inherited from the sparse representation of data. However, the CSPP outperform DSPP, which is caused by the reason that DSPP pay attention to set the smoothness regularization term to measure the loss of the mapping in preserving the sparse structure of data and ignore the defect of sparse reconstruction in SR.

#### 4.3.2 Effect of the size of pairwise constraints sets on the performance

Under different L on different face datasets, pairwise constraints sets are created with the different size PC and calculated the corresponding maximum recognition accuracy. Concrete experimental results are shown in Tables 2–4.

TABLE 2 Experimental results on UMIST

L	PC	Recognition accuracy (%)
4	400	81.50
	800	85.43
	1200	86.25
	1600	86.50
8	500	85.07
	1000	87.35
	1500	89.23
	2000	90.73

TABLE 3 Experimental results on Yale

L	PC	Recognition accuracy (%)
3	100	73.87
	200	74.87
	300	75.87
	400	76.16
6	200	81.06
	400	82.13
	600	83.73
	800	84.80

TABLE 4 Experimental results on AR

<i>L</i>	<i>PC</i>	Recognition accuracy(%)
5	1000	82.87
	2000	83.19
	3000	84.15
	4000	85.65
10	2000	97.08
	4000	97.19
	8000	98.27
	10000	98.85

From Tables 2–4, some conclusions are drawn as follows:

1) With increase in the size of pairwise constraints sets, the recognition accuracy become more higher, which illustrate that more supervised pairwise constraints information is effective for adjustments of sparse reconstructive weight matrix in Equation 6.

2) Although there are less increment of the pairwise constraints set on UMIST and Yale than AR, the performance of CSPP is more sensitive to the increment

## References

- [1] Fukunaga K 1990 Introduction to Statistical Pattern Recognition 2nd ed *Academic Press* Boston USA
- [2] Duda R O, Hart P E Stork, D G 2000 Pattern Classification second ed *John Wiley & Sons* New York
- [3] Sam T R, Lawrence K S 2000 Nonlinear dimensionality reduction by locally linear embedding *Science* **290**(5500) 2323–6
- [4] Joshua B T, Vin de S, John C L 2000 A global geometric framework for nonlinear dimensionality reduction *Science* **290**(5500) 2319–23
- [5] Mikhail B, Partha N 2003 Laplacian eigenmaps for dimensionality reduction and data representation *Neural Computation* **15**(6) 1373–96
- [6] Zhenyue Z, Hongyuan Z 2003 Nonlinear dimension reduction via local tangent space alignment *In Proceedings of Intelligent Data Engineering and Automated Learning (IDEAL '2003)* Springer 21–3
- [7] Xiaoifei H, Partha N 2003 Locality preserving projections *Advances in Neural Information Processing Systems (NIPS '2003)*
- [8] Xiaoifei H, Deng C, Shuicheng Y, Hong-Jiang Z 2005 Neighborhood preserving embedding *In Proceedings Of International Conference on Computer Vision (ICCV2005)* 1208–13 IEEE Computer Society October 2005
- [9] Deng C, Xiaoifei H, Jiawei H 2007 Isometric projection *In Proceedings Of AAAI Conference on Artificial Intelligence* Wikipedia
- [10] Tianhao Z, Jie Y, Deli Zhao, Xinliang Ge 2007 Linear local tangent space alignment and application to face recognition *Neural Computation* **15**(6) 1547–53 (*in Chinese*)
- [11] Yang A Y, Ganesh A, Sastry S S, Yi M 2009 Robust face recognition via sparse representation *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**(2) 210–27
- [12] Patrik O H 2004 Non-negative matrix factorization with sparseness constraints *Journal of Machine Learning Research* **5**(11) 1457–69
- [13] Xiaorong P, Zhang Y, Ziming Z, Wei Z, Mao Y 2005 Face recognition using Fisher non-negative matrix factorization with sparseness constraints *In Proceedings of the Second international conference on Advances in neural networks (ISNN '2005)* Springer-Verlagpp 112–7
- [14] Ran H, Bao-Gang H, Wei-Shi Z, Xiang-Wei K 2011 Robust principal component analysis based on maximum correntropy criterion *IEEE Transactions on Image Processing* **20**(6) 1485–94
- [15] Ran H, Wei-Shi Z, Bao-Gang H 2011 Maximum correntropy criterion for robust face recognition *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**(8) 1561–76

on UMIST and Yale. It is reason for the problem that the performance of CSPP is influenced by the ratio of the increment of the pairwise constraints set to the size of the training sample instead of the absolute increment of the pairwise constraints set.

## 5 Conclusion

Constraints-based Sparsity Preserving Projections (CSPP) is proposed for dimensionality reduction in the paper. On the base of SPP, CSPP adjust the sparse reconstructive weight matrix through the penalty way with supervise pairwise constraints information. Experimental results on UMIST、YALE and AR demonstrate the effectiveness of our proposed algorithm. However, for CSPP, the certain size of the pairwise constraints set is needed. Although attaining pairwise constraints information is simpler than class label information, the work cost us much effort. Therefore, how to introduce new supervised information guide sparse reconstruction is our next work.

- [16] Lishan Q, Songcan C, Xiaoyang T 2010 Parsity preserving projections with applications to face recognition *Pattern Recognition* **43**(1) 331–41
- [17] Lei Z, Meng Y, Zhizhao F, Zhang D 2010 On the dimensionality reduction for sparse representation based face recognition *In Proceedings of International Conference on Pattern Recognition (ICPR '2010)* IEEE Computer Society 1237–40
- [18] Jie G, Zhenan S, Wei J, Rongxiang H, Yingke L, Shuiwang J 2012 Discriminant sparse neighborhood preserving embedding for face recognition *Pattern Recognition* **45**(8) 2884–93
- [19] Nannan G, Mingyu F, Hong Q, Bo Z 2012 Discriminative Sparsity Preserving Projections for Semi-Supervised Dimensionality Reduction *IEEE Signal Processing Letters* **19**(7) 391–4
- [20] Rong Y, Jian Z, Jie Y, Hauptmann A G 2006 A discriminative learning framework with pairwise constraints for video object classification *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**(4) 578–93
- [21] Xing E, Ng A, Jordan M, Russell S 2003 Distance metric learning with application to clustering with side-information *In Proceedings of Advances in Neural Information Processing Systems* MIT Press 505–12
- [22] Aharon B H, Tomer H, Noam S, Daphna W S 2006 Learning a mahalanobis metric from equivalence constraints *Journal of Machine Learning Research* **6**(6) 937–65
- [23] Zhang D, Zhou Z, Chen S 2007 Semi-Supervised dimensionality reduction *In Proceedings of SIAM International Conference on Data Mining (SDM ' 2007)* ACM 629–34
- [24] Chen Y, Rege M, Dong M, Hua J 2007 Incorporating user provided constraints into document clustering *In Proceedings of International Conference on Data Mining(ICDM ' 2007)* Nebraska USA 103–12
- [25] Wei J, Peng H 2008 Neighborhood preserving based semi-supervised dimensionality reduction *Electronics Letters* **44**(20) 1190–1
- [26] Hakan C, Jakob V, Frédéric J, Alexander K 2008 Semi-Supervised dimensionality reduction using pairwise equivalence constraints *In: Proc. Of International Conference on Computer Vision Theory and Applications (VISAPP '2008)* ACM 205–14
- [27] Guoxian Y, Hong P, Jia W, Qianli M 2010 Robust Locality Preserving Projections with Pairwise Constraints *Journal of Computational Information Systems* **6**(5) 1631–36
- [28] Shiguo C, Daoqiang Z Semisupervised 2011 Dimensionality Reduction With Pairwise Constraints for Hyperspectral Image Classification *IEEE Geoscience and Remote Sensing Letters* **8**(2) 369–73

[29] Mariam K, Philippe B, Ludovic M, Denis H 2011 Constraint scores for semi-supervised feature selection: A comparative study *Pattern Recognition Letters* 32(5) 656–65  
[30] Maggini M, Melacci S, Sarti L 2012 Learning from pairwise constraints by Similarity *Neural Networks* 26(10) 141–58

[31] Donoho D L, Tsaig Y, 2008 Fast Solution of  $\ell_1$ -norm Minimization Problems When the Solution May be Sparse *IEEE Transactions on Information Theory* 54(11) 4789–812

## Authors



**Jianjun Chen, born in May, 1966, Zhejiang, China**

**Current position, grades:** associate professor of Shaoxing University, Shaoxing, China.

**University studies:** Bachelor's degree in physics science from Zhejiang University in 1987, the Master's degree from Taiyuan University of Technology in 2004

**Scientific interest:** machine learning and image processing

**Publications:** 11

# An effective time of advent-based scheme for mitigating the influence of the non-line-of-sight propagation

**Qinghui Wang\*, Wangyuan Huang**

*College of Information Engineering, Shenyang University of Chemical Technology, Shenyang, 110142, China*

*Received 1 May 2014, www.tsi.lv*

---

## Abstract

Aiming at the problem that Time of Advent-based wireless location is easily influenced by non-line-of-sight, a set of effective wireless ranging system for mitigating the influence of non-line-of-sight has been proposed in this paper. At first, the Kalman filter model has been established to eliminate the influence of non-line-of-sight, as well as the random interference. Secondly, it carries out the off-line filter simulation with test data to verify the effectiveness of the model. Finally, a set of ranging system has been designed by taking the ATmega1280 microprocessor as the controller and the nanoPAN5375 as the radio frequency chip, and it can carry out the real test on the improved ranging platform. According to the results, the designed system can accomplish the real-time dynamic filter, which can reduce the non-line-of-sight error with a high measurement precision, and it can also be applied in the location under non-line-of-sight environment directly.

*Keywords:* time of advent, wireless location, non-line-of-sight, Kalman filter, nanoPAN5375

---

## 1 Introduction

The E-911 location demand released by Federal Communications Commission (FCC) in 1996 gave rise to an extensive study on the wireless location technology, which turned the location and tracking of MT to be a study field with rapid development [1]. The wireless sensor network location technology has already been applied in various fields, including the business, public security and military field, for instance, it can test the operational condition of the equipment when deploying WSN in the industrial site, it can track the logistic dynamics when deploying in warehouse, and it can even provide the optimal route for the firemen when deploying in the fire rescue scene [2, 3] etc. When compared to the GPS that enjoys the widest application at present, the wireless sensor network location system has its own advantages. At first, GPS equipment cannot work in places that GPS satellite signal cannot reach [4], such as the indoor environment, thick forest, etc. while the wireless sensor network location system is free from the restrictions of the site. Secondly, the cost of GPS is too high, and it is not applicable for the low-end and simple application site.

At present, the study on the wireless sensor network location is to make full use of the current resources in each communication standard to improve the precision of the wireless location in complicated wireless communication environment, among which NLOS transmission caused due to the shielding of physical entity is the key element that influences the location precision. Aiming at reducing the influence of NLOS

transmission on the precision of the TOA-based location, substantial studies have been carried out in both home and abroad. Reference [5] proposed the location algorithm for the particle filter, reference [6] put forward a new iterative minimum residual, reference [7] came up with the residual test, and there are also some other methods, for instance, reference [8] adopts the nonlinear least square method for solving the inequality constraints, modifies the delay caused by NLOS and locates later. All these methods can reduce the unfavourable influence brought by the NLOS error to a certain degree and improves the location precision. But these methods have a huge computation quantity, and some also need to increase the storage capacity of the system, which all increases the cost and complexity of the system.

Aiming at the NLOS transmission in the individual indoor location [9], the KF algorithm that can eliminate the NLOS error has been adopted for removing the influence of NLOS. Moreover, a set of indoor wireless ranging system has been established by taking the ATmega1280 as the processor and nanoPAN 5375 as the radio frequency chip, and the distance between the moving tag and fixed base station has been measured through symmetric double-side and two-way ranging method. Meanwhile, according to the characteristics of NLOS, KF model has been established on the basis of distance, which will carry out the off-line algorithm verification for the test data through matlab simulation software, and then the algorithm will be realized specifically on the designed ranging platform. According to the experiment, for this wireless ranging system, the algorithm is easy to be realized, with strong instantaneity

---

\* *Corresponding author* e-mail: wangqh@syuct.edu.cn

and high precision, and it can also be applied for indoor location directly.

**2 System composition**

The ranging system mainly relies on the anchor of the wireless ranging function and moving tag to accomplish the ranging function, while the data collector control software operates in PC to display and save the distance data measured timely and after filtering, among which, the reference node and moving tag share the same hardware system, consisting of the ATmega1280 SCM[10], radio frequency measurement module nanoPAN 5375 [11] and the interface converting chip. The radio frequency module employs the CCS [12] to finish the communication. In this paper, it mainly focuses on the study of the real-time filtering for the distance obtained by the reference node and moving tag under NLOS, aiming to achieve more accurate distance data.

**2.1 TOA-BASED RANGING MODEL UNDER NLOS**

TOA is a ranging method based on the arrival time [13], which will calculate the distance from the nodes according to the transmission speed and time of the signal. Suppose the reference point participating in the ranging is *A* and the moving tag is *B*, the reference point *A* is fixed, and the moving tag *B* moves towards *C* in uniform linear motion. Establish the model as shown in Figure 1, in which *D* is human or barrier, and it can form the NLOS transmission effectively.

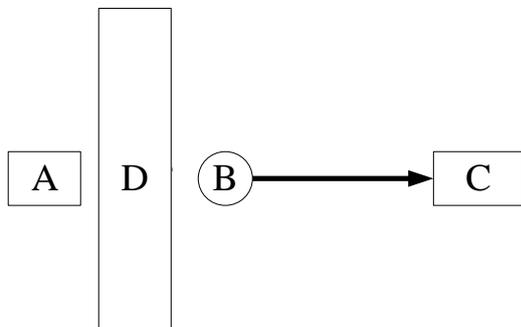


FIGURE 1 TOA-based ranging model under NLOS

Since the direct route of the signal is blocked by the barrier, it delays the signal transmission time, leading to the deviations in results and even zero value, and finally it will decrease the location precision.

**2.2 RANGING PRINCIPLE**

In this paper, the studied ranging algorithm is SDS-TWR ranging method [14] (also called symmetric double-side and two-way ranging method). SDS-TWR method can reduce the ranging error brought by the non-synchronization and drift, which mainly carries out a reverse TWR process on the basis of TWR ranging method. Thus, four time values can be obtained.

The principle of SDS-TWR is shown in Figure 2.

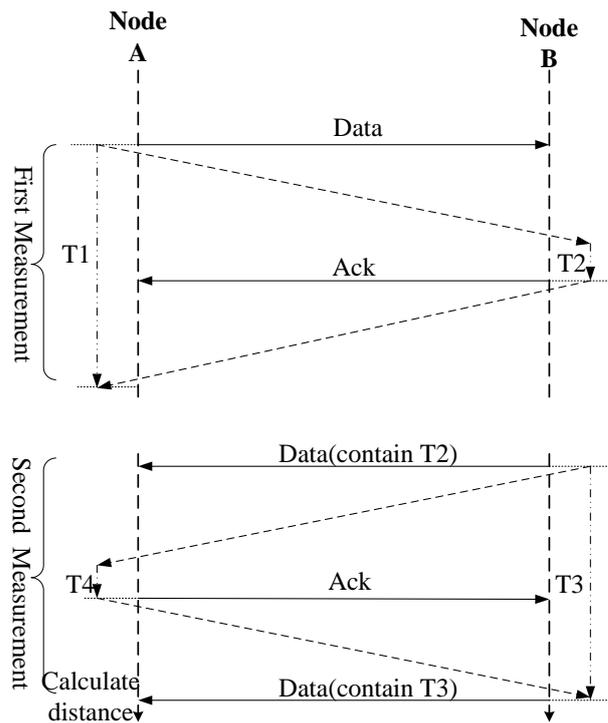


FIGURE 2 Symmetric Double Sided Two Way Ranging

In this method, it includes two symmetric measurements. Firstly, the time from sending data package by node *A* to receiving the package is  $T_1$ ; the node *B* will start timing after receiving the data from node *A*, and stop timing after returning the reply, suppose the processing delay of node *B* is  $T_2$ ; the round-trip time of the data package will be obtained by subtracting  $T_1$  from  $T_2$ . Secondly, node *B* will send data package, and node *A* will receive and reply, suppose the time measured by node *B* and node *A* is  $T_3$  and  $T_4$  respectively, and then the round-trip data of the data package is  $T_3 - T_4$ . Suppose the transmission speed of signal in the medium is  $c$ , then the theoretical distanced between the two nodes are:

$$d = c \times \frac{(T_1 - T_2) + (T_3 - T_4)}{4}, \tag{1}$$

As shown in Figure 1, the reference point *A* achieves the distance from the moving tag *B* according to SDS-TWR, and then carries out filtering for the distance data obtained under NLOS through KF algorithm.

**3 Algorithm implementation**

**3.1 NLOS ERROR AND KF MODEL**

During the process of radio wave transmission, it will reflex or reflect if there are barriers, so that the TOA is extremely delayed. That is to say, it will produce NLOS error.

Under NLOS, the ranging equation can be represented by:

$$R(t_i) = N(t_i) + T(t_i) + S(t_i), \quad (2)$$

where the  $R(t_i)$  refers the distance measurement from moving tag  $B$  to base foundation  $A$  at  $t_i$  moment (obtained by multiplying the transmission speed with the transmission time),  $N(t_i)$  stands for the real NLOS error, since the error caused by NLOS is positive deviance, then  $N(t_i) > 0$ ;  $T(t_i)$  stands for the real distance from the tag to the foundation base value,  $S(t_i)$  stands for the standard ranging error, and it is a random process with zero average value.

The KF model can be designed according to equation (2), in which  $z(t)$  stands for the TOA measured value sequence, namely  $R(t_i)$ .

The related parameter settings of KF Equation (3) and (4) are shown as follows:

$$x(t+1) = Ax(t) + p(t), \quad (3)$$

$$z(t) = Gx(t) + h(t), \quad (4)$$

$$A = \begin{bmatrix} 1 & \Delta \\ 0 & 1 \end{bmatrix}, \quad G = [1 \quad 0], \quad x_t = \begin{bmatrix} R(k) & \dot{R}(k) \end{bmatrix}^T, \\ p(t) = [0 \quad u_n], \quad h(t) = n_m.$$

In which,  $R(k)$  is the TOA waiting to be estimated,  $\dot{R}(k)$  is the first derivative of  $R(k)$ ,  $\Delta$  is the sampling gap of KF,  $u_n$  is the process noise component, and  $n_m$  is the measurement error.

### 3.2 CHARACTERISTIC ANALYSIS

At first, since the NLOS error distribution in TOA measurement is related to the barrier distribution on the radio wave propagation route, NLOS error has a random characteristic. The randomness gives rise to the dramatic changes in NLOS error, so that the deviation in TOA measurement value is extremely huge. These measurement value greatly influenced by NLOS will severely impact the correct estimation for TOA. As a result, if the influence including relatively huge error measurement can be eliminated, it can remove the NLOS error to a great extent.

### 3.3 MEASURED VALUE DROPPING METHOD AND OVERALL SHIFTING METHOD

The measured value dropping method refers to dropping the measured value with great deviations through KF. It will judge if the deviation in the current measured value is within the acceptable error range through comparing the deviation with the threshold value. The overall shifting method makes advantage of the NLOS error and makes the overall measured value has a feature of

positive deviation, and offsets the overall estimation downwardly to eliminate the NLOS error. It has advantages in eliminating the NLOS error of measured value with small deviation, but it cannot eliminate the influence of measured value with a huge deviation on the follow-up estimation value effectively.

### 3.4 REALIZATION OF ALGORITHM SOFTWARE

The system software mainly includes the main program, initial sub-program and functional sub-program, among which, the initial sub-program mainly includes the initialization of clock module, drive initialization of radio frequency chip and initialization of application level, while the functional sub-program mainly includes the ranging sub-program, KF sub-program and display sub-program. The ranging sub-program is mainly applied in receiving and sending signal, as well as achieving the distance data. The KF sub-program is mainly responsible for carrying out the mathematic processing for the distance data obtained, thus to improve the ranging precision. The displaying sub-program is mainly responsible for sending the unprocessed distance data and filtered distance data to the serial port, and then displays it on the upper computer software.

The whole filtering algorithm is realized in reference point A, and it is pre-processed with the measured-value dropping method and overall shifting method, and then the KF algorithm is employed to carry out the dynamic filtering.

## 4 Design of system hardware

The system employs the nanoPAN 5375 radio frequency module of the German nanotron Company, responsible for generating the chirp spread spectrum signal. The master control module employs the ATmega1280 SCM, responsible for controlling the work of nanoPAN 5375 module.

The wireless communication module employs the nanoPAN5375 module with 20dBm radio frequency power, and the longest communication distance can reach 800 meters. The module integrates the nanoLOC TRX receiver, radio frequency switch, power amplifier and other matching and conditioning unit, which is applicable for long-distance measurement and communication. The unique CSS technology of nanotron has been employed to measure the distance between the two points accurately. Within sight distance, the ranging precision is within 0.6 and 1.5 meters. The module provides three un-overlapped 2.4 GHz ISM channels that can be adjusted freely, and it supports several independent physical networks, and improves the co-existence with the current 2.4 GHz wireless technology [15]. Meanwhile it can provide reliable data communication with optimal transmission range. Therefore, the module can be widely applied in the LBS, as well as the high-precision ranging, real-time location and wireless sensor network.

**5 Experiment and analysis**

In this paper, the test data has been employed to testify the feasibility of the whole plan through simulation and real system test.

**5.1 ALGORITHM OFF-LINE SIMULATION**

The system firstly carries out the ranging for moving tag B according to Figure 1 without adding any algorithm. Suppose in the experimental environment, the coordinate of the reference point A is (0, 0), the cabinet is fixed in (1, 0) for making the NLOS transmission between A and B, in which the moving tag B (2, 0) moves to C (13, 0) along the x axis with 1 m/s speed, and the distance data obtained will be stored. It will take advantage of matlab simulation software to write the KF program to process the distance data obtained with offline filtering, thus to verify the effectiveness of the algorithm. The simulation result is shown in Figure 5, in which the sampling time is 20 ms, the threshold is 2.5, the horizontal coordinate stands for the quantity of the distance data obtained, about 575, while the vertical coordinate stands for the moving distance of the tag, and the unit is meter.

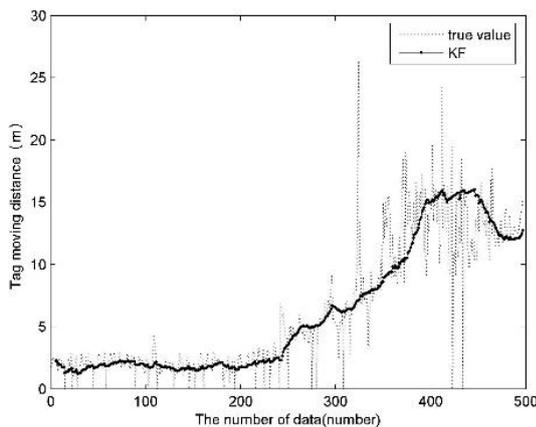


FIGURE 3 True value and filtering of the offline data

In Figure 3, it can be seen obviously that under NLOS, the real test value shakes severely, and there is also zero value. After it is processed with KF algorithm, it can decrease the fluctuation effectively, and eliminate the zero value well, thus it can demonstrate the effectiveness of the algorithm completely.

**5.2 INFLUENCE OF THRESHOLD ON THE MEASURED VALUE DROPPING METHOD**

The selection of threshold value in the measured value dropping method will have a great influence on the measurement results. It will not be able to eliminate the TOA measurement value with great errors if it is set too high, or it will eliminate too much measurement value, and the estimation value will not change with the measurement value, thus the KF will not converge. In Figure 4, it is the offline filtering simulation when the

threshold value is 0.3, while in Figures 5 and 6, it is the offline filtering simulation when the threshold is 3 and 6 respectively. It can be seen from Figure 4 that the threshold value is too small, and the estimation value does not change according to the measurement value, resulting that the KF does not converge. In Figure 5, it is the normal condition about the threshold value selection. In Figure 6, since the threshold value is too great, some great TOA measurement value cannot be eliminated, thus it cannot eliminate the influence of NLOS on TOA estimation, and the KF cannot play a good role.

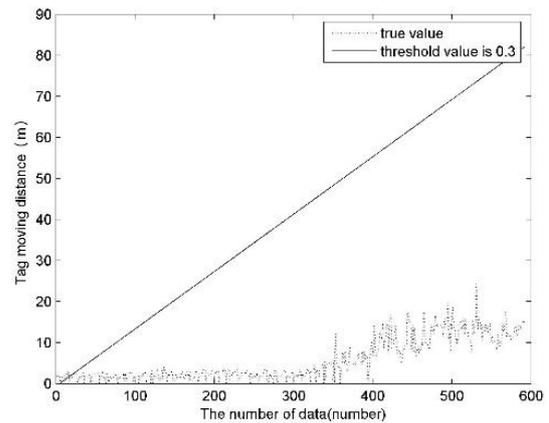


FIGURE 4 Threshold value is 0.3

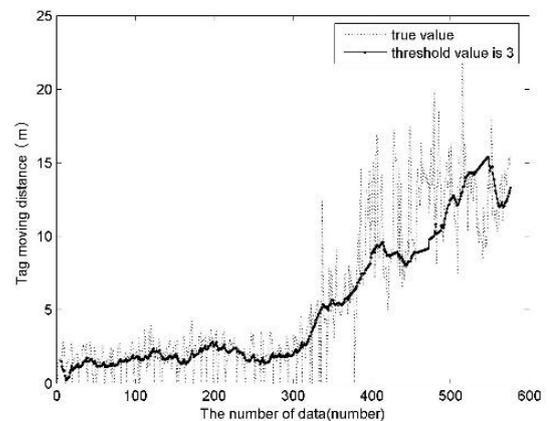


FIGURE 5 Threshold value is 3

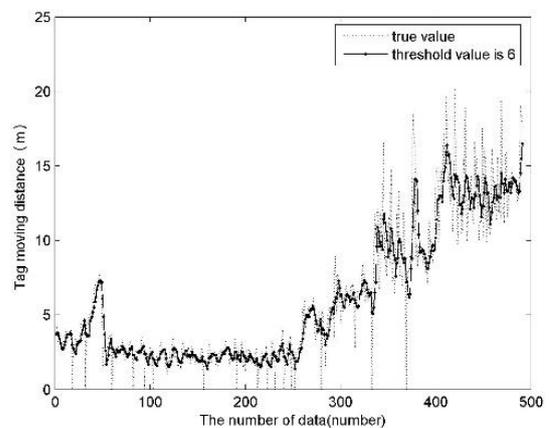


FIGURE 6 Threshold value is 6

In the offline model, the threshold value will be appropriate between 1.5 and 4. However, the threshold value obtained through offline filtering cannot be put into practice. In order to select the threshold value appropriate for the experimental environment, it should pass substantial on-site real tests to determine appropriate threshold value. The sampling time of the system is 20 ms, and the threshold between 0.3 and 1.8 is relatively appropriate for the experimental environment, as shown in Figure 7, the threshold value is 0.9, and from which it can be seen that KF plays an excellent role.

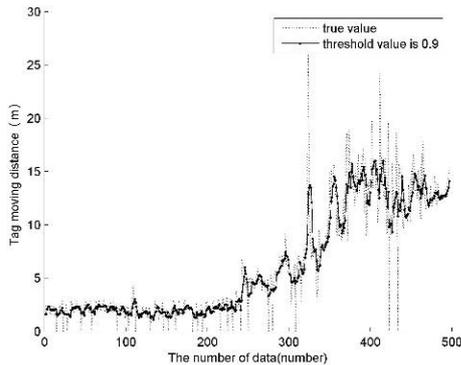


FIGURE 7 Threshold value is 0.9

### 5.3 REAL SYSTEM TEST

In this paper, the designed ranging system is employed to carry out the real-time filtering test, according to the test environment and procedure in 4.1, and it is applied to prove the algorithm is feasible in the real system test. The sampling time selected is 20 ms, the threshold value is 0.3, after comparing the distance before and after the filtering, the result is shown in Figure 8, in which the quantity of the distance data is 575.

### References

- [1] Shikur B Y, Farmani M, Weber T 2012 TOA/AOA/AOD-based 3-D mobile terminal tracking in NLOS multipath environments *9th Workshop on Positioning Navigation and Communication* 201-5
- [2] Lorincz K, Welsh M, MoteTrack 2007 A robust decentralized approach to RF-based location tracking *Personal and Ubiquitous Computing* 11(6) 489-503
- [3] Lee Y W, Stuntebeck E, Miller O C 2006 MERIT: mesh of RF sensors for indoor tracking *3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks* 2 545-54
- [4] Benini A, Mancini A, Longhi S 2013 An IMU/UWB/Vision-based Extended Kalman Filter for Mini-UAV Localization in Indoor Environment using 802.15.4a Wireless Sensor Network *Journal of Intelligent & Robotic Systems* 70(1-4) 461-76
- [5] Jung J, Myung H 2011 Indoor localization using particle filter and map-based NLOS ranging model. *International Conference on Robotics and Automation (ICRA) IEEE* 5185-90
- [6] Fukuda K, Okamoto E 2012 Performance improvement of TOA localization using IMR-based NLOS detection in sensor networks *International Conference on Information Networking* 13-18
- [7] Chan Y T, Tsui W Y, So H C, Ching Pak Chung 2006 Time-of-arrival based localization under NLOS conditions, *IEEE Transactions on Vehicular Technology* 55(1) 17-24
- [8] Wu S, Li J, Liu S 2011 An improved reference selection method in linear least squares localization for LOS and NLOS *Vehicular*

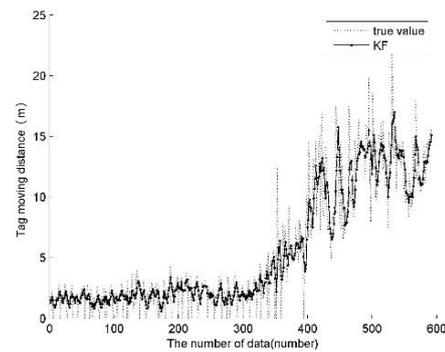


FIGURE 8 The system measured and filtering data

It can be seen from Figure 8 that since the dramatic shaking in the measured value caused by NLOS generates a huge error, and the filtering method is adopted to carry out the real-time processing. When there is a measured value with huge error, the measured value including great NLOS error will be dropped during the iteration process. Meanwhile, it can also eliminate the zero value caused by the inaccessible signal, which reduces the influence of NLOS error.

### 6 Conclusion

In this paper, the TOA-based scheme for eliminating the influence of NLOS proposed establishes the KF model and carries out the pre-processing for the data according to the statistical characteristics of NLOS error, and it reduces the measurement errors effectively. Through offline and real system test, the results have shown that the algorithm receives obvious effect in reducing the NLOS error, improves the ranging precision and provides good data support for further ranging-based location application.

*Technology Conference (VTC Fall) IEEE* 2011 1-5

- [9] Long T 2013 Research on Indoor Positioning System Based on the Application Weighted Screening Method *Information Technology Journal* 12(22) 6717-22
- [10] Cheng X Y, Jiang C L, Li M 2013 Experimental Evaluation of Networked Munitions Communication and Localization *Advanced Materials Research* 774 1690-4
- [11] Charrow B, Michael N, Kumar V 2013 Cooperative multi-robot estimation and control for radio source localization *Experimental Robotics Springer International Publishing* 337-51
- [12] Wang Q, Jiang J 2013 Performances of trigonometric chirp spread spectrum modulation in AWGN & rayleigh channels *Proceedings of the 8th ACM workshop on Performance monitoring and measurement of heterogeneous wireless and wired networks* 2013 203-10
- [13] Shen J, Molisch A F, Salmi J 2012 Accurate passive location estimation using TOA measurements *IEEE Transactions on Wireless Communications* 11(6) 2182-92
- [14] Oh D, Kwak M, Chong J W, 2012 A Subspace-Based Two-Way Ranging System Using a Chirp Spread Spectrum Modem Robust to Frequency Offset. *IEEE Transactions on Wireless Communications*, 11(4) 1478-87
- [15] Gao L, Lan Y D 2013 Transmission Distance Estimation and Testing for 2.4 GHz ZigBee Applications *2013 Fourth International Conference on Emerging Intelligent Data and Web Technologies (EIDWT) IEEE* 27-32.

**Authors**

**Wang Qinghui, born in 1972, Heilongjiang, China**

**Current position, grades:** Faculty in the College of Information Engineering, Shenyang University of Chemical Technology, Shenyang since 2002  
**University studies:** Ph.D. Degree from the Computer Science Department, Northeastern University, China, in 2006  
**Scientific interest:** Wireless sensor network, embedded system



**Huang Wangyuan, born in 1988, Hunan, China**

**Current position, grades:** M.S. degree student in the College of Information Engineering, Shenyang University of Chemical Technology  
**University studies:** B.S. Degree from the the Department of Information Engineering, Huaihua University, in 2012  
**Scientific interest:** Wireless network protocol design, and mobile computing

# Incomplete character recognition technology in the license plate recognition system

**Kai Song, Caihong Niu\***

*Information Science and Engineering College, Shenyang Ligong University, Shenyang, China*

*Received 1 May 2014, www.tsi.lv*

---

## Abstract

The incomplete characters recognition in the license plate, characters can be divided into linear character and curve character. For the curve characters, used a method of statistic the number of character holes as the character characteristics for feature extraction, which extended seed filling algorithm in the Computer Graphics. For linear characters, we proposed a method of extracting a character conversion slope of the line feature by Hough transform, which had a good effect on linear character recognition.

*Keywords:* license plate recognition; extraction of character feature; incomplete character recognition

---

## 1 Introduction

An important sign for license plate recognition is the recognition accuracy, under normal circumstances, the accuracy can reach 90% or even above 95%, but in the practical applications, There are many factors caused accuracy reduced, including artificial factors and natural factors, such as some drivers deliberately paint brush in license plate, with tape paste the license plate, place the preparation tire cover license plate, these are artificial factors, or rainy day, muddy, pollution, the shortage of parked position is wrong, and insufficient light can also result license plate fuzzy, In this case, it is difficult to accurately identify a license plate number [1].

Therefore, extracting effective and easy to implement feature is particularly important for incomplete characters. We presented a character feature extraction method, which can effectively increase the accuracy of fuzzy license plate recognition. The method by extracting characters of part features and build character database, compare the part characteristics of the incomplete characters with the database, and thus achieve the goal of improve the incomplete character recognition.

The commonly used character feature extraction include statistical feature extraction and topological feature extraction, not all these methods of feature extraction for incomplete characters have high effective degree, so bearing this in mind, extracting the number of character holes and character strokes straight slope characteristics as a key object of study. Firstly, through the license plate location, tilt correction and segmentation process, extracting independently the license plate character, and then through the character of thinning algorithm processing to extract characters, at this point, we can extract the feature of character [2, 3]. The number of character holes and Hough transform algorithm for line

slope detailed below instructions. It obtained better effect on the incomplete character recognition by these two features [4].

## 2 The number of holes of character for feature extraction

This article presents a method that search the number of circles and straight lines as characteristics, the main idea of this method is that since many of the characters have circles or holes, we use a recursive seed filling algorithm to fill these holes statistically, [5] determines the number of characters circles as a feature of the character.

This article uses the recursive seed-filling algorithm in computer graphics to calculate the number of character circles [6]. This algorithm firstly assumes that some points in the closed contour line is known, and then search the adjacent points around the seed point, and these points must be in the contour lines. If it is found that adjacent point is not in the contour line that is the outline of the edge; if it is found that the adjacent point in the contour line, this point will become a new seed point, and continue to search as above method.

4-connected algorithm and 8-connected algorithm: seed filling algorithm is essentially a recursive algorithm, It specify a seed point firstly and use the seed point as the benchmark to search in all directions, each pixel is processed, until it meet the boundary, the differences among all sorts of seed filling algorithm are in the colour and edge processing mode. We introduced two concepts, 4-connected algorithm and 8-connected algorithm. We start from any point in the region, if it was just for up, down, left and right to search the four directions and reach any pixel of the region, so the region is filled by this method is called the 4-connected regions, the filling method is called 4-connected algorithm. If we trigger

---

\* *Corresponding author* e-mail: niucaihong1010@163.com

from any one point in the region, and search the up, down, left, right, bottom-left, top-left, bottom-right and top-right under the eight directions and reach any pixel

of the region, the region is 8-connected regions, this algorithm is called 8-connected algorithm.

Recursive seed filling algorithm process is shown in Figure1 below.

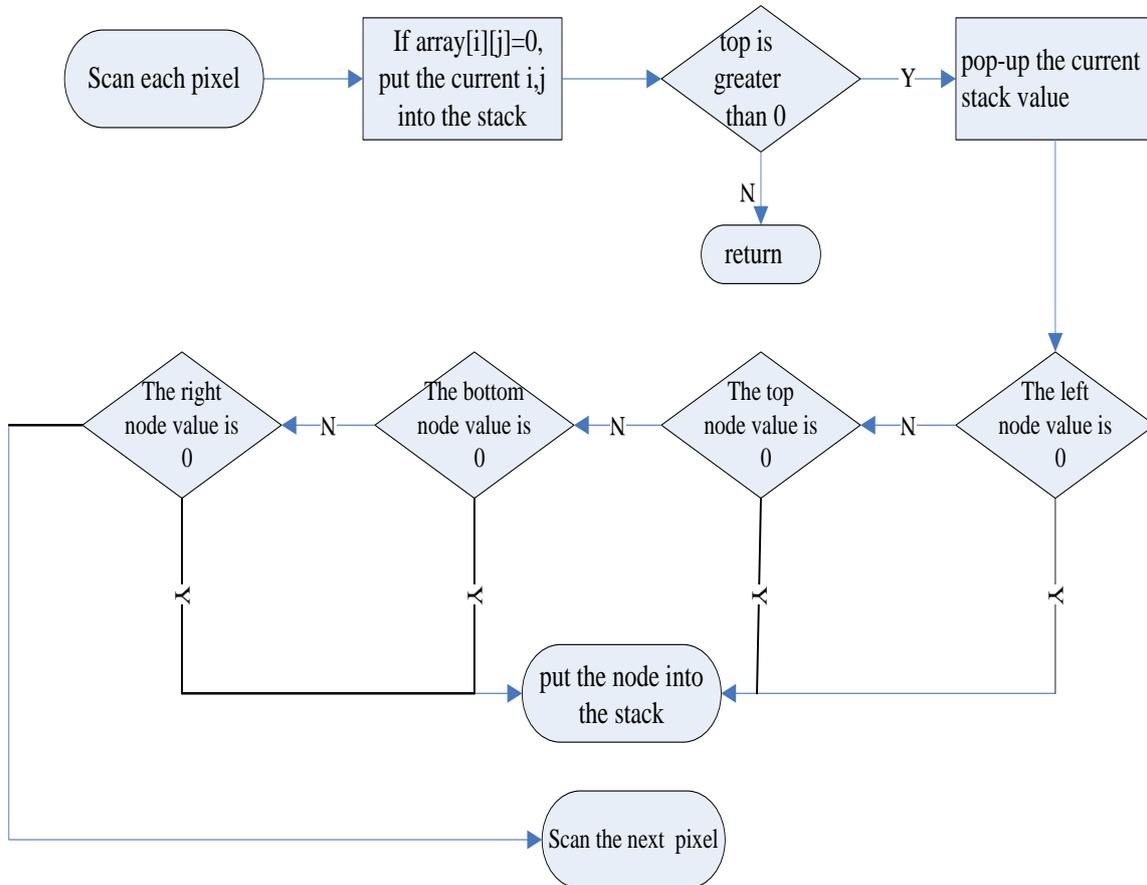


FIGURE 1 The algorithm process of extracting the characters' feature

We use the 4-connected filling algorithm, Specific steps of the algorithm are:

- 1) Selecting seed , push the seed into the stack. Scan each pixel value and set up two stack stack1 and stack2, if the image data array  $[i][j] = 0$ , then the pixel is a white pixel, select it as a seed, the horizontal and vertical coordinates were pushed into stack1 and stack2, so  $stack1[+ + top] = i$ ,  $stack2[top] = j$ .
- 2) If the stack is empty or the stack top value is less than 0, then jump out.
- 3) If the stack is not empty, then the pixels into a filling colour, namely the array  $[i][j] = flag$ , the flag value represents a kind of colour , and judge whether the four connected pixels adjacent to the pixel are boundary colour or have set into a polygon fill colour, if not, then push the pixel coordinates into stack, as a new seed. If  $array[i][j-1] = 0$ , then push the top node into the stack,  $stack1[+ + top] = i$ ,  $stack2[top] = j$ . If  $array[i-1][j] = 0$ , then push the left node into the stack, if  $array[i+1][j] = 0$ , then push the right node into the stack , If  $array[i][j+1] = 0$ , then push the bottom node into the stack, when a hole filling is

completed, the flag automatically add 1, as a new colour, repeat the first two steps of the algorithm.

- 4) End of the algorithm.

The Figures 2-4 show the results that deal with seed filing algorithm , we can see the different circles are dyed different colours, we take advantage of this to calculate the number of circles in a character, and calculate where these circles in the character according to the colour.

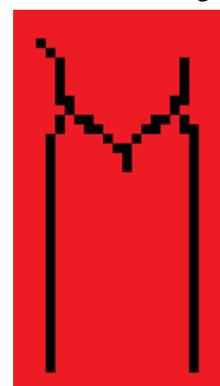


FIGURE 2 The character with no circle

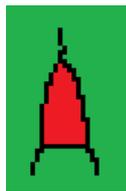


FIGURE 3 The character with one circle

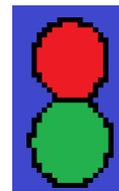


FIGURE 4 The character with two circles

The algorithm process of extracting the feature of the characters is shown in Figure 5.

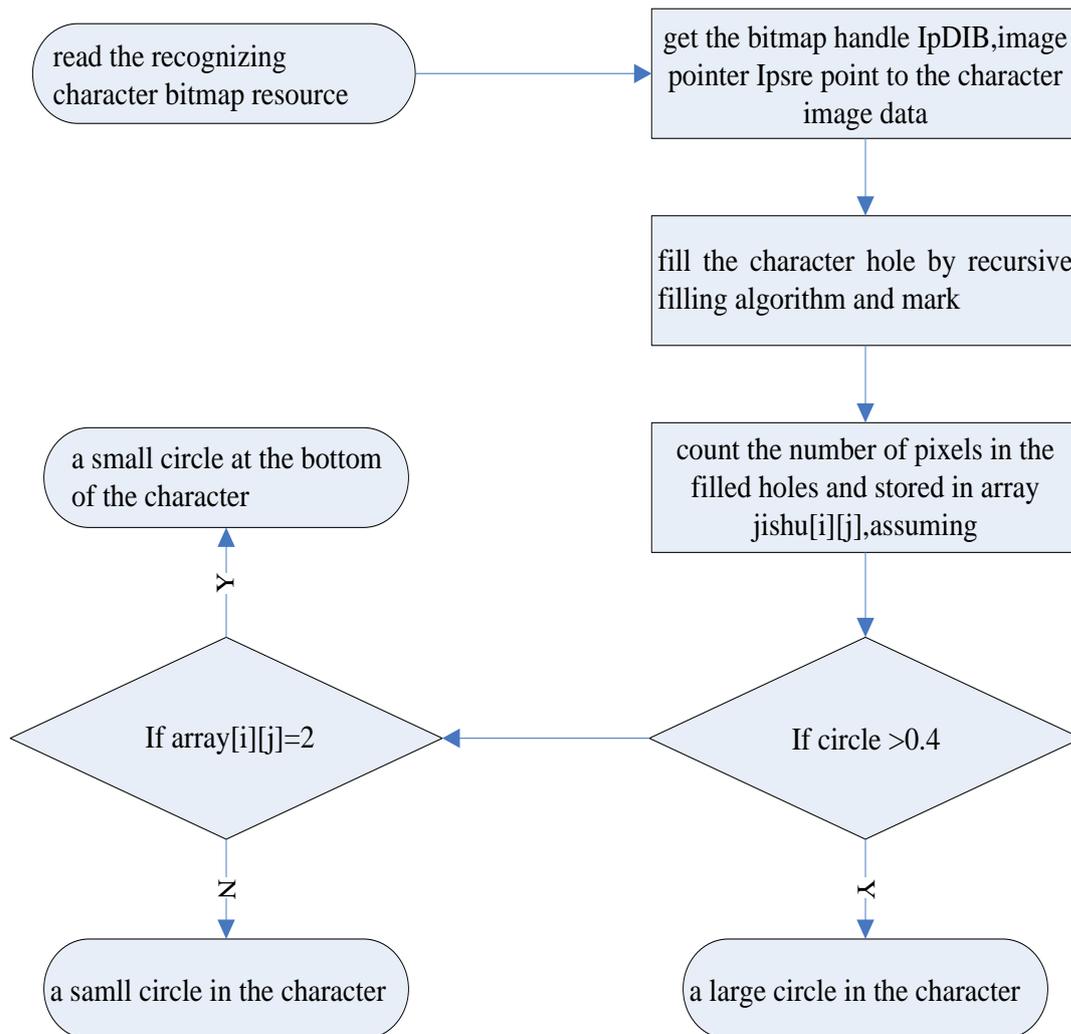


FIGURE 5 The algorithm process of extracting the characters' feature

### 3 Feature extraction of the character slope

According to the above discussion on character, we know its ability of recognition for the complete character is good, but for the incomplete character is rather poor, in order to improve the system's ability to recognize, this section is put forward four character features in view of the incomplete characters, which are the character stroke slope feature, the angle feature of the stroke intersection, the central angle feature and the curvature feature [5, 6]. For incomplete characters, we need to extract the local features, because we cannot estimate the damaged parts

and degree of disability of the characters. Local features will have a better recognition effect. For example, the slope feature of the stroke, as long as we have a bit remnants of the straight line in the direction can we find the slope and recognize. But this kind of method have a large amount of calculation, it has influence on the system speed, so we should combine this feature with the previous feature to recognize the characters, so that you can balance the ability of recognition and the speed of recognition.

We use the method of Hough transforms to detect the straight line in the character and find the slope of the line

[7, 8]. Due to the principle of Hough transform has been elaborated in the past, no longer described here, the reader can refer to relevant reference material [18].

Usually the process of detecting lines by Hough transform is as shown in Figure 6 below.

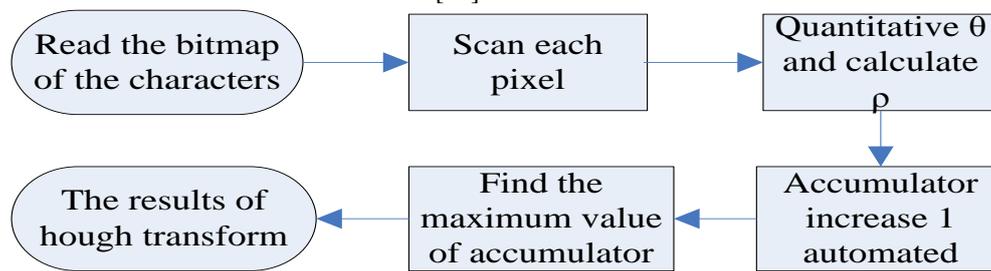


FIGURE 6 The process of detecting lines by Hough transform

1) Constructing an accumulator  $A(\rho, \theta)$ . Each element of the array represents the transform space location of the point and the number of straight lines through this point, initialize  $A$ , each element is zero. Here we apply for a pointer to the transform region, and apply large enough memory space for accumulator through the pointer,  $\text{int}^* \text{lpTransArea} = (\text{int}^*) \text{malloc}(\text{lwidth} * \text{lheight} * \text{sizeof}(\text{int}))$ , then we can save pointer address of the transform region  $\text{lpTrans} = \text{lpTransArea}$ , finally the accumulator reset  $\text{memset}(\text{lpTrans}, 0, \text{lwidth} * \text{lheight} * \text{sizeof}(\text{int}))$ .

2) If the linear feature points were detected, then corresponding to each feature point  $(x, y)$  calculated the value of  $\theta$  at 0 degree to 180 degrees, loop in each pixel scanning line, if the value of the data pointer  $\text{lpSrc}$  is 0, it indicates that the pixel is a black pixel, we use variable  $I$  AngleNumber represents the quantitative value of  $\rho$  and calculated the value of  $\rho$  from 0 degree and 180 degrees using  $\rho = x \sin \theta + y \cos \theta$ .

3) The value of the corresponding accumulator adds one. We will make the pointer point to transform region as the base pointer and make  $*(\text{lpTransArea} + \text{iDist} * \text{iMaxAngleNumber} + \text{iAngleNumber})$  add one. This expression represents the corresponding memory space of the accumulator.

4) We find the maximum value in the accumulator, the value represents the linear parameter what we're detecting. Using the parameters to find the slope of the straight line  $k = -\tan \theta$  and put the value of slope into the feature database.

In order to improve the precision and speed, we adopt an improved linear detection method, algorithm is as follows:

1) Searching in the image, find out a non-zero value point  $p$  as a seed.

2) Make  $p$  as the starting point of a rectangle, the size of the rectangle is  $M * N$ , search the non-zero points in the rectangle, then calculate linear parameters  $(\rho, \theta)$  of each non-zero point.

3) Make the deviation value of  $\rho$  for  $\Delta\rho$  and  $\theta$  for  $\Delta\theta$ , we vote selection in the rectangle, then statistics the number of parameters within  $(\rho_i + \Delta\rho, \theta_i + \Delta\theta)$ , we get the value  $n_i$ .

4) Find the vote maximum  $n_{\max}$  range in this region and average to the parameter which belongs to the range,

so it can reduce the quantization error, the average value of this parameter is the line parameter which through a point  $p$  in this region.

5) Set  $T_1$  to line length threshold value, if  $n_{\max}$  is bigger than  $T_1$ , go to step 6, search other points of the line; Otherwise thinking about the line which through point  $p$  does not exist, reset, go to step 1 and search again.

6) Search point  $p$  in the image and calculate the corresponding value of  $\rho$ , if  $\rho - \bar{\rho} < \Delta\rho$ , then count the number of cumulative voting and set the grey value of this point to 0.

7) After the full search, if the number of line vote is bigger than the setting threshold value, then we think this line exist, then find the parameters of the line and put them into the array  $\text{line1}[n]$ ,  $\text{line2}[n]$ , these two arrays store the value  $\rho$  and  $\theta$ .  $N$  is the number of detected line.

According to this algorithm, we can calculate the number of all lines in the binary images. In this way, we counted the number of straight lines, [9-11] and found the slope of the straight lines, recorded the slope of each character one by one and put them into the feature database [12-15].

For a complete or an incomplete character, we can still extract the feature of slope by Hough transform.

Experimental results are as follows:

1) Searching in the image, find out a non-zero value point  $p$  as a seed.

2) Make  $p$  as the starting point of a rectangle, the size of the rectangle is  $M * N$ , search the non-zero points in the rectangle, then calculate linear parameters  $(\rho, \theta)$  of each non-zero point.

3) Make the deviation value of  $\rho$  for  $\Delta\rho$  and  $\theta$  for  $\Delta\theta$ , we vote selection in the rectangle, then statistics the number of parameters within  $(\rho_i + \Delta\rho, \theta_i + \Delta\theta)$ , we get the value  $n_i$ .

1) Find the vote maximum  $n_{\max}$  range in this region and average to the parameter which belongs to the range, so it can reduce the quantization error, the average value of this parameter is the line parameter which through a point  $p$  in this region.

2) Set  $T_1$  to line length threshold value, if  $n_{\max}$  is bigger than  $T_1$ , go to step 6, search other points of the

line; otherwise thinking about the line which through point  $p$  does not exist, reset, go to step 1 and search again.

3) Search point  $p$  in the image and calculate the corresponding value of  $\rho$ , if  $\rho - \bar{\rho} < \Delta\rho$ , then count the number of cumulative voting and set the grey value of this point to 0.

4) After the full search, if the number of line vote is bigger than the setting threshold value, then we think this line exist, then find the parameters of the line and put them into the array line1 [n], line2 [n], these two arrays store the value  $\rho$  and  $\theta$ . N is the number of detected line.

According to this algorithm, we can calculate the number of all lines in the binary images. In this way, we counted the number of straight lines, [9-11] and found the slope of the straight lines, recorded the slope of each character one by one and put them into the feature database [12-15].

For a complete or an incomplete character, we can still extract the feature of slope by Hough transform.

Experimental results are as follows:



FIGURE 7 The Slope feature extraction of incomplete "A"

Because of the quantization error, it will lead to some errors, but this is entirely result from the quality of the refinement algorithm, if the processing effect for the part of character refinement algorithm is better, it will greatly reduce the problem. FIGURE 8. The Slope feature extraction of incomplete "N" reduces the problem.

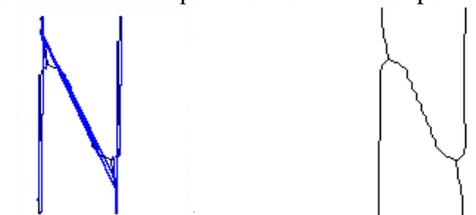


FIGURE 8 The Slope feature extraction of incomplete "N"



FIGURE 9 The Slope feature extraction of incomplete "4"

Because the origin of coordinates in the image is starting from the up-left corner, we will use this coordinate system, so the calculated value of slope may be different from common coordinate system, but there is no effect on the slope of the character. The number 4 with

two special lines is respectively the horizontal and vertical straight line, we calculated the slope of a straight line, respectively  $k_1 = 2.14286$ ,  $k_2 = 2.18182$ ,  $k_3 = 2.05$ . We calculated the average value  $k = 2.1249$  of the three. The algorithm has to meet the requirements of extracting its feature of slope for linear character.

Finally, we give some characters slope characteristics and the range of error, the slope of the straight line, horizontal and vertical not considered here. [13, 14] Selecting  $(k-0.3, k+0.3)$  for the error interval, if the slope of the straight line fall in this range, judging from the value of  $K$ , we can know the character, the value of  $K$  represents the value of slope for the character.

- 1) The slope of a slash in the number 2 is  $k = 1.4$ ;
- 2) The slope of a line in the number 4 is  $k = 2.1249$ ;
- 3) The slope of two lines in the letter A is  $k_1 = 5.25$ ,  $k_2 = 5.25$ ;
- 4) The slope of two lines in the letter K respectively is  $k_1 = -1.85$ ,  $k_2 = 4.4$ ;
- 5) The slope of two lines in the middle of the letter M respectively is  $k_1 = 1.107$ ,  $k_2 = 1.18$ ;
- 6) The slope of a line in the letter N is  $k_1 = 2.64$ ;
- 7) The slope of a line in the bottom- right side of the letter Q is  $k = 1.66$ ;
- 8) The slope of a line in the bottom- right side of the letter R is  $k = 2.52$ ;
- 9) The slope of two lines in the letter V respectively is  $k_1 = 4.8$ ,  $k_2 = 5.14$ ;
- 10) The slope of four lines in the letter W respectively is  $k_1 = -5.68$ ,  $k_2 = 16.25$ ,  $k_3 = 5.77$ ,  $k_4 = -17$ ;
- 11) The slope of two lines in the letter X respectively is  $k_1 = 2.51$ ,  $k_2 = -2.49$ ;
- 12) The slope of two lines in the letter Y respectively is  $k_1 = 1.96$ ,  $k_2 = -1.96$ ;
- 13) The slope of a line in the letter Z is  $k = 2.38$ .

#### 4 Conclusion

Extraction the number of holes in the character, extends the application of seed filling algorithm, this method is easy to implement, with high accuracy, in addition, it is an innovation to calculate the slope of the character stroke by the method of Hough transform, we obtained a relatively standard form about the slope of characters through calculation, in the incomplete character recognition, the slope has a great difference between the characters, so it has a good degree of differentiation, we achieved better results through experiment.

#### Acknowledgments

My sincere thanks should go to Liaoning province technology hall plan projects (2012217005) and Liaoning province Science of public research funds (2012004002). It is because of their sponsorship, this project can be completed so smoothly. Any progress that I have made is the result of their profound concern and selfless devotion.

## Reference

- [1] Wang Xiaoxue 2010 Application of Digital Image Processing in License Plate Recognition *Process Automation Instrumentation* **31**(7) 22-8
- [2] Han Liming 2010 Research and implementation on key technology in license plate recognition system *Computer Engineering and Design* **31**(17) 3919-23
- [3] Tao Xun 2011 License Plate Recognition System Design and Realization, *Electrical Automation* **33**(4) 77-80
- [4] Wang Yulei, Jiang, Lixing 2009 Digital Character Recognition Algorithm Based on Character Feature *Marine Charting* **29**(1) 56-8
- [5] Dong Lingjiao 2008 Character feature extraction for car plate recognition *Electrical Engineering* **25**(9) 106-8
- [6] Rui Ting 2004 License plate character recognizing under high noise using sam *Pattern recognition and artificial intelligence* **17**(2) 467-70
- [7] Ye Chenzhou 2000 Number-Plate Character Recognition *Journal of Shanghai Jiaotong University* **34** (5) 672-5
- [8] Zhang Jian 2011 Research on character recognition of license plate recognition *Information Technology* **9**(4) 109-20
- [9] Yu Lasheng, Shen 2004 Deyao A Refinement of the Scan Line Seed Fill Algorithm *Computer Engineering* **29**(10) 70-2
- [10] Wen Y Lu, Y Yan J Zhou, Z von Deneen K M, Shi P 2011 An Analog Feedback Associative Memory *IEEE transactions on intelligent transportation systems* **12**(3) 117-26
- [11] Zou Mingming, Lu Di 2010 Recognition algorithm of car license plate characters based on modified template match *Foreign Electronic Measurement Technology* **29**(1) 59-61
- [12] Zhao Kun 2010 An Improved Preprocessing Algorithm of Vehicle License Plate *Henan Sciences* **28**(3) 329-32
- [13] Wang Jianyong 2006 An Improved Algorithm for Line Detection *Computer Engineering* **32**(16) 172-4
- [14] Pasi Franti, Alexey Mednongov, Ville Kyrki, Heikki Kalviainen 2000 Content-based matching of line-drawing images using the Hough transform *International Journal on Document Analysis and Recognition* **3**(2) 117-24
- [15] Zheng L, He X, Samali B, Yang L T 2013 Accuracy enhancement for license plate recognition *Journal of Computer and System Sciences* **79**(2) 511-6
- [16] Hsu G-S, Chen J-C, Chung Y-Z 2013 Application-Oriented License Plate Recognition *IEEE Transactions on Vehicular Technology* **62**(2) 552-61
- [17] Jiao J B, Ye Q X, Huang Q M 2009 A configurable method of multi-style license plate recognition *Pattern Recognition* **42**(3) 358-69
- [18] Khader Mohammad, Sos Agaian, Hani Saleh 2011 Practical automatic Arabic license plate recognition system *SPIE Conference on Multimedia on Mobile Devices* 2011.

## Authors

**Kai Song, born in 1964, Liaozhong, China**

**Current position, grades:** professor and a PhD supervisor, Party committee secretary of the College of Information and Engineering in Shen Yang Ligong University now

**University studies:** B.S. degree in computer control theory from Xiamen University, Xiamen, China, in 1986, M.S. degree in network and communication from Northeastern University, Shenyang, China, in 1989 and Ph.D. degrees in biological environmental monitoring and control engineering from Shenyang Agricultural University, Shenyang, China, in 2008.

**Scientific interest:** computer vision, image processing and analysis.

**Publications:** about 20 academic papers about image processing in core journals in recent three years

**Caihong Niu, born in 1987, Xinxiang, China**

**Current position, grades:** M.S. degree student in Shenyang Ligong University

**University studies:** B.S. degree in Communication engineering from Henan University, Kaifeng, China, in 2011, M.S. degree in Shenyang Ligong University now

**Scientific interest:** image processing and analysis.

# Research and implementation of 3D reconstruction base on multi-contours

Kunliang Liu<sup>1</sup>, Jinming Huang<sup>2\*</sup>

<sup>1</sup>*School of Computer Science and Software Engineering, Tianjin Polytechnic University, Tianjin 300387, China*

<sup>2</sup>*Development and Research Center, China Geological Survey, Beijing 100037, China*

*Received 22 May 2014, www.tsi.lv*

---

## Abstract

3D Reconstruction often faces to a serial of 2D contour lines but not to volume data, which as we often processed, so study of 3D reconstruction based on multi-contours has important practical values. In the process of 3D-reconstruction based on multi-contours, contours correspondence, contours splicing, branch problem, and terminal contours closing are all its key technologies. In this paper we give the concrete solutions on every step of 3D construction of multi-contours. According to winding issue of contours we provide means of gauging sum of angles of contour's edges adjacent to each other, which avoided error judging of winding of contours. As to branch issue of one contour corresponding to several contours, we give the way of splitting contour based on ratio in circumference of corresponding contours. We also give the mean of maximal field angle to reduce the calculation time on triangulating terminal contours. The solution we provided can give correct result of contours splicing under any kind of contours. It proves that every step of the solution is correct and effective. The solutions we designed are more general than other solutions.

*Keywords:* contour, 3D reconstruction, Delaunay triangulate, convex hull

---

## 1 Introduction

We consider a 3D-Reconstruction problem, in which a sequence of contour lines is to be sewed up to construct a closed surface; with the surface, we can obtain three-dimensional shape of geometric objects. 3D-Reconstruction problem based on multi-contours arise in a wide area of applications, including medical data visualizing, 3D geological modelling and biological science displaying. In practical application, the data information we get often not a volume data using which we can construct object shape easily base on MC algorithm, but a serial of contour lines information. For instance, in the process of medical data visualization, if the distance of two medical slices of CT or MRI image is much greater than the image resolution, that is, the sample interval in Z direction is very sparse. In this case, we should outline contour lines to enclose the region we are interested in on every image slices and reconstruct the triangulation network model from the serial contour lines. Another example is in the field of medical image three-dimensional data visualization, though the distance between two adjacent tomographic images are very small, the computer cannot distinguish different co-exist material from one to another based on current material classification techniques. Therefore, we need medical specialist to outline the areas of interest. In the medical treatment system, In order to ensure the accuracy and reliability of the treatment we should firstly outline organs (blood vessels, nerves, etc.) contour lines, and

then reconstruct the tissue surface via these contour lines. Therefore, the tissue surface reconstructed from organs contour lines can assist medical diagnosis greatly, provide a reliable basis for the elaboration of best surgical options to improve surgical quality, reduce malpractice [1]. Again also in 3D geological model reconstruction technology, the solution should interlink the contour lines of adjacent geologic sections to generate the 3D geological model. Therefore, under the need of the practical application of the above, the study of three-dimensional reconstruction has very important practical value. In the process of 3D-reconstruction based on multi-contours, contours match, contours splicing, contours divergence handling, and terminal contours closing are all its key technologies. In [2], contours splicing are based on OBB projection transformation, authors first determine the vertices of the polygon convexity, for concave vertex, convert it to the corresponding convex hull, then calculate the Oriented bounding box of convex hull, rotate and pan the bounding box, compute the inscribed ellipse of bounding box, project proportionally every vertices of contour lines on this oval, splice the contour lines base on project vertices, tell correspondence between the vertices of adjacent contour lines; Finally, restore the actual coordinates, perform a three-dimensional reconstruction of the original model. The approach can resolve the problem of cross-stitching better than the shortest diagonal method, but under certain circumstances, this approach can also occur distortions, for the reason of not considering

---

\* *Corresponding author* e-mail: 2003\_asecliu@126.com

contour centre alignment problems and the changes of the contour size, the model sometimes will be even worse compared with the original model, and that the method did not solve the problem of splitting of contour lines too. In [3] Although author considered the centre alignment and contours size change and proposed contours branch solutions, When processing branch issues the solution required manual intervention, reducing its automatic and did not deal with the end of the contour line plugging problems which would result in empty in the end of the contour lines, so the spliced results can not to form a closed surface. Current splicing resolutions of multi-lines provided by all the paper involved focused to achieve some particular steps, resolve some particular problem. It only suit limited scope and is not universal. This paper, aiming at the reconstruction problem of contour lines, effectively resolved the cross-cutting issues in stitching contour lines, achieved the automatic processing of branch contour and closing of terminal contour.

**2 Pre-treatment contours**

As for adjacent contours lines, if the centre is aligned, the shape resembling and the twining direction is identical, Splicing results generally cannot go wrong. In fact, the contour lines for three-dimensional reconstruction have diversity shape, in order to improve the effect of reconstruction, avoid wrong connection, it is necessary to pre-process them before stitching.

**2.1 CENTER ALIGNMENT AND SIZE ADJUSTMENT OF CONTOURS**

Adjacent contour lines shape and the centre position may be largely different, so in order to correctly splice the contour lines it is necessary to convert the contour lines by panning and zooming coordinates, so that the centre is identical to each other and the size ratio coincides with each other, otherwise, it will produce cones phenomenon, as Figure 1 shows. Assuming two adjacent contour lines are Contour1 and Contour2, whose enclosing rectangle centre coordinate are center1(x1,y1) and center2(x2,y2), in which PanFactor and ZoomFactor denote the translation distance and scaling factor of the contour line, then there is formula as Equation (1) shows.

$$\begin{cases} PanFactor.x = x2 - x1 \\ PanFactor.y = y2 - y1 \\ ZoomFactor.x = width2 / width1 \\ ZoomFactor.y = height2 / height1 \end{cases} \quad (1)$$

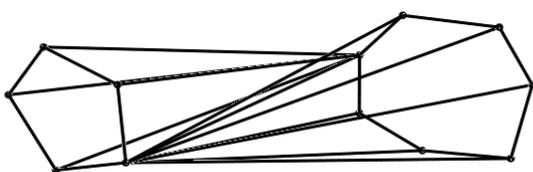


FIGURE 1 Incorrect splicing, resulting in cones

All coordinates of Contour line of Contour 1 maintains unchanged, each point coordinate of Contour 2 are adjusted as Equation (2) shows.

$$\begin{cases} Contour2.x = (contour1.x + PanFactor.x) * ZoomFactor.x \\ Contour2.y = (contour1.y + PanFactor.y) * ZoomFactor.y \end{cases} \quad (2)$$

**2.2 TWINING DIRECTION CONSISTENCY ADJUSTMENT OF CONTOUR**

For adjacent contour lines, after their position and size have been adjusted, if their twining directions are not identical, some of the contour lines direction should be reversed to make their twining direction identical. These judgments and treatment are problems often encountered in computer graphics processing, pattern recognition, CAD and other areas. In most cases, calculating the contour line normal vector if the adjacent contour lines' normal vector are inconsistent, then the twining direction of contour line are different, so there needs to reverse the twining direction of some contour lines, this can be obtained by reverse the control points' arrangement. As the contour lines shape are diverse, the contour line may be convex polygon or concave polygon, the method judging the twining direction via contour line's normal vector only suits for convex polygon, there is likely to get wrong result when contour line is concave polygon, one solution is firstly projecting concave polygon onto it's convex hull, then calculating the normal vector of it's convex hull, but this method increases the amount of computation. This paper adopts the method of judging the sum of angles to solve the twining problem. This method suits for diverse shapes of contour line, without having to calculate concave polygon's convex hull. When using this means each edge of the polygon of contour lines can be seen as a directed edge whose direction coincides with the direction of the vertex arrangement. After calculated the sum of deflection angles between current edge and the next edge of the contour line, then judged the signs of adjacent contour lines' deflection angle sum, if their sign are identical, their twinning direction coincident, otherwise inconsistent. In order to get deflection angle sum, the normal vector direction of the plane in which the contour line lies should be ensured identical, then computing respectively the dot product and cross product of the two adjacent edge of the contour line, based on dot product the angle between two adjacent edges of contour line can be ascertained, based on the cross production of the two adjacent edges of contour line and the normal vector of the plane the sign of the angle can be ascertained, if cross production vector has the same direction with the plane normal vector the sign of angle is positive otherwise negative. No matter concave polygon or convex polygon the accumulative sum of deflection angle of contour line is or near  $2\pi$  or near  $-2\pi$ . The process is as Figure 2 shows and the formula is as Equation (3) shows. In which  $M_i$  is the dot product of

two adjacent edges,  $N_i$  is the cross product of two adjacent edges,  $N_{face}$  is the plane normal vector in which the contour line lies. The value of angle (i) is positive if  $N_i$  and  $N_{face}$  have the same direction, otherwise, negative.

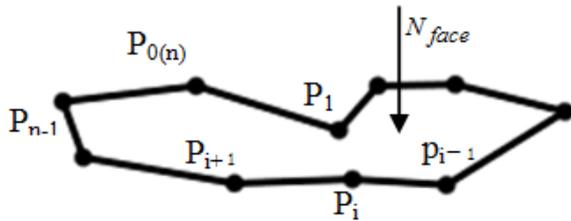


FIGURE 2 Calculate deflection angles sum

$$\left\{ \begin{array}{l} M_i = p_{i-1}p_i \bullet p_i p_{i+1} \\ N_i = p_{i-1}p_i \times p_i p_{i+1} \\ Angl(i) = \pm \arccos\left(\frac{M_i}{|p_{i-1}p_i| \bullet |p_i p_{i+1}|}\right) \\ SumAngle = \sum_{i=0}^n Angl(i) \end{array} \right. \quad (3)$$

### 2.3 COMPUTING CONVEX HULL OF CONTOUR AND PROJECTING COORDINATES TO CONVEX HULL

Convex contour can be stitched directly using splicing technique, but concave contour needs to firstly computing

$$\#define ConvexDeg(v_i, v_j, v_k) \quad (v_k.x - v_j.x) * (v_j.y - v_i.y) - (v_k.y - v_i.y) * (v_j.x - v_i.x)$$

<pre>void CalConvexHull(V, n, &amp;S) input: contourline's vertex set V, vertex number n output: contourline's convexhull vertex set S {   InitStack(S);   push(S, v_0); push(S, v_1);   k = 2;   a = v_0; b = v_1; c = v_k;   while( k != 1)   {     while(ConvexDeg(a, b, c) &lt;= 0 &amp;&amp; (Top &gt;= 1))     {       Pop(S);     }     a = S[Top]; priorpoint in V     b = S[Top];     c = S[Top]; subsequentpoint in V     while(ConvexDeg(a, b, c) &lt; 0 &amp;&amp; Top &gt;= 1)     {</pre>	<pre>V = V - v_k; k = (k+1)%n a = S[Top-1]; b = S[Top]; c = v_k; while(convexDeg(a, b, c) &gt; 0) {   V = V - v_k; k = (k+1)%n c = v_k; } Pop(S); } Push(S, v_k); } a = S[Top-1]; b = S[Top]; c = v_1; if( convexDeg(a, b, c) &lt; 0)   Pop(S); a = S[Top]; b = S[1]; c = S[2]; if( convexDeg(a, b, c) &lt; 0 &amp;&amp;</pre>	<pre>S[1] is the first point in V for( i = 1; i &lt; Top - 1; i++) {   s[i] = s[i+1]; } Pop(S); }</pre>
---	--	---

FIGURE 3 Main algorithm of converting arbitrary contour into convex one

its convex hull, and then projects its concave vertices on to the convex hull. The contour lines of connection area of ore body delineation and stratigraphic section are usually concave polygon, so it needs to disposal concave polygon to convex polygon.

Convex hull is an important tool when describing the shape and extracting the feature of object. It has been widely used in the research field of pattern recognition, image processing etc. The definition of convex hull is very simple, as for an arbitrary point set s or a polygon p, the convex hull is the minimal convex polygon that can encircle the points set s or polygon p. we design and implement the algorithm to convert arbitrary contour into convex one, the main algorithm framework is as Figure 3 shows.

After having computed the convex hull of concave contour line, we project points, which are concave points of contour onto the line segment, which is determined by the two nearest convex points of the contour line; we can obtain a convex contour from an arbitrary contour. There are two projection modes, one is according to the length proportion, and the other is vertical projection. Due to vertical projection has the result of folding line, so we select length proportion projection based on the ration, that the distance between current concave point and the previous convex point compares to the distance the previous convex point and the successive convex point on contour line. Compared to the vertical projection mode the length proportion project can effective avoids the problem of cross-stitching due to folding line.

### 3 Contour line splicing

Assuming two adjacent parallel planes each have a contour line, the upper contour line point array is  $p_0, p_1, \dots, p_{m-1}, p_m$  (in which  $p_m$  and  $p_0$  are the same point), the lower contour line point array is  $q_0, q_1, \dots, q_{n-1}, q_n$  (in which  $q_0$  and  $q_n$  are the same point), the point array is arranged in counter clockwise. Every line segment  $p_i p_{i+1}$  or  $q_i q_{i+1}$  is called line segment of contour line. We can get a Triangular facets through connecting two control points of one line segment to an adjacent contour's control point, as Figure 4 shows.

We define the line segment, which connects the point on upper contour and the point on lower contour as span. Obviously, a contour line segment and the two spans, which connect the two control points of the line segment to the adjacent contour line's control point form a triangle facet, it is called elementary triangle facet. The two spans are called left span and right span respectively. The three-dimensional shape reconstruction based on two convex contour lines is to use a series of triangular facets interconnecting the upper and the lower contour lines. But how to guarantee the connected three-dimensional shape reasonable and has a good properties are the issues that need careful study. The numerous elementary facets interconnecting the upper and the lower contour lines must compose the interconnected three-dimension surface and must not intersect each other inside the triangular facets.

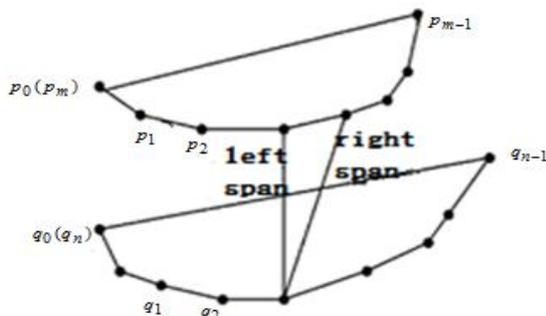


FIGURE 4 Connecting control points

Therefore, if the triangular facets are reasonable they must satisfied the two conditions as follow:

1) A contour line must and can only appear in an elementary basic triangular facet. Therefore, if the two upper and lower contour lines respectively have  $m$  and  $n$  line segment, the reasonable three-dimension surface of reconstructed shape must have  $m+n$  elementary triangular facets.

2) If the span is the left span of an elementary triangular facet, it must and only be the right span of another elementary triangular facet.

The triangular facets meet the above criteria set are acceptable body surface. For two adjacent contour lines and points array on it, the acceptable body surface meets, the above criteria can have a variety of different combinations. In so many combinations of acceptable

surface, in order to determine a combination of need, through the development of different optimization objective function, many scholars proposed different optimization methods. For example: H. Fuchs proposed the algorithm of smallest surface area; E. Kepple proposed the algorithm of maximum volume; Ehristiansen proposed the algorithm of shortest diagonal; Gannapathy proposed the algorithm of adjacent contour lines synchronously advancing [6]. Among them, the maximum volume algorithm, the shortest diagonal algorithm and the adjacent contour lines synchronously advancing algorithm belong to heuristic method; the minimum surface area algorithm and the largest volume algorithm are all global optimum surface reconstruction algorithm which need large amount of calculation and more time-consuming; The shortest diagonal algorithm and the adjacent contour lines synchronously advancing algorithm belong to a local optimum determination algorithm which need the smaller amount of calculation and can improve computing speed. The shortest diagonal algorithm suite the situation of the upper and the lower contours' size and shape are similar and centre points are relative close. The shortest diagonal algorithm chooses a shorter one of the two diagonals  $p_1 q_2$  and  $p_2 q_1$  of the quadrilateral  $p_1 p_2 q_1 q_2$  as the next triangular facet's edge to generate the triangular facet, as Figure 5 shows.

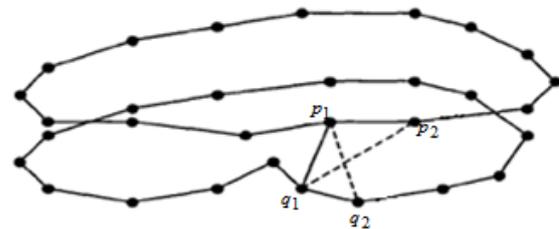


FIGURE 5 The shortest diagonal algorithm

After pre-processed the adjacent contour lines, their size and shape are relatively approximate and centre points are relative close, so they suit the shortest diagonal algorithm.

When stitching the adjacent contour lines using the shortest diagonal algorithm, we start from the first point  $p_0$  of contour line which has the less control points in the two contour lines, then find the nearest point  $q_i$  in the corresponding contour line to the point  $p_0$ , look the line segment  $p_0 q_i$  as starting side, and execute the shortest diagonal algorithm.

### 4 Branching problem

When the number of two adjacent contour lines is unequal, there needs to solve the branch problem. When solving the branching problem caused by a number of non-overlapping contour lines there is need to transfer multi-branch problem to a group of single branch problem. Contour lines merging and contour lines

splitting are two means to solve the branch problem, the way of contour lines merging is to merge the multiple contour lines into one contour line, then to splice the contour lines one corresponding one; the way of contour lines splitting is to split a single contour line into multiple contour lines via particular means, then to splice them one corresponding one.

4.1 THROUGH INTERPOLATED EDGE SOLVING BRANCH PROBLEM (WHEN BRANCH NUMBER=2)

The way of solving the branching problem via introducing interpolated edge belongs to contour lines splitting method. It is suitable for solving the problem of independent branching. The method utilizes a perpendicular bisector to split single contour line into two contour lines, then corresponds the two contour lines to adjacent two contour lines. When using this method under the circumstance of the shortest distance between the adjacent contour lines and the difference in size are relative great, the majority region of a single contour line connects to the smaller one among adjacent contour lines, thus reducing the accuracy of reconstruction, moreover, when there are too many branch, it will be very difficult to implement the method.

4.2 SOLVING BRANCH PROBLEM BASED ON RATIO OF CIRCUMFERENCE (WHEN BRANCH NUMBER ≥ 3)

Since the adjacent contour lines' distance is small and the shape of the upper and lower contour line should have a certain similarity, so we can transfer multi-branch problem into several single branch problem by using the ratio of the circumference of multiple contours. Assuming there is only one closed contour line in lower layer and there are several closed contour lines in upper layer, then we should split the single contour line in lower layer into several contour lines with the number same to the number of contour lines in upper layer according to the ratio of circumference of the upper contour lines, then we splicing the contour lines according to the followed algorithm steps. As Figure 6 shows, we assuming there are three contour lines  $C_1, C_2, C_3$  in upper layer and there is one contour line  $C_0$  in lower layer, then method of splitting contour line is as followed:

Step1: Computing barycentric coordinates of contour lines  $C_1, C_2, C_3$ , assuming to be  $B_1, B_2, B_3$ , and calculating barycentric coordinate of contour line  $C_0$ , assuming to be  $A_0$ .

Step2: Computing the barycentric coordinate of the polygon whose vertices are barycentric  $B_1, B_2, B_3$ , assuming to be  $B_0$ .

Step3: connecting the line segments  $B_0B_1, B_0B_2, B_0B_3$ .

Step4: making line segments  $A_0A_1, A_0A_2, A_0A_3$  parallel to line segments  $B_0B_1, B_0B_2, B_0B_3$  in contour line  $C_0$ .

Step5: computing the circumference of contour lines  $C_1, C_2, C_3$ , assuming to be  $P_1, P_2, P_3$ .

Step6: computing the coordinate of point  $M_1$  in contour line  $C_0$ , making  $M_1$  splitting the  $A_1, M_1, A_2$  in contour line  $C_0$  based on the circumference of contour lines  $C_1, C_2$ , that is  $A_1M_1 : M_1A_2 = P_1 : P_2$ . Using the same way, Splitting  $A_1, M_3, A_3$  and  $A_3, M_2, A_2$  each into two parts respectively based on the circumference of contour lines  $C_1, C_3$  and  $C_2, C_3$ .

Step7: matching contour lines  $C_1, C_2, C_3$  with  $M_1A_0M_3A_1M_1, M_1A_0M_2A_2M_1$  and  $M_2A_0M_3A_3M_2$ , which are split parts of contour line  $C_0$ .

Step8: Reconstructing the 3D entity using the way of 3D reconstructing of single-contours.

Note that: The split contour lines must be pre-treated also prior to splicing the contour lines.

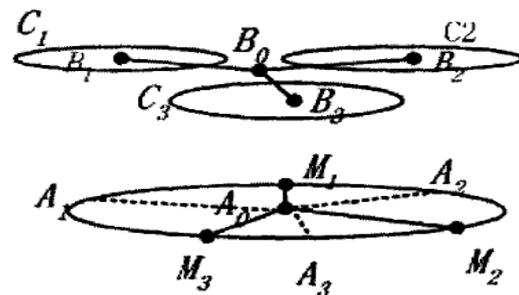


FIGURE 6 Splitting when one line corresponds to multi-lines

5 Closure processing of the end contour lines

In order to achieve the closure processing of the end contour lines, the paper designed and implemented an arbitrary polygon triangulation algorithm. In implementing the algorithm, we look on contour line as an arbitrary polygon and triangulate the contour line resorting to the triangulation algorithm of arbitrary polygon. As for details of arbitrary polygon triangulation algorithm, we can refer to the literature [7]. In order to reduce the amount of calculation, as this paper implements the algorithm we use the maximum opening angle triangle method in search of the triangle which has the minimum circumscribed circle radius, that is to say that the point which has the maximum angle to the current edge is to be the selected point, as Figure 7(a) shows. As for line segment  $AB$ , there are angles  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$  between the line segments and the other points, in which the angle of  $\alpha_4$  is maximum, and so we select point  $C$  and the line segment  $AB$  to form a triangle.

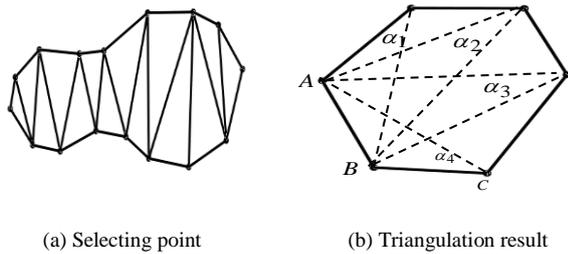
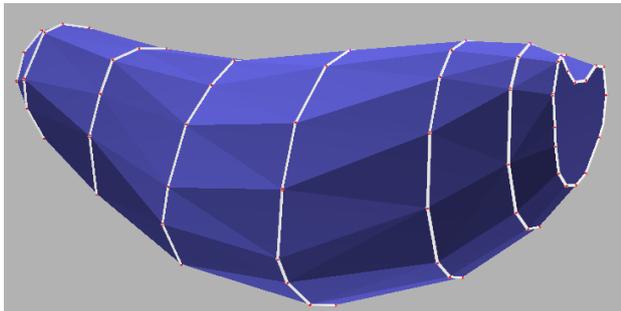


FIGURE 7 Triangulating contour line

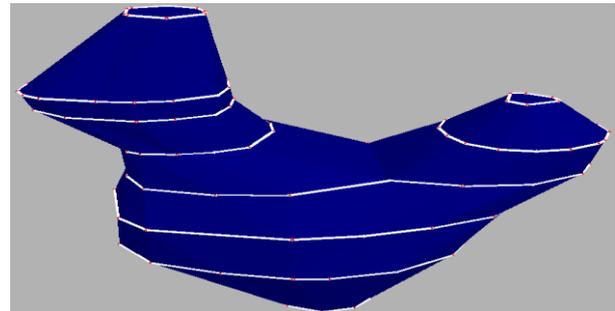
**6 Achieving result**

This paper selects VisualC++2010 and OpenGL as development tools. Using a series of contour lines data obtained by 2D inversion in the field of solid mineral resources development, after contour lines pre-processing, contour lines branch, contour lines splicing and the end contour lines closing we get the closed three-dimensional shape model, as Figure 8 shows. The stitching models of various form of contour lines are correct, avoiding the phenomena of cross-stitching. The case of one corresponding to multiple contour lines as Figure 8(b) and Figure 8(c) shows (when one contour line corresponding to more than three contour lines this method is also effective.) can automatically achieve a single contour divided by the ratio of the circumference

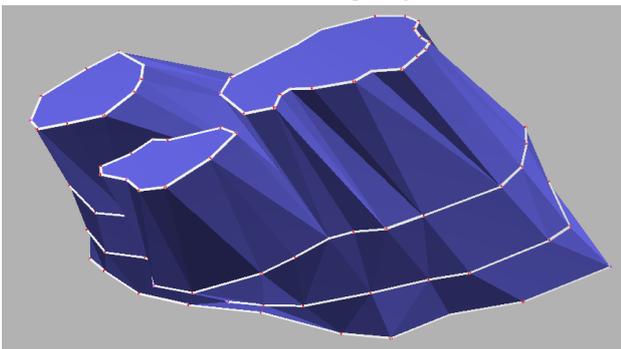
of contour lines, completely avoiding manual intervention. The segmentation results meet engineering requirements. Various form of the end contour lines are regarded as the arbitrary polygon. Using the triangulation algorithm of arbitrary polygon presented in this paper, the triangulations of end contour lines are as Figure 8(a), Figure 8(c) and Figure 8(d) showing, achieving the processing of closing the end contour lines, and there are no empties and triangulated triangles intersections. The time complexity of the algorithm is the same to the algorithm of literature [7], about  $O(N^2)$ , but when searching for triangle with the smallest circumcircle radius, using the method above can reduce the amount of computing, the reason is that obtaining the radius needs square root demanding large amount of calculation, and the maximum opening angle triangle method requires only multiplication and division demanding less amount of calculation, and when in search of candidate points under the circumstance of engineering field in which the control point coordinates absolute value is very large but relative value is very small, due to calculation accuracy problem, the minimum circumcircle radius method will triangulate in error and arise empties or triangulated triangles intersection.



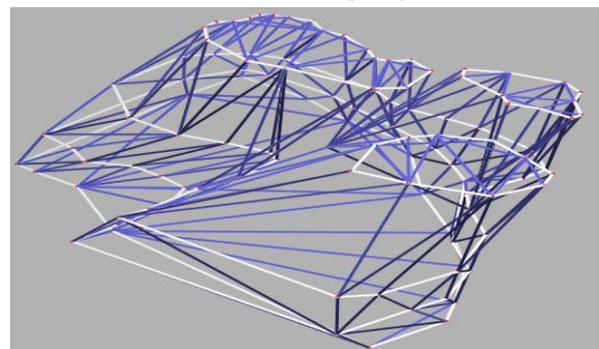
(a) No branch splicing



(b) Two branch splicing



(c) Three branch splicing



(d) Mesh line of splicing

FIGURE 8 Contour lines splicing and triangulation result

## 6 Conclusions

This paper studies the contour lines corresponding, contour lines stitching, contour lines branching, the end contour lines closing and other key technologies used in multi-contour reconstruction, and provides the specific solution of every implementation steps in three-dimension reconstruction. For contour lines winding direction problem the paper provides the method of detection sum angle of polygon avoiding the misjudgement of the contour line winding; First projects the concave polygon onto its convex hull then splices the contour lines in the process of stitching of contour lines avoiding the cross-stitching; the paper provides the method of basing on circumference ratio of contour lines to solve the branching problem; as for the end contour lines, we uses the arbitrary triangulating algorithm to implement the closure of them, when searching for minimal circumradius of triangle we uses the method of maximum opening angle triangle in triangulating reducing the amount of calculation. It proved that the

methods above are practicable, intuitive, fast, versatile compared with other contour lines stitching method. When contour lines are not parallel, we can firstly convert them to parallel or approximately parallel contour line relative to the reference point or reference plane in the pre-processing stage, then convert the coordinates to its' original coordinates after stitched the contour lines using above method to get the correct reconstruction result. In the future, we will study the three-dimension shape reconstruction under the circumstance of the contour lines not parallel and the smoothing of reconstruction result to make the reconstruction algorithm more general and the result of reconstruction more natural and beautiful.

## Acknowledgments

This paper belongs to the project from Institute of Geophysical and Geochemical Exploration. Task Book Number: Science [2008] 05-03-18, Project Number: 1212010660305

## References

- [1] Huang Yong-li 2004 Theory and technology research of 3D reconstruction based on slice-images data Master Degree Thesis 5
- [2] Li Yun-feng, Liu Xiu-guo 2011 Algorithm of contours splicing based on OBB projection transformation *Journal of Computer Applications* **31**(12) 3353-6
- [3] Mao Xian-cheng, Lu Xiao-qin, Xu Zhi-qiang 2009 Triangle tiles reconstruction from drill data-based complex contours *Journal of Computer Engineering and Applications* **45**(23) 179-81
- [4] Ma Hong-bin, Guo Jia-teng 2007 Cut-and-Sew Algorithm: a New Multi-Contour Reconstruction Algorithm *Journal of Northeastern University (Natural Science)* **28**(1) 111-4
- [5] Wu Zhong-hai, Ye Cheng-qing, Pan Yun-he 1997 An Improved algorithm of convex hull computing *Journal of Computer-Aided Design & Computer Graphics* **9**(1) 9-13
- [6] Tang Zhe-sheng 1992 Visualization of 3D spatial data sets Beijing Tsinghua university press
- [7] Tu Zhi-hong, Sang Nong 2005 Delaunay Triangulation Algorithm of Arbitrary Polygons with Visual C Language *Computer & Digital Engineering* **33**(1) 34-6
- [8] Li Mei, Mao Shan-jun, Ma Ai-nai 2006 Building Orebody Solid Model from Planar Contours, *Journal of Computer-Aided Design & Computer Graphics* **18**(7) 1017-21
- [9] Shu An, Ran Shu-hang, Wu Zhang-wen, Zhang Li, Huang Liang 2009 Three dimensional surfaces reconstruction based on shape of adjacent layer contours *Journal of Computer Applications* **29**(2) 450-2
- [10] SeungWon Shin, Damir Juric, 2002 Modeling Three-Dimensional Multiphase Flow Using a Level Contour Reconstruction Method for Front Tracking Without Connectivity *Journal of Computation Physics* **180** 427-70
- [11] Charles Soussen, Ali Mohammad-Djafari 2004 Polygonal and Polyhedral Contour Reconstruction in Computed Tomography *IEEE Transactions on Image Processing* **13**(11) 1507-23
- [12] Huang Jin-Ming, Liu Kun-Liang 2011 Design and Implementation on Volume Data Visualization System *Advance Material Research* **433-440** 5680-5

## Authors



**Liu Kunliang, born in 1976, Shandong, China**

**Current position, grades:** lecturer

**University studies:** Computer Application Technology

**Scientific interest:** Analysis and Design of Algorithms, 3-D Data visualization



**Huang Jinming, born in 1970, Jiangxi, China**

**Current position, grades:** Senior Engineer

**University studies:** Geodetection and Information Technology

**Scientific interest:** Geodetection and Information Technology, Computer Graphics

# An abnormal user behaviour detection method based on partially labelled data

**You Lu<sup>\*</sup>, Xuefeng Xi, Ze Hua, Hongjie Wu, Ni Zhang**

*School of Electronic and Information Engineering, Suzhou University of Science and Technology, Suzhou, P.R. China*

*Received 1 March 2014, www.tsi.lv*

---

## Abstract

Detecting abnormal user behaviour is of great significance for a secured network, the traditional detection method, which is based on machine learning, usually needs to accumulate a large amount of abnormal behaviour data for training from different times or even different network environments, so the data gathered is not in line with practical data and thus affects accuracy, and that increases overhead for data labelling. In light of these disadvantages, this paper proposes the detection method based on collaborate learning, it uses under-sampling method based on distance and distribution to generate training sample from imbalanced data, and semi-supervised learning method combined by ensemble classifying method to reduce demand for labelled data, it also uses differentiated member classifiers based on mixed perturbation method for collaborate training and selectively build ensemble classifier according accuracy to detect abnormal user behaviour. Experiments based on data from simulation and real network showed that this method can effectively detect abnormal behaviour and outperform traditional methods in several evaluating indicators.

*Keywords:* abnormal user behaviour detection, collaborative learning, support vector machine

---

## 1 Introduction

Abnormal user behaviour has become an increasingly serious threat to network security, behaviour such as worm, DDoS attack and botnet will burden network load, leading to dramatic drop of service quality, or even collapse of network. Therefore, accurate detection and in-time warning place an important role in network management [1, 2].

Abnormal user behaviour detection has always been a hot topic for network research. Thanks to the progress of machine learning, there are many different machine-learning methods been used in abnormal user behaviour detection. Among these methods, SVM (Support Vector Machine) [3-7] has gained more attention from researchers due to its high efficiency, stability and strong generalization ability; it can also overcome disadvantages as over-fitting, local extreme and curse of dimensionality in neural network and other methods. For example, Kim et al. proposed anomaly detection method based on SVM [4], and evaluated its performance via KDD99 data; Laskov et al. put forward one-class SVM method for intrusion detection [5], which performed well in respect of false alarm rate; Tsang et al. held up core vector machine CVM [6], which can finish fast training based on large data set; Khan et al. combined SVM and hierarchical clustering [7], which could improve the SVM method's efficiency and achieve high detection rate when dealing with large data set. Though most of current detection methods based on SVM have high efficiency, their performances are not perfect in real network environment. This is because, on one hand, the existing

methods usually need to accumulate large amounts of abnormal behaviour data as training sample from different times or even different network environment, so the data gathered is not in line with practical condition and thus affects accuracy, but if practical data gathered in targeted environment and over continuous time is used as training sample, there is a new problem: abnormal user behaviour only accounts for a small part of traffic in real environment, which will cause imbalance of training data and lead to over-fit of SVM classifier, and again affects accuracy of classified detection. On the other hand, it is very expensive to obtain the label, with increasing and changing abnormal behaviour's model; the large overhead may lead to detection methods' late response to abnormal behaviour and consequently affects the effect of detection application.

There are many specific sampling methods such as under-sampling can construct the training data from imbalance traffic. However, traditional under-sampling method based on random sampling does not consider the selected subset's effect on accuracy of SVM classifier. For problem about overhead of labelling, the semi-supervised learning method can reduce the demand for labelled data by training the classifier by part-labelled sample data, but methods based on single classifier such as Self-Training [8] has low accuracy, so researchers combine the collaborative method with semi-supervised learning, such as Co-Training [9] based on two classifiers, Tri-Training [10] based on three classifiers, CoForest [11] based on n classifiers, and so on. But in iteration process of these methods, using 10-fold cross-validation to calculate the label's confidence can generate

---

<sup>\*</sup> *Corresponding author* e-mail: luyou.china@gmail.com

large overhead. Moreover, all member classifiers are used in detection application, so some classifiers affected by noise accumulation should reduce the accuracy of detection application.

To solve the above two problems, this paper proposes an abnormal user behaviour detection method based on collaborative learning. First, in order to improve the traditional under-sampling methods, we calculate the sampling ratio based on distribution of majority class and distance between majority class's subsets and minority class in real data, thus balanced training sample is built on the premise that real data distribution is retained as much as possible, and classification accuracy is improved as well. Secondly, we combine collaborative learning method with semi-supervised learning method, trains member classifiers based on partially labelled data to reduce the need for labelled data. In the process of training, cross-validation is replaced by the integration of member classifiers' results in order to reduce the overhead. Finally, we use selective ensemble method to build the ensemble classifier according to the member classifiers' accuracy gradually calculated in the process of semi-supervised learning, and avoid the low accuracy member classifiers' affection to the effect of detection application. The experiment results based on simulation and practical network data showed that our method performs better in several evaluating indicators, compared with traditional methods.

The rest of the paper is organized as below: we present the basic concepts and abnormal user behaviour detection model in Section II. In section III, we introduce the methods of under-sampling, generation method of member classifiers, training and ensemble methods of member classifiers, and the process of detection. In section IV, we present the experiment, including the experiment environment and results analysis, and in section V, we make a conclusion and present some future works.

## 2 Model of abnormal user behaviour detection

### 2.1 RELATIVE CONCEPTS

Different user behaviour' network traffic has different statistical characteristics, which reflects the intrinsic characteristics of behaviour. The detection method based on machine learning is to train classifiers with labelled training samples, making it adapt to normal and abnormal behaviour's differences in terms of statistical characteristics, and then use them to classify real traffic in order to detect abnormal user behaviour. To better understand our detection method, we provide the following definitions:

**Definition 1:** behaviour characteristics. Factors of user's behaviour that could reflect differences between normal and abnormal behaviour and be used in statistics study, such as duration of flow, time between packet arrivals and so on. It can be represented by vector

$C_{index}=\{C_1,C_2,\dots,C_n\}$ , in which  $C_i$ ,  $i \in [1,n]$  represents No.i recognition clues.

Current research literatures of traffic classification and network security provide many behaviour characteristics sets, Moore et al. even gives a list of 246 types of behaviour characteristics [12]. But in specific situations, these characteristics are usually redundant or irrelevant, and some of them need to be removed through feature selection. In this paper, in consideration of efficiency, principal component analysis is adopted as feature selection method. In light of the length of this paper, we are not going into details.

Detection method based on machine learning needs a certain amount of labelled user behaviour data as training samples, the basic procedure is as follows: capture user traffic according to behaviour characteristics, analyse behaviour data manually or in other methods, and label the data. Since SVM is a two-category classification method, the label could be set as  $t \in \{1,-1,0\}$ , in which 1 is positive tag and means normal behaviour, and -1 is negative tag and means abnormal behaviour, and 0 means unknown type. So the definition of training sample could be concluded.

**Definition 2:** training sample. Labelled user behaviour data that could be used to train classifiers, the training sample that consists of m entries of labelled data could be shown as follow:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} & t_1 \\ x_{21} & x_{22} & \dots & x_{2n} & t_2 \\ \dots & \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} & t_m \end{bmatrix}_{m \times (n+1)}$$

Every line of this sample is made up of measured value  $x_{i,j}, i \in [1,m], j \in [1,n]$  on  $C_{index}=\{C_1,C_2,\dots,C_n\}$ , with corresponding label  $t_i, i \in [1,m], t_i \in \{1,-1,0\}$ .

### 2.2 SUPPORT VECTOR MACHINE THEORY

Support vector machine is a machine learning method put forward by Vapnil et al. [3] in the 1990s. It minimizes structure risk on the basis of statistical theory and overcomes the barrier of empirical risk minimization in traditional methods, so it has good generalization ability even with small sample. Its core theory is to replace a nonlinear mapping with a kernel function that satisfies Mercer condition, which allows sample point imported to map a high dimensional feature space, and uses linearly separable plane to obtain approximate ideal classification result. If the linearly separable training samples:  $S = \{(x_i, y_i), i = 1, 2, \dots, r\}, x_i \in R^d, y_i \in \{1, -1\}$ .

Optimal separating hyper-plane in d-dimensional space is:

$$w \cdot x + b = 0. \tag{1}$$

Then, seek optimal hyper-plane can be transformed into the problem of constrained optimization:

$$\begin{aligned} \min \varphi(\omega) &= \frac{1}{2} \|x\|^2 \\ \text{s.t. } y_i [w \cdot x_i + b] - 1 &\geq 0, i = 1, 2, \dots, r \end{aligned} \quad (2)$$

The optimal classification function obtained at last is:

$$f(x) = \text{sign} \left( \sum_{i=1}^r a_i y_i (x_i, x) + b \right) \quad (3)$$

In this function, if  $a_i$  does not equal to 0, then the sample is called support vector;  $b$  could be calculated when the support vector is selected. In linear inseparable samples, a slack variable  $\xi$  and a penalty parameter  $c$  could be added to the constraint condition in Equation (2), which turns it into:

$$\begin{aligned} \min \varphi(\omega, \xi) &= \frac{1}{2} \|x\|^2 + c \sum_{i=1}^r \xi_i \\ \text{s.t. } y_i [w \cdot x_i + b] - 1 + \xi_i &\geq 0, i = 1, 2, \dots, r, \xi_i \geq 0 \end{aligned} \quad (4)$$

In this way, the minimum risk requirements for minimum misclassified samples and maximum class interval have been compromised, and optimal classification plane in broad sense is obtained.  $C > 0$  is a constant, which controls penalty for misclassification. According to functional theory, as long as kernel function  $K(x, x')$  satisfies Mercer condition, it corresponds with some transformation space's inner product. Appropriate kernel function can transform the nonlinear separability

problem in previous space into linear separability problem in feature space, therefore, appropriate  $K(x_i, x')$  can transform nonlinear classification into linear classification without increasing computation complexity. After replacing inner product with kernel function, the classification decision-making function is:

$$f(x) = \text{sign} \left( \sum_{i=1}^r a_i y_i K(x_i, x) + b \right) \quad (5)$$

### 2.3 ABNORMAL BEHAVIOUR DETECTION MODEL

Abnormal user behaviour detection procedure can be described as follows: first, constructing training data on the basis of real traffic. Since the imbalance of abnormal behaviour data can affect classifier's accuracy, a appropriate sampling method is needed to build a balanced training data on the premise that real data distribution is maintained as much as possible; then use semi-supervised learning technology to train classifiers because this method can reduce reliance on labelled data by using more unlabelled data, collaborate learning and selective ensemble are also incorporated in semi-supervised learning process to make up both sampling method and semi-supervised learning's adverse effects on classification accuracy and overhead. The last procedure is to design detection process. So the abnormal user behaviour detection model proposed in this paper, which includes sample processing, member classifiers building, semi-supervised learning, selective ensemble and abnormal behaviour detection, is shown in Figure 1.

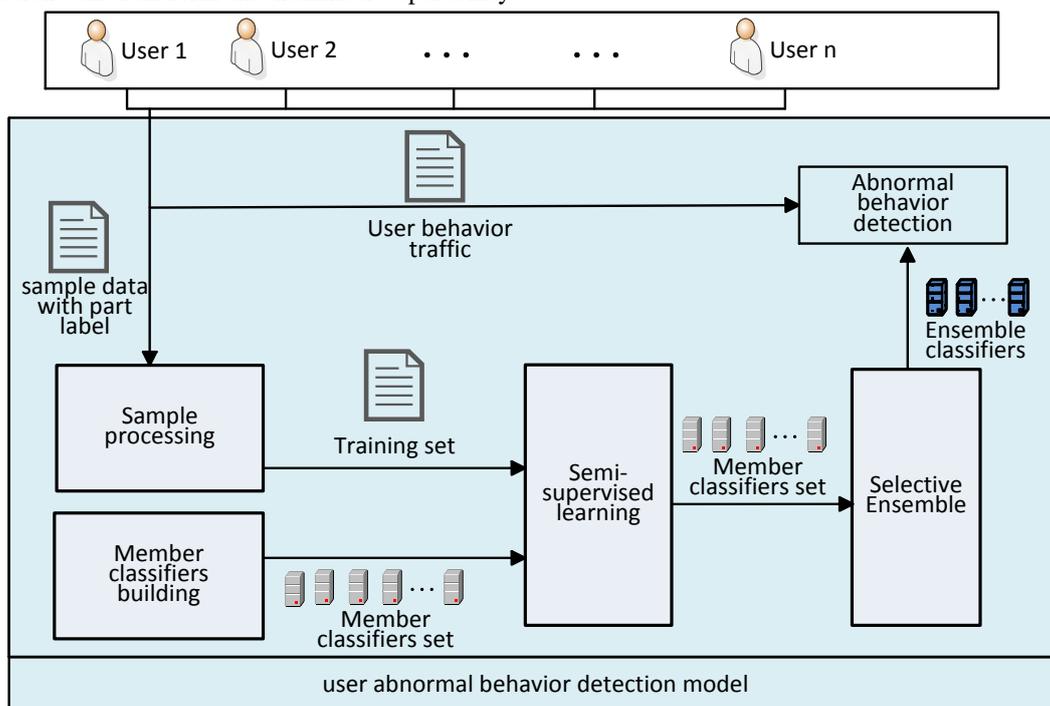


FIGURE 1 Abnormal Behaviour Detection Model

Sample processing module: using under-sampling method based on distance and distribution to process user behaviour traffic and construct training sample.

Member classifiers building module: generate a certain number of member classifiers by mixed perturbation method on the basis of feature and parameter for followed semi-supervised learning.

Semi-supervised learning module: using member classifiers' collaboration to conduct semi-supervised learning.

Selective ensemble module: select member classifiers according accuracy and integrate them into ensemble classifier.

Abnormal behaviour detection module: classify user behaviour traffic by ensemble classifier and detect abnormal behaviour.

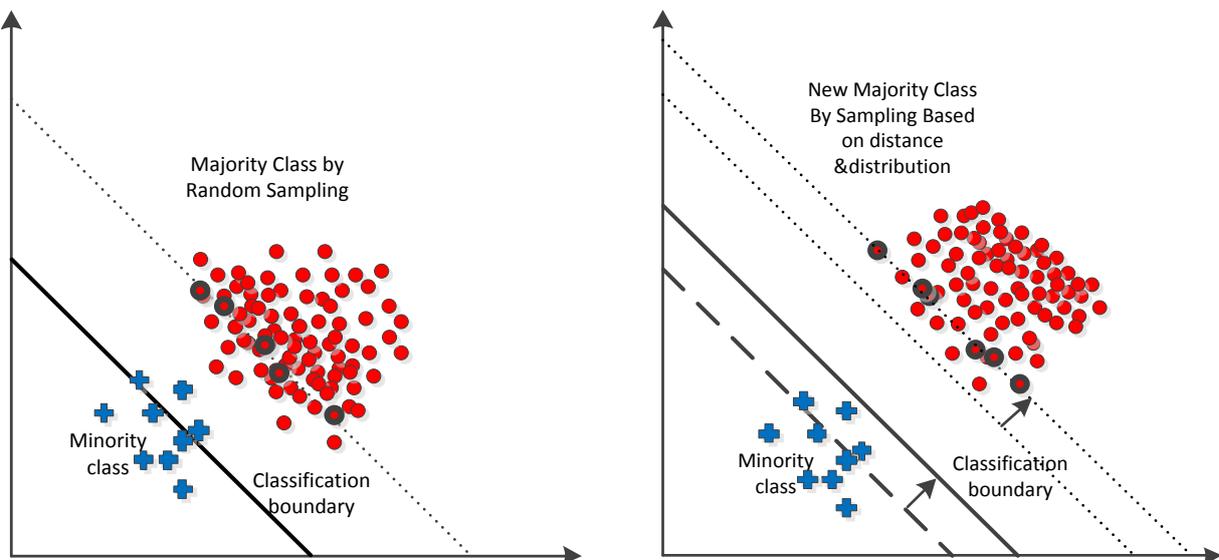
### 3 Abnormal behaviour detection method

#### 3.1 CONSTRUCTION OF TRAINING SAMPLE

The quality of training sample is of great importance to the accuracy of detection. The traditional machine learning method usually accumulates abnormal behaviour data as sample, since these data is gathered over a long period of time or even from different network environments, they may be not in line with practical condition and affect sample quality. In order to obtain high quality sample, it is better to sample and label data from the targeted network and over a continuous period of time. However, abnormal behaviour data only accounts for a small part of real traffic and training sample may be imbalanced if using uniform sampling method, which may seriously affect detection ability of balanced-data based SVM classifier. The traditional sampling methods

for imbalanced data are oversampling and under-sampling, the former is not appropriate because it requires that minority class must be a convex set, but the nature of abnormal data is unknown. The latter obtains balanced sample number by reducing the number of majority class sample, it has no requirement for the distribution properties of majority class, so our method uses under-sampling to construct training sample.

However, for the SVM based abnormal behaviour detection method proposed in this paper, traditional under-sampling methods [13], such as random under-sampling, DROP and CNN algorithm, still have deficiencies, because these methods only randomly select a subset of the majority class, and do not consider the selected subset's effect on the accuracy of SVM classifier. In fact, inappropriate subsets of majority class may lead to disappointed classification result, as shown in Figure 2(a), its subsets are too close to minority class, which causes SVM classification boundary moving to minority class and consequently reduces classification accuracy. If the distance between subsets of majority class and minority class is taken into consideration while sampling, and reduce subsets close to minority class and increase those far away, as shown in Figure 2(b), then the classification boundary can return to correct position. Besides, distance cannot be the only deciding condition of sampling ratio, distribution of majority class data also affects classification accuracy and should be considered as well, that's to say, if cluster majority class data, then majority subset data should account for a higher proportion in the sample, and minority subset should account for a lower proportion. Training sample constructed by this way can retain distribution of majority class to the largest extent, and ensure accuracy of classification.



(a) Classification boundary of random sampling

(b) Classification boundary of sampling based on distance

FIGURE 2 The effect of sampling methods to classification boundary

In light of above analysis, this paper proposed under-sampling method based on distance and distribution, the main idea is to cluster majority class (normal user behaviour data) in data to be sampled and obtain its distribution information, then calculate the distance between different subsets of majority class with minority class (abnormal user behaviour data), at last set sampling ratio based on the size of subset and its distance. The principle is that the more items the subset has, the higher sampling ratio; and the farther the subset is from minority class, the higher the ratio, thus enabling training sample to reach a compromise between retaining as many data distribution information as possible and making sampled data of majority class being as far away from minority class as possible.

Another noteworthy problem is that our method uses semi-supervised learning method (it will be introduced later) which uses partially labelled sample for training, therefore, not all data to be sampled is labelled. This makes it even more difficult to determine majority class and minority class. This paper uses semi-supervised clustering technology to deal with it, which means clustering all data to be sampled (both labelled and unlabelled data) into two categories, then study the number of data labelled as -1 (abnormal behaviour) in both subsets, the subset with more data labelled as -1 is minority class, otherwise it is majority class.

In conclusion, sampling procedure used in this paper is as follows:

**Step 1.** Sample practical traffic according to uniform proportion or equal proportion, form data set to be sampled, label part of the data manually or in other ways (for training effect, data labelled as -1 needs to be accumulated to certain threshold before stops labelling). Assume there are  $s$  entries of data to be sampled, which is

shown as follow: 
$$Source = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} & t_1 \\ x_{21} & x_{22} & \dots & x_{2n} & t_2 \\ \dots & \dots & \dots & \dots & \dots \\ x_{s1} & x_{s2} & \dots & x_{sn} & t_n \end{bmatrix},$$

$t_i \in \{1, -1, 0\}$  are labels and  $x_{i,j}, i \in [1, s], j \in [1, n]$  are user behaviour data based on  $C_{index} = \{C_1, C_2, \dots, C_n\}$ .

**Step 2.** Cluster  $Source$  into two subsets (we use Spherical K-Means algorithm), and study the number of data labelled as -1 (abnormal behaviour) in both subsets, set the subset with more data labelled -1 as majority class Mayor, the other one is Minor. Assume there are  $s_1$  entries of data in Mayor and  $s_2$  entries in Minor, and

$s_1 + s_2 = s$ , so: 
$$Major = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n'} & t_1 \\ x_{21} & x_{22} & \dots & x_{2n'} & t_2 \\ \dots & \dots & \dots & \dots & \dots \\ x_{s_1 1} & x_{s_1 2} & \dots & x_{s_1 n'} & t_{n'} \end{bmatrix},$$

$$Minor = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n'} & t_1 \\ x_{21} & x_{22} & \dots & x_{2n'} & t_2 \\ \dots & \dots & \dots & \dots & \dots \\ x_{s_2 1} & x_{s_2 2} & \dots & x_{s_2 n'} & t_{n'} \end{bmatrix}.$$

Calculate the central value of minority class:

$$\bar{x}_i = \frac{\sum_{j=1}^{s_2} x_{ij}}{s_2}.$$

$$\overline{Minor} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_{n'}), \text{ in which } \bar{x}_i = \frac{\sum_{j=1}^{s_2} x_{ij}}{s_2}.$$

**Step 3.** Cluster majority class Mayor (we use Clique algorithm) into  $K$  subsets  $A_1, A_2, \dots, A_k$ , assume there are  $Count(A_i)$  entries of data in subset  $A_i$ , calculate central value of every category  $\bar{A} = (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_{n'})$ , in which

$$\bar{a}_i = \frac{\sum_{j=1}^{Count(A_i)} x_{ij}}{Count(A_i)},$$

calculate the distance between  $A_i$  and

central value of minority class  $\overline{Minor}$  :

$$Dist_{(Minor, A_i)} = \sqrt{(\bar{x}_1 - \bar{a}_1)^2 + (\bar{x}_2 - \bar{a}_2)^2 + \dots + (\bar{x}_{n'} - \bar{a}_{n'})^2}.$$

**Step 4.** Calculate the sampling ratio of subset  $A_i$  in Mayor:

$$Ratio_{A_i} = \frac{Dist_{(Minor, A_i)}}{\sum_{j=1}^k Dist_{(Minor, A_j)}} \cdot \frac{Count(A_j)}{\sum_{j=1}^k Count(A_j)}. \tag{6}$$

According to ratio, number of sample  $A_i$  can be calculated:

$$Size(A_i) = s_0 \cdot \frac{Ratio_{A_i}}{\sum_{j=1}^k Ratio_{A_j}} + count(A_i). \tag{7}$$

In which  $count(A_i)$  is the number of labelled data in subset  $A_i$ ,  $s_0$  is the pre-set number of data item after under-sampling of majority class, and  $s_0 \approx s_2$ .

**Step 5.** Randomly sample unlabelled data in subset  $A_i$  according to the number  $Size(A_i)$ , and add all labelled data, after processing data in all subsets, combine majority class's processing result with minority class's data, thus forming the training sample  $Y$ :

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n'} & t_1 \\ x_{21} & x_{22} & \dots & x_{2n'} & t_2 \\ \dots & \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn'} & t_m \end{bmatrix}_{m \times (n'+1)}, \text{ in which}$$

$$m = s_0 + s_1, t_i \in \{1, -1, 0\}.$$

### 3.2 Member classifier training and ensemble

In order to reduce demand for labelled data in training process, semi-supervised learning method is adopted. It is a reasonable choice to use partially labelled data to train SVM classifier. The main idea of semi-supervised learning is to train classifier with labelled data in the sample and classify unlabelled data, then add classification result with high confidence to labelled data for future iterative learning, thus using the “knowledge” obtained from unlabelled data to further strengthen classifiers. Traditional collaboration based semi-supervised learning (such as two classifiers based Co-Training [9] and three classifiers based Tri-Training [10]) still face problems like noise accumulation and computational overhead. Therefore, some researchers combined ensemble classification with collaborative learning, such as n classifiers based Co-Forest [11] method. It uses member classifiers’ ensemble classification result as confidence to reduce overhead, but in process of the iteration, for every member classifier  $F_i(i \in [1, n])$ , all the other classifiers’  $F_j(j \in [1, n] \text{ and } j \neq i)$  ensemble classification results need to be calculated and determined whether the result satisfies their own condition of convergence, which consequently generates a large overhead. In view of this, this paper further improves this method by calculating confidence on the basis of all member classifiers’ ensemble classification results in every iteration process, then updates all member classifiers’ labelled data and calculates overall condition of convergence to reduce overhead. However, this method cannot ensure optimization of every member classifier, so selective ensemble method is introduced, we increase the number of member classifiers (n) in semi-supervised learning. When constructing ensemble classifier at last, select member classifiers according accuracy, and exclude member classifiers that fail in fully optimized. Since research shows that when member classifiers reach optimal performance, there is an upper limit [14] for the number of member classifiers needed (20-30), so ensemble classifier based on accuracy can still assure accuracy.

#### 3.2.1 Member classifiers construction based on mixed perturbation

Since the nature of selective ensemble is still ensemble learning, which integrate the classification results of member classifiers to determine final classification, and obtain better performance than single classifier. Schapire et al. proved that the necessary and sufficient condition for ensemble classifier’s higher accuracy than any other member classifier is that all member classifiers have higher accuracy and differences [14]. Since selecting member classifiers on the basis of accuracy only guarantees their accuracy, this paper, according to characteristics of SVM classifiers, designs a constructing

method for member classifiers that ensures their differences.

It has already proved that SVM classifiers are characteristic-sensitive and parameter-sensitive [15]. Characteristic-sensitive means that different training sample according different subsets selected from feature space corresponds can generate different classifiers, and parameter-sensitive means that Gaussian kernel based SVM’s classification ability is closely related to its parameter ( $\zeta$  and penalty parameter C), and there is a “low discrepancy area” in parameter  $\zeta$  and C’s figure region, giving them the feature of low discrepancy in member classifiers within this area, the “low discrepancy area” is called  $Reg_{Low}$ . On the basis of above conclusion, this paper proposes the construction method of member classifiers based on mixed perturbation: first use feature perturbation technology to select different subsets (the number is u) from the user behaviour characteristics set  $C_{index}=\{C_1, C_2, \dots, C_n\}$ ; then with the help of parameter perturbation technology, randomly select v parameter  $\zeta$  and w parameter C for Gaussian kernel within the  $Reg_{Low}$  region; at last combine them together, which can generate different member classifiers (the number of classifier is  $u*v*w$ ), the detail of method is:

---

#### Algorithm 1: member classifier generation based on mixed perturbation

---

**Input:** training sample Y, behaviour characteristic set  $C_{index}=\{C_1, C_2, \dots, C_n\}$ , u (number of characteristic subspace), v(number of parameter  $\zeta$ ), w (number of parameter C)

**Output:** member classifier set  $F_{all}=\{f_1, f_2, \dots, f_{u*v*w}\}$  and characteristic subspace set  $C_{all}=\{C(f_1), C(f_2), \dots, C(f_{u*v*w})\}$

---

- 1: **For** i = 1 to u
- 2: Randomly select m=n/2 characteristics entry from  $C_{index}$ , form characteristic subspace  $C_{index}(i)=\{C'_1, C'_2, \dots, C'_m\}$ , then build new m-dimensional sample  $Y_i$  from sample Y according the characteristics in set  $C_{index}(i)$ .
- 3: Analyse sample  $Y_i$ , calculate its  $Reg_{Low}$  by the method in literature [20], select v parameters  $\zeta$  and w parameters C
- 4: **For** j = 1 to v
- 5: **For** k = 1 to w
- 6: Use parameter  $\zeta_j$  and  $C_k$  to generate member classifier  $f_{(i-1)*u*v+(j-1)*v+k}$ , add it to  $F_{all}$  and add  $C_{index}(i)$  to  $C_{all}$  as  $C(f_{(i-1)*u*v+(j-1)*v+k})$
- 7: **Return**  $F_{all}$  and  $C_{all}$ .

---

#### 3.2.2 Algorithm of collaboration-based semi-supervised training and selective ensemble

The basic procedure of the collaboration-based semi-supervised training and selective ensemble is:

- i) after using labelled data in training data to train all member classifiers, use these member classifiers to classify unlabelled data in training data;
- ii) integrate classification results, calculate the confidence of data’s label, the value is the ratio of the number of member classifiers supporting this label to the total number of classifiers;

iii) select data with highest confidence (set the number as h) from those with confidence higher than threshold (0.5 in this case), and add them to training sample;

iv) iterate above steps until it reaches the maximum iteration number or can no longer update training data. Since confidence of classification results obtained in the iteration process can also reflect accuracy of different member classifiers, so classifiers' accuracy is also updated based on results integration during iteration. When training is completed, a certain number of member classifiers with highest accuracy can be directly selected to construct ensemble classifier, which will be used to detect abnormal behaviour. Algorithm detail is as follow:

Algorithm 2: collaboration-based semi-supervised training and selective ensemble	
<b>Input:</b>	$F_{all}=\{f_1, f_2, \dots, f_{u^*v^*w}\}, C_{all}=\{C(f_1), C(f_2), \dots, C(f_{u^*v^*w})\}, Y,$ iteration number <b>Max</b> , ensemble classifier number <b>z</b> , number of renewed data <b>h</b>
<b>Output:</b>	ensemble classifier $F_{resemble}=\{f_1, f_2, \dots, f_z\}$
<b>1</b>	<b>For every</b> member classifier $f_i \in F_{all}$ , build new m-dimensional sample $Y_i$ from sample $Y$ according the characteristics in set $C(f_i)$ , and set $f_i$ 's accuracy $Correct(f_i)=0$
<b>2</b>	Use labelled data in $Y_i$ to train member classifier $f_i$
<b>3</b>	Use all member classifiers to classify unlabelled data in sample $Y$
<b>4</b>	Integrate classification results of unlabelled data by bagging method, and calculate its confidence with $Degree = \frac{Agree}{u * v * w}$ in which <i>Agree</i> is the number of member classifiers that support the label 1
<b>5</b>	Select classification results whose confidence exceed 0.5 and form renewed set <i>Result</i> , arrange items of <i>Result</i> in descending order based on confidence
<b>6</b>	<b>If</b> ( <i>Result</i> = $\phi$ ) or (iteration number > <b>Max</b> ) go to 8
<b>7</b>	<b>Else</b> use top h items in <i>Result</i> to update all member classifiers' sample $Y_i$ check every item of updated data, if a member classifier labels this item correctly, add the classifier's accuracy 1 go to 2.
<b>8</b>	Select top z member classifiers according accuracy and form ensemble classifier $F_{resemble}=\{f_1, f_2, \dots, f_z\}$ .
<b>9</b>	<b>Return</b> $F_{resemble}$

### 3.2.3 Abnormal behaviour detecting procedure

The procedure of using ensemble classifier to detect abnormal user behaviour is:

- i) capture user traffic;
- ii) classify the data with ensemble classifier;
- iii) using the bagging method to vote for the classification results;
- iv) determine whether the user behaviour is abnormal or not on the basis of simple majority rule (for

convenience of judgment, set z, the number of member classifiers in ensemble classifier, as singular), detailed procedures can be described as:

**Step1.** Measurement: measure user behaviour traffic according to behavioural characteristics, and obtain data vector to be detected  $D=\{d_1, d_2, \dots, d_n\}$

**Step2.** Classification: input data vector D into member classifiers (z) to classify it.

**Step3.** Voting: vote to the data's label out coming from every member classifier.

**Step4.** Judgment: on the basis of simple majority rule, if output is labelled 1, then it represents normal behaviour; if -1, then it represents abnormal behaviour.

Since user behaviour traffics constantly, the detection process is in iteration, as shown in Figure 3.

## 4 Experiment and analysis

### 4.1 EXPERIMENT INTRODUCTION

This paper uses data from simulation and real network environment to verify the detection method. Simulation experiment uses 10% subset of KDD99 data set, which is adopted by many researchers as the benchmark of abnormal user behaviour detection. Real network data come from the computer room network in Suzhou University of Science and Technology, the network topology is shown in Figure4.

a) Simulation experiment.

Since 10% subset of KDD99 data set includes 97278 entries of normal behaviour data and 396743 entries of abnormal behaviour data, which is not in line with them balance of abnormal behaviour data in real network, so our experiment sampled KDD99 training data set to form the imbalanced data to be sampled, and set some of the data's label as empty. At last, construct the test data set from KDD99 by the same way. The detailed condition of the simulation experiment is shown as Table 1.

TABLE 1 Condition of simulation experiment

data set	sum	abnormal data	labelled data
set 1	2000	1%	30%
set 2	2000	5%	30%
set 3	2000	10%	30%
set 4	2000	20%	30%
set 5	2000	30%	30%
set 6	2000	40%	30%
set 7	2000	20%	5%
set 8	2000	20%	10%
set 9	2000	20%	15%
set 10	2000	20%	20%
set 11	2000	20%	25%
set 12	2000	20%	30%

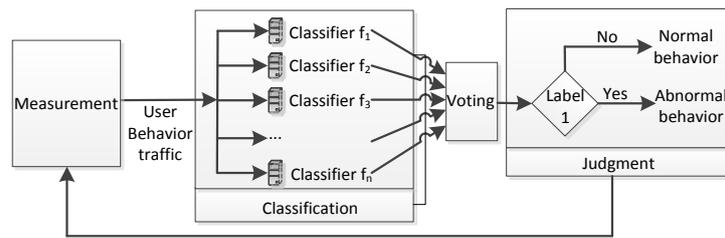


FIGURE 3 The process of detection

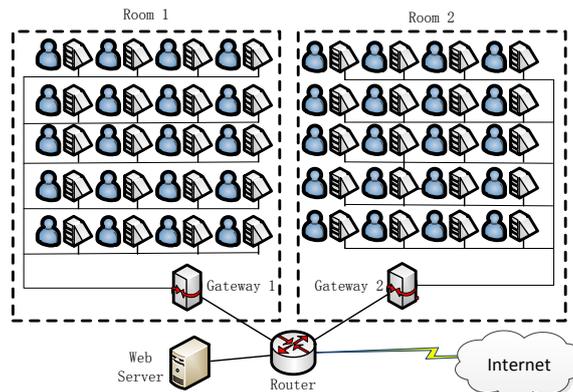


FIGURE 4 Topology of real-environment network

b) Experiment in real network environment

According to the University’s course arrangement, we collect the network traffic when there were students doing network attack trials in network security course. There were 6 students conducting DDoS attack to the server in Classroom2, while 20 students in Classroom 1 doing normal activities (including Web search, VOD and P2P download). It took 45 minutes for traffic data collection (one class). Among traffic data collected, according to IP address, student behaviour in room 2 was defined as abnormal behaviour, while data collected from room 1 was defined as normal behaviour data. Data analysis found that abnormal behaviour data only accounted for 20% of total traffic, so the imbalance feature was satisfied. Behaviour characteristics used in collection is based on the KDD99’s setting(use the characteristics that can be collected in our network environment), and construct over 3000 entries of behaviour data, select 1000 as data to be sampled and 2000 as test set. The detail is shown in Table 2.

TABLE 2 Condition of Real-environment experiment

data set	sum	labelled data
set 1	1000	5%
set 2	1000	10%
set 3	1000	15%
set 4	1000	20%
set 5	1000	25%
set 5	1000	30%

c) Contrast method.

Two contrast methods were adopted: the detection method based on single SVM classifier, and detection method based on Naive Bayesian classifier. The SVM classifier uses svm lib’s tool to optimize parameters.

d) Evaluating indicator

Precision, Recall and F-Measure were used as evaluating indicators. Precision and Recall reflect detection method’s ability to classify abnormal behaviour, and F-Measure was the harmonic mean of Recall and Precision, which could better evaluate detection method in a comprehensive way, therefore, these three indicators were widely used by researchers.

e) Hardware and software platform

The software is behaviour detection application developed by ourselves integrated with tools as svm lib, Weka, and so on, the database is SQL Server 2005, the hardware platform is Intel Core2 Quad 2.3GHz, 4GB memory, and the OS is Windows XP SP3.

4.2 ANALYSIS OF EXPERIMENT RESULT

In simulation experiment shown in Figure 5, our method is much more stable and performs better than the contrast methods. Specifically, analyse the results of different proportion of abnormal data in training data, i.e. the different imbalance degree between abnormal behaviour data and total traffic.(meanwhile the proportion of labelled data is fixed at 30%), in Figure 5(a) (Precision), Figure 5(b) (Recall), Figure 5 (c) (F-Measure) shows the comparison results with abnormal data proportion at 1%(use data set 1 in Table 1), 5%( set 2 in Table 1), 10%( set 3 in Table 1), 15%( set 4 in Table 1), 20%( set 5 in Table 1), and 30%( set 6 in Table 1). As we can see from these results, contrast method 2 performs better than contrast method 1 if there is less abnormal data, because SVM is based on balanced data. As the proportion of abnormal data rises, contrast method 1’s performance gradually gets close to contrast method 2 or even

outperforms it. But our method performs better than both of contrast methods in various situations, because of its advantages coming from the collaborative learning, ensemble classification, and special treatment to imbalance data as well. Then analyse the results of different proportion of labelled data in training data (meanwhile the proportion of abnormal data is fixed at 20%), in Figure 5(d) (Precision), Figure 5(e) (Recall), Figure 5(f) (F-Measure) shows the comparison results with labelled data proportion at 5%(use data set 7 in Table 1), 10%(set 8 in Table 1), 15%( set 9 in Table 1), 20%( set 10 in Table 1), 25%(set 11 in Table 1), and 30%( set 12 in Table 1). These results showed that when

there is less labelled data in the total data to be sampled, contrast method 1 is better than contrast method 2, because SVM has better generalization ability than Naive Bayesian method, but with labelled data increases, performance improvement of contrast method 2 is faster than that of contrast method 1, while our method performs better than both of the contrast methods and is more stable, because when there is less labelled data, our method can rely on collaboration-based semi-supervised learning technology, and when the number of labelled data increases, it can maintain stable due to the advantage brought by ensemble classification.

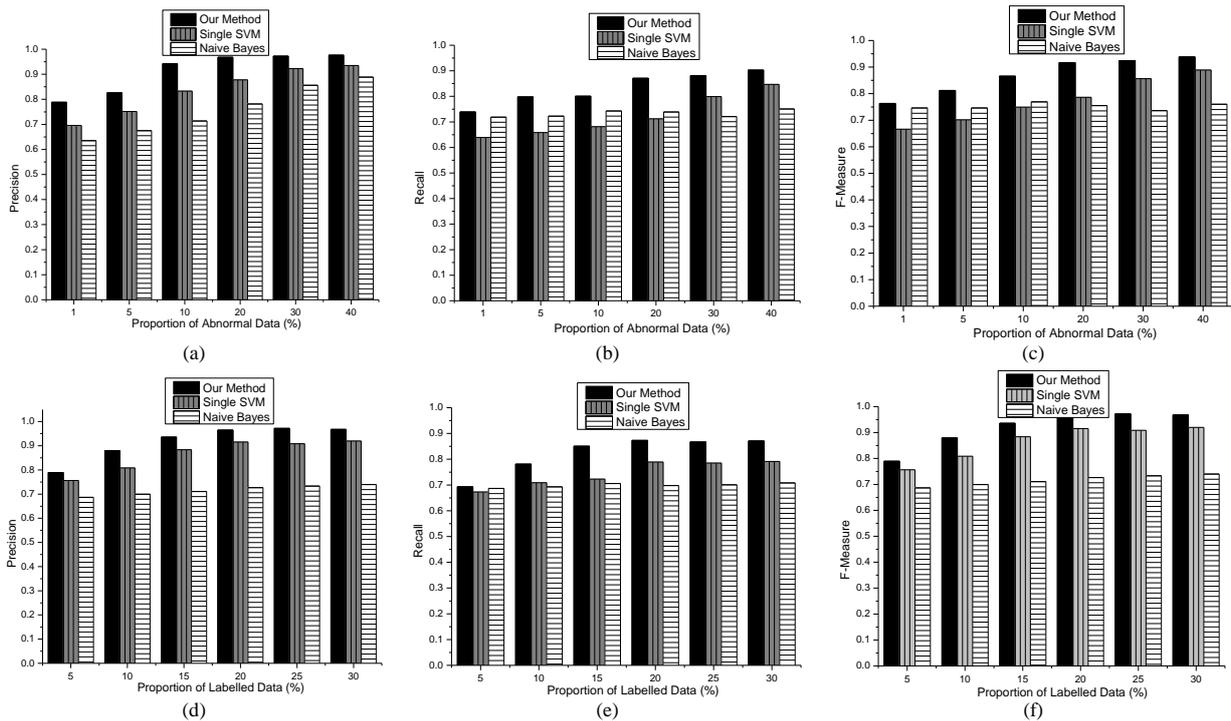


FIGURE 5 Results of Simulation Experiment

The result of real-environment experiment is shown in Figure 6. As showed in Figure 6, compared with simulation experiment result, our method maintains its advantages in all indicators. Since the proportion of abnormal behavioural data is fixed in real-time environment (about 20%), we analyse the results of different proportion of labelled data in training sample, in Figure 6(a) (Precision), Figure 6(b) (Recall), Figure

6(c) (F-Measure) showed the comparison results with labelled data proportion at 5% (use data set 1 in Table 2), 10% (set 2 in Table 2), 15% (set 3 in Table 2), 20% (set 4 in Table 2), 25% (set 5 in Table 2), and 30% (set 6 in Table 2). Results shows, thanks to the collaboration-based semi-supervised training and selective ensemble technology, our method is better than contrast methods when there is less labelled data.

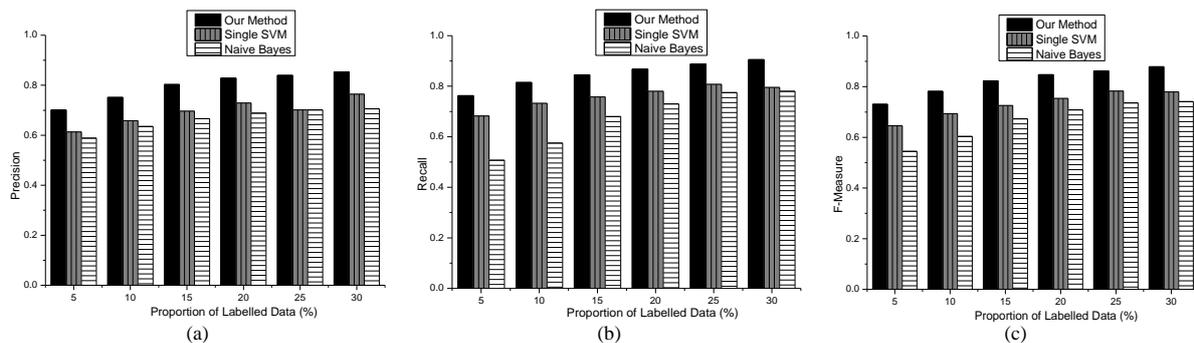


FIGURE 6 Results of Real-Environment Experiment

## 5 Conclusion

Traditional machine learning based abnormal user behaviour detection method need accumulating a large amount of abnormal behaviour data from different period of time or even different network environment, the data gathered is not in line with practical condition, and that increases overhead for data labelling, so they cannot detect abnormal user behaviour quickly or accurately. This paper proposes an abnormal user behaviour detection method based on collaborative learning to improve the traditional methods. It uses distance and distribution based under-sampling method to construct training sample from imbalance real data gathered in targeted environment over continuous time, trains different member classifiers by collaborative learning

## References

- [1] Luo J, Han Z, Wang L 2009 Trustworthy and controllable network architecture and protocol framework *Chinese Journal of Computers* 3(3) 391-404 (*In Chinese*)
- [2] Lin C, Lei L 2007 Research on Next Generation Internet Architecture *Chinese Journal of Computers* 30(5) 693-711
- [3] Vapnik V 1998 Statistical Learning Theory *Wiley New York*
- [4] Deepaa A J, Kavitha V 2012 A Comprehensive Survey on Approaches to Intrusion Detection System *Procedia Engineering* 38 2063-9
- [5] Davisa J J, Clark A J 2011 Data preprocessing for anomaly based network intrusion detection A review *Computers & Security* 30 353-75
- [6] Tsang W, Kwok J T, Cheung P M 2005 Core Vector Machine: fast SVM training on very large datasets *Journal of Machine Learning Research* 6 363-92
- [7] Khan L, Award M, Thuraisingham B 2007 A new intrusion detection system using support vector machines and hierarchical clustering *VLDB Journal* 16(4) 507-21
- [8] Hady M F A, Schwenker F 2013 *Semi-supervised Learning Handbook on Neural Information Processing* Springer Berlin Heidelberg 215-39

method on partially labelled data to reduce the overhead and labelled data, and constructs ensemble classifier based on accuracy to detect abnormal user behaviour accurately. The experiment results show that our method performs better in several evaluating indicators than traditional methods. Our next work includes optimize detection efficiency, and study user behaviour control mechanisms on the basis of abnormal behaviour detection result.

## Acknowledgments

This work is supported by Jiangsu Provincial Natural Science Foundation of China under Grants No. BK20131154.

- [9] Teichman A, Thrun S 2012 *The International Journal of Robotics Research* 31(7) 804-18
- [10] Zhou Z-H, Li M 2005 Tri-Training Exploiting Unlabelled Data Using Three Classifiers *IEEE Transactions on Knowledge and Data Engineering* 17(11) 1529-41
- [11] Li Ming, Zhou Z-H 2007 Improve Computer-Aided Diagnosis With Machine Learning Techniques Using Undiagnosed Samples *IEEE Transactions on System* 19(11) 1479-93
- [12] Moore A W, Zuev D 2004 Discriminators for use in Flow-based classification *Technical Report IRC-TR-04-028 Intel Research Cambridge*
- [13] Bastista G E, Prati R C, Monard M C 2004 A study of the Behaviour of Several Methods for Balancing Machine Learning Training Data *ACM SIGKDD Exploration News letter* 6(1) 20-9
- [14] Schapire R E, Freund Y, Bartlet P 1998 Boosting the classification boundary a new explanation for the effectiveness of voting methods *The Annals of Statistics* 26(5) 1651-86
- [15] Valentini G, Dietterich T 2004 Bias variance analysis of support vector machines for the development of SVM based ensemble methods *Journal of Machine Learning Research* 5(6) 725-75

Authors	
	<p><b>You Lu, born in July, 1977, Suzhou, China</b></p> <p><b>Current position, grades:</b> Ph.D. candidate, lecturer.  <b>University studies:</b> School of Electronic and Information Engineering Suzhou University of Science and Technology, School of Computer Science and Engineering, Southeast University, Nanjing.  <b>Scientific interest:</b> next generation network architecture and user behaviour control.  <b>Publications:</b> 5 papers in Chinese top journals.</p>
	<p><b>Xuefeng Xi, born in February, 1978, Suzhou, China</b></p> <p><b>Current position, grades:</b> Ph.D. candidate, Associate Professor.  <b>University studies:</b> School of Electronic and Information Engineering Suzhou University of Science and Technology, School of Computer Science and Engineering, Southeast University.  <b>Scientific interest:</b> network application, natural language processing, information extraction, parallel and distributed computing.  <b>Publications:</b> 3 papers in Chinese top journals.</p>
	<p><b>Ze Hua, born in May, 1972, Suzhou, China</b></p> <p><b>Current position, grades:</b> Ph.D. candidate, Associate Professor  <b>University studies:</b> School of Electronic and Information Engineering Suzhou University of Science and Technology.  <b>Scientific interest:</b> network application, parallel and distributed computing.  <b>Publications:</b> 3 papers in Chinese top journals</p>
	<p><b>Hongjie Wu, born in June, 1977, Suzhou, China</b></p> <p><b>Current position, grades:</b> PH. D. Associate Professor.  <b>University studies:</b> School of Electronic and Information Engineering Suzhou University of Science and Technology, School of Computer Science and Technology, Soochow University.  <b>Scientific interest:</b> network application, parallel and distributed computing.  <b>Publications:</b> 8 papers in International and Chinese top journals.</p>
	<p><b>Ni Zhang, born in August, 1976, Shanxi, China</b></p> <p><b>Current position, grades:</b> Ph.D. candidate, lecturer.  <b>University studies:</b> School of Electronic and Information Engineering Suzhou University of Science and Technology, School of Computer Science and Engineering, Southeast University, Nanjing.  <b>Scientific interest:</b> network application, natural language processing, information extraction, parallel and distributed computing.  <b>Publications:</b> 2 papers in Chinese top journals.</p>

# Landscape based on three-dimensional SketchUp modelling to get visualization applications

**Xiaoxiong Wang\***

*School of Architecture, Chang'an University, Xi'an, Shaanxi, 710075, China*

*Received 1 May 2014, www.tsi.lv*

---

## Abstract

Landscape designing CAD software are now mostly in two-dimensional draft stage in domestic, and three-dimensional visualization of designed landscape models and applications are still in the initial stage. The essay based on precise three dimensional solid modelling SketchUp software, introduces SketchUp software functions and features, and focuses on three aspects of garden design, hand-painted TIN terrain design, planting design and planning and design of garden objects. Through engineering examples to import the landscape of two-dimension vector data into SketchUp software, use SketchUp techniques flexibly to solve different details and avoid redundant and diverse work. After rendering and post processing, establishing a realistic visualization of three-dimensional entity model.

*Keywords:* SketchUp software, landscape, three dimensional solid model, engineering examples

---

## 1 Introduction

With the development of science and technology, concerning for geographic information are increased, the requirement for access to information is getting higher and higher. Due to the three dimensional environment closer to people's vision habitat, more geographical information can be shown, it can express more complex spatial structures. Establishment of three-dimensional geologic models in two ways: use a geographical model of three dimensional to display topography; use three-dimensional model animation techniques [1] to display formation and process of the development of models.

Development of three-dimensional modelling technology, three dimensional landscape modelling technique bases on SketchUp software has gotten more and more attention by Garden Lovers [2]. SketchUp is a collection of design tools that Google introduced in 2006 for design-oriented process, operator interface is simple and powerful this gives many landscape designers and enthusiasts even more ideas of experience, breaking the shackles of traditional two-dimensional landscape design ideas, and create different styles of three dimensional model programmes. Establishment of three-dimension model in SketchUp software is as simple as a pencil on paper drawing, the software will automatically identify the composition of line and capture, you can quickly import or export to dwg, dxf, jpg and other formats files [3]. Due to landscape, design has a lot of detail in textures, you could find texture mapping function in the SketchUp software, use material of the software itself, or you can import your own production material, to create the landscape model more realistic.

Landscape designing is the process of combining aesthetics and technology, science and the environment. Designing works not only need to have visual aesthetic, and get more interaction with the surrounding environment. Modern landscape designing including urban greening, plaza designing, road conditions, and river attractions planning, landscaping, etc. [4]. It emphasizes both the landscape and life, geography, cultural integration, but also the vegetation development, conservation and the using of natural resources [5]. However, designing and application of computer-aided landscape was introduced relatively late, many landscape design in secondary developed of Auto CAD. This tool has been unable to meet current requirements and ideas of the designers. So a variety of model building software in 3D have been appeared much, and are applied to all works of life. This article focuses on modelling methods of the SketchUp software, and shows the embodiment of engineering to display the results.

## 2 SketchUp software overview

SketchUp has a unique nature sketches and interaction designing thinking, making designers have more inspiration in the designing process, powerful function bring designers illuminating harvest [6].

### 2.1 INTRODUCTION OF SKETCHUP FUNCTION

According to the current mainstream trend of multimedia, three dimensional images and text content is favoured by learners. Three dimensional multimedia can present plane content more specific and detailed for learners to

---

\* *Corresponding author* e-mail:wxxbobo@163.com

understand an object more intuitively in different perspectives and aspects. Three-dimensional character presented by media factors break through previous content presentation form. Learners can have different mental feeling when they receive the content presented by multimedia courseware.

TABLE 1 SketchUp function introduction

Function modules introduction	Tools in module
Essential tools modules	Selection tool, eraser tool, painting tool
Comment module	Measuring tool, text tool, dimension tool
Drawing module	Circle tool, line tool, polygon tool, free drawing tool
Camera module	View, track, zoom tool
Correction module	Mobile tool, follow tool, scale tool
Sandbox modules	TIN is used to simulate terrain
Model management module	Materials browser, parameter settings and parameter selection
Google module	Share tools on Google Earth
Navigation module	Touring tool, walking tool, camera setting tool
Import/export module	Supports various data formats for import and export

2.2 SKETCHUP SOFTWARE CHARACTERISTICS

- 1) Feature of SketchUp software is his interface is simple, powerful, and easy to learn and use.
- 2) SketchUp provides new methods for computer-aided architectural designing, meet the requirements of each stage of the building designing, show better details, and get three dimensional works more realistic.
- 3) SketchUp provides a powerful rendering tool and several models displaying mode, you can set the daylight effects rendering and shadow analysis at different times.

- 4) SketchUp software has a unique geometry creation and editing techniques, and full of randomness.
- 5) SketchUp through creating groups and components to achieve three-dimensional modelling, using it widely and flexibly will not only improve efficiency, also can display optionally implement an object hidden.
- 6) SketchUp software entities that defines rich parameter settings can be made, each entity has a intelligent autonomous relationship with each other.
- 7) SketchUp software has more real actions, shadow analysis, special effects presentation can be done more quickly and efficiently.
- 8) SketchUp software provides the secondary development interface, with extended capabilities [7, 8].

3 Modelling of three dimensional designing

Three-dimensional SketchUp landscape designing, take landscape design as a subject, terrain design, planning design, planting design for the main content, integrate elements such as buildings, roads, culture, whole architecture is shown in Figure 1.

3.1 DUAL CHANNELS ASSUMPTION

Terrain designing of three dimensional landscape designing software SketchUp modules including topography, terrain analysis and geomorphology transform these basic steps. First use graphic data to simulate the original topography to get terrain mesh. Analyse according to software modules such as terrain, water, and slope.

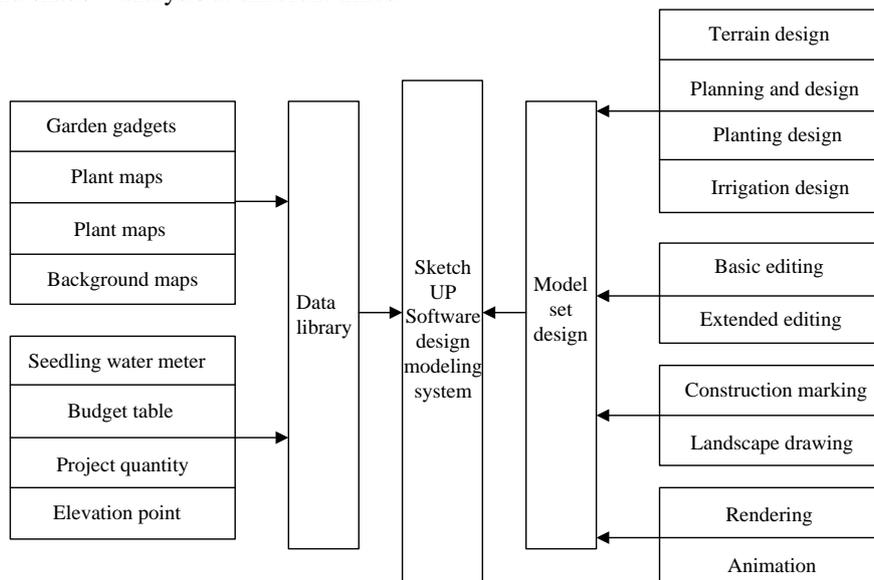


FIGURE 1 Design modelling system architecture

In the sandbox modules, we use TIN to make terrain models. TIN is the triangular plane which linking to each other, after smoothing, it looks like a continuous smooth surface. The orientation of the triangle in TIN can be different, some triangles are level headed, and some

triangles are orientated vertically, the orientation of the triangle is known as triangulation. Triangulation [9] is very important, because with some tools of the Sandbox, you can change the orientation of the triangle, so as to make the TIN more smooth as shown in Figure 2.

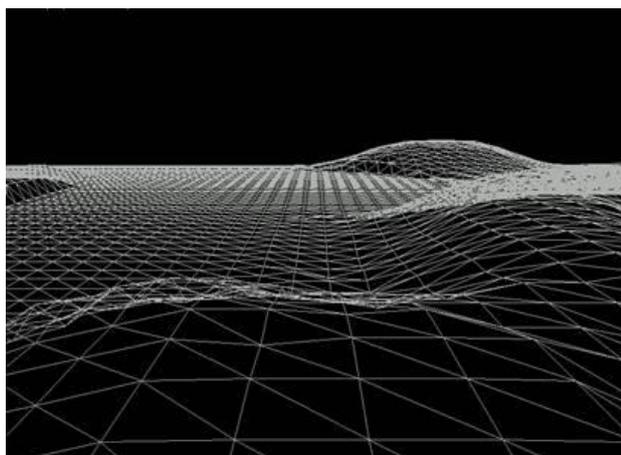


FIGURE 2 TIN making terrain model

### 3.2 PLANTING DESIGNING

In landscape construction process, usage of the plant is usually very large. There are various planting patterns in the landscape model, include solitary planting, stroll planting, turf planting, unitizing and intermingle. In SketchUp software, you can edit the result of planting, planting plants in any terrain.

Flowers, plants, shrubs, trees and other plants in landscape are described by point features. Many three dimensional modelling software have problems in details, with free hand drawing in SketchUp, trees are more realistic, details are more sophisticated, as shown in Figure 3.



FIGURE 3 SketchUp Hand-painted plants

### 3.3 PLANNING AND DESIGNING

Road landscape planning and designing is included in SketchUp software, handle road intersection automatically and set round corners at road crossings, modify road width and update the road borders and rounded corners at intersections automatically; get the block Specific, custom site name; setting garden gadgets models parametrically, the system can adequately sum up the curb, flower beds, pools, steps, Pergola, fence walls, landscape architecture feature information, as shown in Table 2.

TABLE 2 Biologically treated effluents (mg/l)<sup>a</sup>

Category	Element, Ebenename	Space Type	Data Type
Vegetation	Single-tree	Punctiform	Ponit Feature
	Path	Threadiness	Line Feature
	Green-lawn	Facet	Polygon Feature
Sidewalk	Sidewalk	Facet	Polygon Feature
	Centre-road	Threadiness	Line Feature
River	River	Facet	Polygon Feature
Main-Building	Main-building	Facet	Polygon Feature
	Surroundings	Punctiform	Ponit Feature
Infrastructure	Benches	Threadiness	Line Feature
	Stree-lights	Punctiform	Ponit Feature

Some parameters of landscaping or gardening gadgets can also be generated by selecting the input or parameter, as shown in Figure 4 for landscaping small objects.

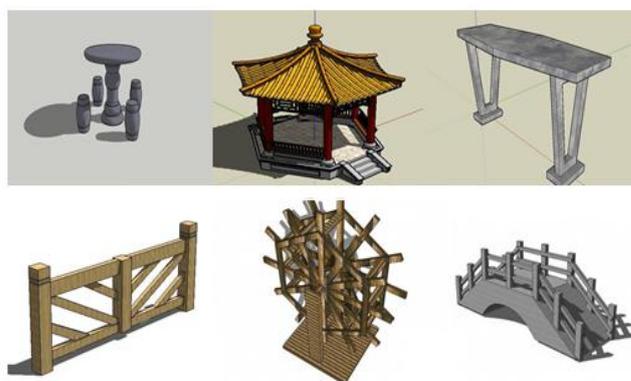


FIGURE 4 Garden gadgets maps

4 Engineering examples

Getting ready for a modelling job that needs to be modelled into components of architecture, landscape architecture, and any component can be saved as a single file, modifying or assembling a deviation can be modified at a later time.

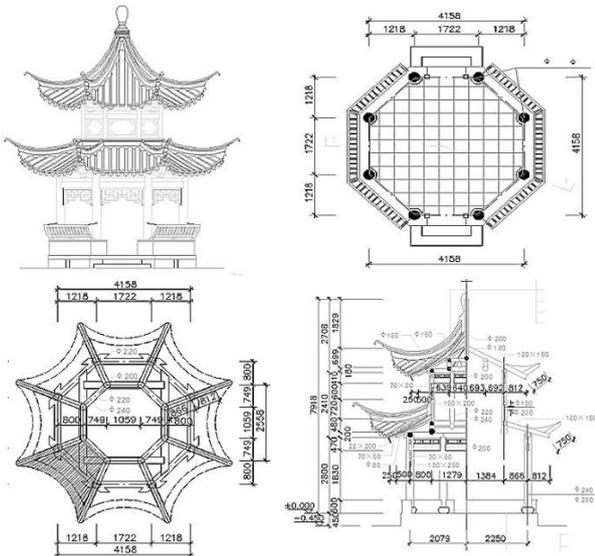


FIGURE 5 Required to establish a model of CAD files

As shown in Figure 5 Select the modelling of architectural CAD files, import into SketchUp software, and through simple depiction of each line to generate it, then make in a three dimensional model of each side. This stage demands very high quality of the CAD drawing. Representation of contours in two-dimensional are in very mess. We need to import SketchUp by CAD drawing software as a base map, change details, as shown in Figure 6. Some overlapping lines of unnecessary details in CAD, this is unnecessary in SketchUp, so make the whole graphic into component avoids complexity of works. Pull the base, without the exploded diagram above, based on maps that are not exploded. On this basis, select the staggered lines for the base model; Import ARC paths to follow and create the geometry model; Use method of creating groups to copy multiple objects of the same; Selection exploded after the merge component, removing unwanted clutter line (as shown in Figure 7).

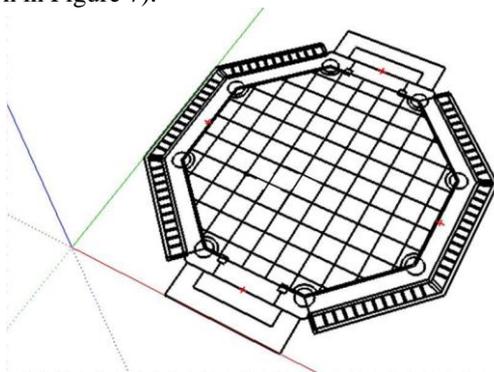


FIGURE 6 Base maps modified

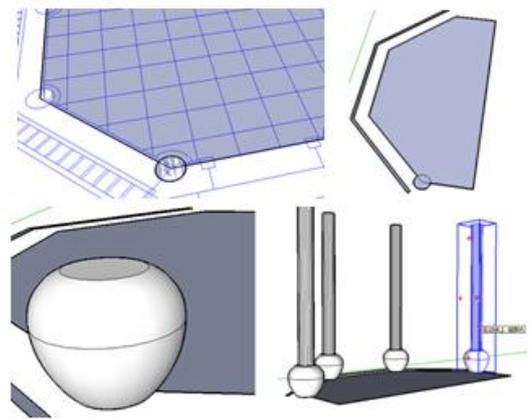


FIGURE 7 Base model

We used a group quickly copying approach in the modelling process. Intercropping in the base map guides, by rotating a certain degree, to achieve the positioning that you want to copy. Finally, drawing plane or make combination to implement a three-dimensional view of the component, as shown in Figure 8. When model's appearance substantially completed, we can get post processing. In the SketchUp get appropriate materials, click on the pattern and then attach the material model to be rendered on the needed part.

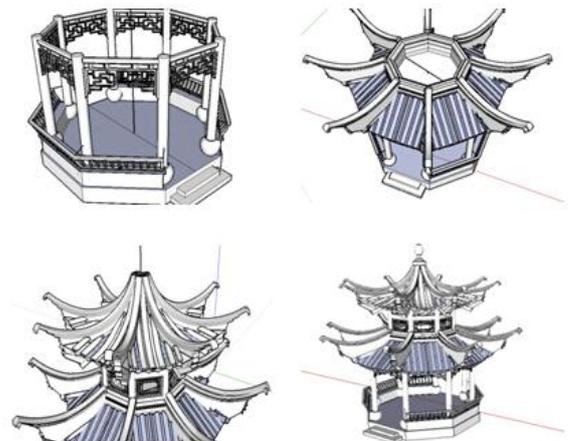


FIGURE 8 Three dimensional model stretch mix

5 Conclusion

This paper introduced variety of function modules and various tools in SketchUp software, make an overview of SketchUp software is simple to understand, powerful, compatible with good features. From three-dimension modelling aspect outlines three main designing including landscape design, planting design, planning and designing. From TIN terrain designing, hand-painted designing of landscape plants, parametric or texturing the landscape planning of the image of the object to show SketchUp modelling tool's variety and convenience. Finally, through a detailed analysis of engineering examples, from the base map locating, to the combination of individual components, to show

SketchUp software's flexibility and application. Two-dimensional image after imported into SketchUp, many lines do not work because SketchUp intelligent, various issues could appear through models from two-

dimensional graphics into a three-dimensional. Through flexible use of SketchUp software, you can easily avoid unnecessary work, this reflecting its strong functionality.

## References

- [1] Dai Y H 2010 Application of surface modelling based on three-dimension animation *Journal of Hunan Institute of science and technology* 31(12) 91-2 (in Chinese)
- [2] Du B Q 2011 Sketchup's application in landscape design *Modern gardening* (21) 130 (in Chinese)
- [3] Chen S L, Liu Z, Dong S G 2011 Application of ERDAS and SketchUp to build urban three-dimension model *Geo-spatial information* 9(1) 46-50
- [4] Ikeuehi K 2004 Constructing Virtual Cities by Using Panoramic images *International Journal of Computer Vision* 13(3) 237-47
- [5] Deng T 2007 Tourism landscape design principles *Beijing China architecture and building press* 8 the first vision (in Chinese)
- [6] Hu J Q 2009 Research in building virtual environmental towel in Google Earth Hubei Central China Normal University master degree thesis 10-6 (in Chinese)
- [7] Hashim M, Marghany M, Mahmud M R, Anuar M H 2010 Utilization of LiDAR and IKONOS Satellite Data for Security Hotspot Analysis Based on Realism of 3D City Model *The 2010 International Conference on Computational Science and Applications (ICCSA2010) Fukuoka Japan 2010 Springer* 331-45
- [8] Fu Z; Wang N; Wang Q 2010 The Virtual Campus System Based On SketchUp and OpenGL *2010 Second IITA International Conference on Geoscience and Remote Sensing (GRS2010) Qingdao China 2010 IEEE Press* 536-9
- [9] Wang B J, Shi B, Song Z 2009 A Simple approach to 3D geological modelling and visualization *Bull Eng Geol Environ* (68) 559-65

## Authors



**Xiaoxiong Wang, born in 1973, Shannxi Province, China**

**Current position, grades:** lecturer

**University studies:** Master degree

**Scientific interest:** artistic research of urban landscape

# The robot path optimization of improved artificial fish-swarm algorithm

Jiansheng Peng<sup>1, 2\*</sup>

<sup>1</sup>National Key Laboratory of Communication, UEST of China Chengdu 610054, China

<sup>2</sup>Department of Physics and Mechanical & Electronic Engineering, Hechi University, Yizhou 546300, China

Received 12 June 2014, www.tsi.lv

## Abstract

The robot path optimization solution is seek a collisionless path from starting point to end point to make the robot get the shortest route go along with planning path. Let robot path optimization problem map to mathematical model TSP (Traveling Salesman Problem) to resolve it, and make the corresponding algorithm realize robot path planning optimization was introduced in this paper. According to existent insufficiency of traditional artificial fish-swarm algorithm, using improved artificial fish-swarm algorithm optimizes the robot path planning, and stands out the superiority of improved artificial fish-swarm algorithm. The main improved aspect of artificial fish-swarm algorithm is increases examine link in foraging behaviour.

*Keywords:* robot path optimization, Traveling Salesman Problem, improved artificial fish-swarm algorithm, foraging behaviour

## 1 Introduction

Optimization issues are always the researchers' potluck, no matter in engineering research field or scientific research field, its raging flames is never disappearing. As traditional optimization method cannot satisfied the requirement of persons solve the complex questions. So many home and abroad researches step on explorer multi-intelligence optimization algorithm journey, uncovered all kinds of mystery of optimization algorithm. According to characteristics of all sorts of optimization algorithm, select the superior and eliminate the inferior, perform the multi-mixture intelligence optimization algorithm climax [1]. Robot path optimization is one of most important robotics' researches fields, not only can make robot walk optimization path route, but also can reflect the robot accomplish the working performance well. It is play an important role in path planning optimization. Robot path planning optimization can apply to robots obstacle avoidance walk in known environment, robots put out fire, robots rescue and relief work, robots process route planning [2], mobile robots clean tank [3-5], robots carry out the weld and assembling and so on. It is save robots work time and reduce devoted fund. TSP is a particular case miniature of robot path planning, which is to say the TSP is equivalently to robots go N different places to pick up goods, and then go back to original location, every place just can go over once to seek the shortest path. So it can abstract to the research of robot path optimization from traveling salesman problem. There is scientific feasible evidence based to verify the effectiveness of improved artificial fish-swarm algorithm,

and it also brings energy that applies to improve artificial fish-swarm algorithm.

## 2 The Description of traveling salesman problem

TSP is one of typical combinatorial optimization problem, which is easy to describe but hard to solve. The complexity of problem rises with the way of problem's scale increase is in the exponent, which even may produce combination explosion. It is play an important role to research it. TSP problem can describe as: if there are n cities, the distance between every two cities is known, a merchant go out to sell his goods from any city, it's require to go around these n cities and every city just can visit only once, go back to original at last. How to traversal can make all journeys be the shortest. The substance of TSP is finds the shortest Hamilton path in n nodes' complete graph, and the robot path planning is so.

The mathematical model of TSP problem describes [6] as:

$$\left\{ \begin{array}{l} \min \sum d_{ij} x_{ij} \\ s.t. \sum_{i=1}^n x_{ij} = 1, i = 1, 2, \dots, n \\ \sum_{j=1}^n x_{ij} = 1, j = 1, 2, \dots, n \\ \sum_{i,j \in s} x_{ij} \leq |s| - 1, 2 \leq |s| \leq n - 2, s \subset \{1, 2, \dots, n\} \\ x_{ij} = \{0, 1\}, i, j = 1, 2, \dots, n, i \neq j \end{array} \right.$$

\* Corresponding author e-mail: Sheng120410@163.com

TSP has symmetry and asymmetry, when  $d_{ij} = d_{ji}$ , it is symmetry distance TSP, or it is asymmetry distance TSP.

The first formula above is an objective function, which expresses the distance between merchant traverse path of all cities and go back again is the minimum. The second formula expresses the merchant just right arrive at city i one time. The third formula expresses merchant just leave city j one time, together with the second formula, they are express every city has already visit once, thus avoid sub-loop happened. The fourth formula carries the merchant's point that cannot form loop in any city proper subset. Among them,  $|s|$  expresses the number of set s elements. The decision variable in fifth formula  $x_{ij} = 1$ , expresses the trail path include in merchant from city i to city j,  $x_{ij} = 0$  means the merchant is not select this way to travel path. The constraint of  $i \neq j$  making the number of decision variable reduced as  $n*(n-1)$ .

**3 Fish-swarm algorithm introduction and correlation definition.**

**3.1 FISH-SWARM ALGORITHM INTRODUCTION**

Fish-swarm algorithm introduction, which was put forward by Li Xiaolei [1] and others, is a new type swarm intelligent optimization algorithm, which is, imitates the fish's behaviour in nature. Use the top-down design idle has such characters: the requirement of objective function's properties, initial value and parameter is not high, parallelism, global superiority, speediness, traceability and so on. It also has shortcomings: The earlier stage of convergence speed faster than later obviously; the choice of algorithm parameters can effect on convergence speed and result accuracy; the more artificial fish-swarm number and the time need more; wide find optimization range and small changes may lead to convergence speed slow, and the search efficiency will be unsatisfactory.

There are four basic behaviours for artificial fish: (1) foraging behaviour: artificial fish swim to the high food concentration; this is the foundation of algorithm convergence. (2) Bunching behaviour: artificial fish swim to the high food concentration and not very crowd around fish-swim centre; this is the stability and global safeguard of algorithm convergence. (3) Rear-end behaviour: artificial fish rear-end individual fish which with high food concentration and not very crowd around individual fish; this is add the algorithm convergence's speed ability and global. (4) Random behaviour: artificial fish swim freely in water, expand the search range. According to the requirement and property of problem, rational planning and setting artificial fish-swarm algorithm model to solve it. Lump together, the main idea is generate a certain number of artificial fish-swarm, search optimize path in search field by imitate fish behaviour. In TSP, the

quantity of cities is proportional to artificial fish-swarm scale, means the more the cities quantity, which robots need to travel, the more the artificial fish-swarm individual, otherwise less.

**3.2 FISH-SWARM ALGORITHM CORRELATION DEFINITION**

Definition 1: Suppose artificial fish's current location state is  $X_i = (x_1, x_2, \dots, x_n)$ , artificial fish's next state is  $X_j = (x_1^v, x_2^v, \dots, x_n^v)$ , and then this process can be expressed as:  $x_j = x_i + visual * r$ ,  $i=1,2,\dots,n$ ,

$$X_{next} = \frac{X_j - X_i}{\|X_j - X_i\|} * Step * r$$

Definition 2: Suppose G is artificial fish's set, there is  $N(X, k) = \{X' | d(X, X') \leq k, X' \in G\}$  to artificial fish X,  $N(X, k)$  called X's k neighbourhood,  $X' \in N(X, k)$  expresses the X's neighbourhood in distance k.

Definition 3: The centre location of artificial fish  $X_1, X_2, \dots, X_q, \dots, X_m$  is:

$$Center(X_1, X_2, \dots, X_q, \dots, X_m) = \left( \frac{\sum_{q=1}^m x_{q1}}{m}, \frac{\sum_{q=1}^m x_{q2}}{m}, \dots, \frac{\sum_{q=1}^m x_{qj}}{m}, \dots, \frac{\sum_{q=1}^m x_{qn}}{m} \right)$$

$(i = 1, 2, \dots, m, j = 1, 2, \dots, n)$ .

In this form,  $Center(X_1, X_2, \dots, X_q, \dots, X_m)$  expresses centre location of artificial fish  $X_1, X_2, \dots, X_q, \dots, X_m$ , shows the every components average of selected m fish.

Definition 4: The distance between artificial fish expressed as  $d_{i,j} = \|x_i - x_j\|$ .

Definition 5: step expresses moving step of artificial fish, date is crowding factor of artificial fish, try\_number is artificial fish's maximum explore time, visual is visual scope of artificial fish and maxgen is artificial fish's maximum iterate time.

Definition 6:  $X_i$  is current artificial fish,  $X_j$  is next state artificial fish,  $n_f$  means number of partner,  $X_c$  is fish-swarm centre,  $X_e$  is food concentration in fish-swarm centre.

In TSP, initialize every artificial fish is one of random permutations which including all cities number to every artificial fish, that is encoding reasonably. Provide a city number to arrange randomly can initialize an artificial fish. According complexity of problem and algorithm need to create a certain artificial fish. This can be used to similar robot path optimization issues. Its nature is same with TSP. Artificial fish centre is the number of city serial number which appear the most times, statistics n places that are different from each other and the number appear more times, the centre position is appear the most times in the same place. The distance between two fish is the count of different city serial number in one place.

These two fish are neighbour if the distance less than visual, means a fish in another fish's field. The city sort of two artificial fish is:  $A = \{a_1, a_2 \dots a_n\}$  and

$B = \{b_1, b_2 \dots b_n\}$  the distance between them is

$$\text{distance}(A, B) = \sum_{i=1}^n \text{sign}(|a_i - b_i|)$$
 in it,

$$\text{sign}(x) = \begin{cases} 0, & x = 0 \\ 1, & x > 0 \\ -1, & x < 0 \end{cases}$$

### 3.3 THE BASIC BEHAVIOR OF ARTIFICIAL FISH

#### 3.3.1 Foraging behaviour

*Foraging behaviour* is searched arrange information including city position serial number in field by current artificial fish. Change position among  $m$  cities by itself randomly. It is need to satisfied  $m \leq \frac{1}{2} \text{visual}$ , thus can ensure the artificial fish foraging in visual range. If there is more optimize individual fish status than current fish status when foraging, then use artificial fish which more optimize than current fish status to take place of it. That is to say, travel city path is shorter than current artificial fish information walk city route. If execute `try_number` times with above behaviour in common, foraging behaviour failure, then carry out random behaviour.

#### 3.3.2 Bunching behaviour

Find all neighbours of current fish at first, then statistic every city serial number, which appears the most same city serial number in one place all neighbours in current artificial fish field. The most appear serial number in each column formed arrangement is artificial fish centre. If there is some serial numbers appeared times are greatest at the same time, value a number among them randomly. If there is repeated city serial number in one row  $n$  column, disposed it with missing code replace repeated serial number to make it as without repeat arrangement from 1 to  $n$ . If the centre city path less than current city path and satisfied with  $\frac{Y_c n_f}{Y_i} < \delta$ , use several bunching

centre status position serial number replace corresponding current artificial fish status position serial number, the new status artificial fish act as next status artificial fish. If there is repeated code in arrangement number after replacing one row  $n$  column, it should be processed. It will execute foraging behaviour once bunching behaviour failed.

#### 3.3.3 Rear-end behaviour

Search the highest food concentration individual  $X_{min}$  in  $n_f$  neighbourhoods of current artificial fish

neighbourhood space  $N(X, \text{visual})$ , the food concentration is  $Y_{min}$ . If  $\frac{Y_{min} n_f}{Y_i} < \delta$  and  $Y_{min} < Y_i$ , make

the status of most optimize artificial fish  $X_{min}$  as current fish status. That is to say, if there is information which the neighbourhoods of artificial fish included in travelled city walk route shorter than itself and around it not crowded, then use several city position serial numbers of artificial fish short city route randomly instead of current artificial fish city position serial number. Use the same way to process it to generate reasonable city position serial number as next status fish. It will execute foraging behaviour if rear-end behaviour failure.

#### 3.3.4 Random behaviour

Random behaviour provides artificial fish a city arrangement randomly in artificial fish sensing range. Current fish  $X_i$  swim in visual range randomly; this is a default behaviour of foraging behaviour. Realize random behaviour rather simple, it increase artificial fish search range.

### 3.4 THE IMPROVEMENT OF FISH-SWARM ALGORITHM

Add examine link: artificial fish of improved foraging behaviour add a examine link on the basis of primary artificial fish foraging behaviour. It means artificial fish seek  $k$  different status in its visual range, choose searched most optimize status as current artificial fish's next status. Thus can decrease blindness of artificial fish find optimization, and it can make artificial fish have multiple choices in every swimming. This can enhance efficiency greatly, and get better value faster. In TSP, `try_number` each time, current artificial fish searching in neighbourhood field, it will change randomly among  $m$  cities' position. After repeating  $k$  times, choose the best status enter into next `try_number` until reach the most tempter times. The possibility will be more likely to find current status individual fish status as next status better than current fish status.

The pseudo-code descriptions of adding examine link foraging behaviour as follows:

```
float Artificial_fish::AF_pre()
{
    Xmin=Xi; //Xmin is own the most optimize food
    //concentration which search in visual range
    Ymin=Yi; //Ymin express abundance food that
    //own the most optimize food concentration
    //which search in visual range.
    for(i=0;i<try_number;i++) // try_number is artificial
    //fish tempter times
    {
        for(k=0;k<10;k++) //examining other fish status
        //10 times in visual range
        {
```

```

Xj = Random(N(Xi,visual)); //search next status
//randomly in neighborhood field
If (Ymin>Yj) //if it satisfied searched fish own
//food concentration better than most
//optimize fish food concentration
Ymin.
{
Xmin=Xj;//value the status of fish Xj to Xmin
Ymin=Yj; // value the food concentration of
//fish Xj to Xmin
}
}
If (Ymin<Yi) //judge searched most optimize food
//concentration whether better than
//primary food concentration
Xinext = Xj; //use artificial fish which searched
//own most optimize food concentration as next status
else //or moving randomly
Xinext = Random*step;
}
Return AF_foodconsistence ( Xinext ); //return food
//concentration value of next status artificial fish
}
    
```

Step (6): draw curve graph of every optimize process and most optimize solution path diagram.

The algorithm flow diagram is followed 3.6.1:

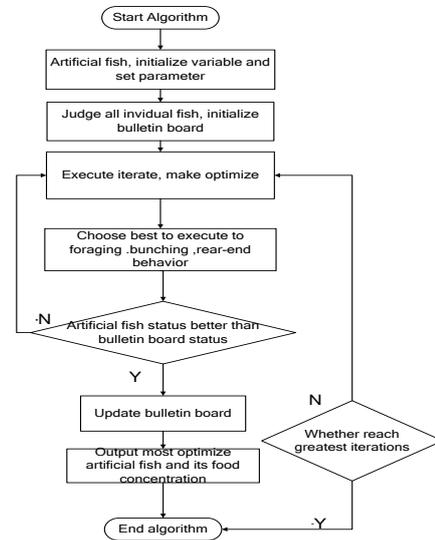


FIGURE 1 Algorithm flow diagram

### 3.5 ALGORITHM PROCEDURE AND FLOW DIAGRAM

Step (1): initialize fish-swarm AF and current fish  $y_u$  and set parameter: crowding factor  $\delta$ , tempter times  $try\_number$ , examine times  $k$ , sensing range  $visual$ , greatest iterations  $maxgen$ , the number of artificial fish-swarm individual  $N$ , city numbers  $M$ , moving step length  $step$ .

Step (2): initialize variable: store most optimize artificial fish status  $zui\_you$  on bulletin board, store most optimize artificial fish food concentration  $zui\_min$  on bulletin board, store food concentration  $Y$  of  $N$  artificial fish, store data  $shu\_ju$  of total distance, store next status  $AF\_next$  of  $N$  artificial fish, store initialize food concentration  $Y\_next$  of artificial fish next status.

Step (3): output random solution, which get of walk route map and total distance of current fish  $y_u$ .

Step (4): make optimize, search most optimize artificial fish individual in initialized artificial fish, update bulletin board and record total distance in the meantime.

Step (5): begin iteration, make behaviour judgment choose. Call functions foraging behaviour, bunching behaviour, rear-end behaviour and random behaviour, compare and judge which behaviour should be execute for artificial fish to get the best next status and food concentration, then choose the best one to execute. At the same time, update the best artificial fish status and food concentration on bulletin board, as well as update artificial fish-swarm status and them food concentration. Record total distance with each iteration get.

### 4 Simulation experiment and result analysis

Adapt to MATLAB R2010a as programming software, solve on computer which CPU is 792MHZ, memory is 2.00GB, operating system is Windows XP.

Table 1 shows the data [7] for 14 cities BURMA14 position coordinate in standard TSPLIB library to use to validate the effectiveness of improved artificial fish-swarm algorithm. It also uses to provide thinking which solve robot path optimization problem better by the application of this algorithm. Set artificial fish-swarm individual number  $N=10$ , cities number  $M=14$ , visual range  $visual=9$ , tempter times  $try\_number=8$ , examine link times  $k=10$ , crowding factor  $\delta=6$ , moving step length  $step=3$ , maximum iterate time  $maxgen=20$ , let 10 fish search optimization. These 14 cities may exist the route is  $\frac{(14-1)!}{2} = 3113510400$  if use enumeration method, the best known solution is 30.8785.

TABLE 1 the data for the TSP problem of 14 cities

City Number			City Number		
	abscissa X	ordinate Y		abscissa X	ordinate Y
1	16.47	96.10	8	17.20	96.29
2	16.47	94.44	9	16.30	97.38
3	20.09	92.54	10	14.05	98.12
4	22.39	93.37	11	16.53	97.38
5	25.23	97.24	12	21.52	95.59
6	22.00	96.05	13	19.41	97.13
7	20.47	97.02	14	20.09	94.55

The most optimize path orbit diagram and iterate change of optimize process which get by improved



- [3] Zhou Likun, Liu Hongzhao 2012 The Application of Adaptive Artificial Fish Swarm Algorithm in Clean Tank Mobile Robots Path Plan *Journal of Xi'an University of technology* **31**(7) 1085-9
- [4] Zhou Y Q, Huang X S, Yang Y, Wu J Z 2012 Hybrid Optimization Algorithm based on Mean Particle Swarm and Artificial Fish Swarm *Information-An International Interdisciplinary Journal* **15**(2) 763-77
- [5] Peng Jiansheng, Li Xing, Luo Guan 2012 Two Kinds of Improved Artificial Fish Swarm Algorithms *Journal of Nanjing University of Science and Technology* **36** 212-6
- [6] Shen W, Guo X P, Wu C, Wu D S 2012 Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm *Knowledge-Based Systems* **24**(3) 378-85
- [7] Standard TSPLIP [OL/DB] 1997 Online Available: <http://elib.zib.de/pub/Packages/mp-testdata/tsp/tsplib/tsplib.html>
- [8] Huang Xiaoyan, Wen Zhan, Fu Kechang, etc 2009 Based on Improved PSO Automobile Path Optimization *Journal of Xiangtan University Natural Science* **31**(2)

## Authors



**Jiansheng Peng**

**Current positions, grades:** The Hechi university teaching in 2004, Associate Professor in 2013, Engineers titles in 2010, China Institute of Atomic Energy Ph.D. in 2010, University of Electronic Science and Technology postdoctoral fellow in 2013.

**Research direction:** Robot path planning navigation, development and application of embedded.

# Research on lead-time reduction of two-stage supply chain based on Stackelberg game

**Hao-ran Shi<sup>1, 3</sup>, Kejian Liu<sup>2\*</sup>**

<sup>1</sup>*School of Energy and Environment, Xihua University, 610039, Chengdu, P.R.China*

<sup>2</sup>*School of Mathematics and Computer Engineering Xihua University, 610039 Chengdu, P.R.China*

<sup>3</sup>*School of Transportation and Logistics, Southwest Jiaotong University, 610031, Chengdu, P. R. China*

*Received 12 June, 2014, www.tsi.lv*

---

## Abstract

In the two-stage supply chain, under the model of lead time reduction management cost shared by upstream and downstream based on Stackelberg Game, when suppliers have the priority of decision right rather than retailers, it is more advantageous to reduce the cost and the lead time and can reach the maximum profit for the whole supply chain.

*Keywords:* Supply chain, Lead time, Decision order, Cost sharing

---

## 1 Introduction

Lead time refers to the interval from ordering to receiving goods in the downstream delivered by suppliers in the upstream. This is also called inventory replenishment lead time. Those in the downstream hope suppliers to reduce the lead time so as to reduce inventory and cost. To reduce the lead time, suppliers usually needs some extra investment, for example, buy new equipment, improve or set up new information system or upgrade inventory equipment. However, many enterprises can't afford such a huge investment and have to shift the cost to those in the downstream. When the cost is on buyers in the downstream, some suppliers take it for granted that buyers should shoulder all cost for reducing the lead time. If buyers are willing to do so, then they are granted with the right to reduce the lead time and decide the length of it.

When suppliers decide the length of inventory lead time, there are usually two decision orders: one is that retailers decide how many goods to order and inform it of suppliers, and leave it to suppliers to decide the length of the lead time; the other is that suppliers decide the time of arrival and retailers decide when to order and how many to order based on the delivery situation.

Suppliers are facing the following questions: when to decide the lead time, before retailers' order or after? What is the best lead time for ordering so as to reduce the cost as much as possible? What will be the influence on the cost if the decision order between the upstream and the downstream is exchanged?

## 2 Literature review

Many researchers have focused on the importance of reducing the lead time from several perspectives. Perry, M. Ben-Daya and Zhang describe the random and swift response model [1-3].

Swift response model is necessary when orders are given at the same time [4]. In two recent articles, some researchers propose an effective qualitative model for the supply chain [5-7]. Many researches study the ordering lead time decision [8-11]. They suppose that those in the downstream decide the ordering lead time and shoulder the cost for reducing the lead time. Moreover, researchers also study the lead time decision made by retailers for the maximum profit. However, although suppliers have shifted the cost for reducing the lead time to retailers, it doesn't mean that cost of suppliers is free from the influence of retailers' decision on ordering. It is because under the condition that suppliers have the priority to decide the length of the lead time, the cost for reducing it will be affected by the cost of ordering and further, the quantity of ordering which may lead to a drop of profit for suppliers. Thus, suppliers do not favour such strategy. To move a step further, if retailers cannot afford the cost alone, they will give up reducing the lead time.

## 3 Model description and establishment

### 3.1 STACKELBERG GAME MODEL

Suppose there are two producers who take the turn to decide the production in a two-stage Stackelberg Game. In the first stage, producer 1 as the leader takes the priority to plan for the production. In the second stage,

---

\*Corresponding author e-mail: liukejian@gmail.com

producer 2 as the follower plans for the production under the principle of obtaining the maximum profit after learning about the yield level of the leader. Suppose the marginal costs of two producers are the same,  $c_1 = c_2 = c$ , the market demand function is  $D = a - (q_1 + q_2)$ , in which  $a > 0$ ,  $a$  is a constant.  $q_i$  is the production of producer  $i$ ,  $i = 1, 2$ . This function is known by two producers.

By the backward induction, we consider the second stage first. Suppose the production of producer 1 is  $q_1$ , the optimal production  $q_2^s$  of producer 2 is:

$$q_2^s \in \arg \text{gMax}_{q_2} \{ \pi_2(q_1, q_2) = [a - (q_1 + q_2) - c]q_2 \}$$

Based on the first order condition, we can get the optimal reaction function for producer 2:

$$q_2 = R_2(q_1) = \frac{a - q_1 - c}{2}$$

Then, we consider the first stage and predict the reaction function for producer 2  $q_2 = R_2(q_1) = \frac{a - q_1 - c}{2}$ .

The optimal production  $q_1^s$  of producer 1 is

$$q_1^s \in \arg \text{gMax}_{q_1} \{ \pi_1(q_1, q_2) = [a - (q_1 + R_2(q_1)) - c]q_1 \}$$

$$= [a - (q_1 + \frac{a - q_1 - c}{2}) - c]q_1$$

By the first order condition, we can get the optimal production  $q_1^s = \frac{a - c}{2}$  of producer 1. So,

$$q_2^s = R_2(q_1^s) = \frac{a - q_1^s - c}{2} = \frac{a - c}{4}$$

Thus, the result of Stackelberg Game is  $q = q_1^s + q_2^s = \frac{3(a - c)}{4}$ ;  $D^s = a - q = \frac{a + 3c}{4}$ . The profits of

$$\text{two producers are: } \begin{cases} \pi_1^s = (D^s - c)q_1^s = \frac{(a - c)^2}{8} \\ \pi_2^s = (D^s - c)q_2^s = \frac{(a - c)^2}{16} \end{cases}$$

However, in actual economic activities, producers cannot know exactly about the market demand function but only estimate it. Nevertheless, such estimation varies from person to person. Here we suppose that every producer takes it for granted that the estimation of his counterparts is the same as his.

Suppose the market demand function estimated by producer 1 is  $D = a_1 - (q_1 + q_2)$ , and that he thinks producer 2 estimates the same.

The market demand function estimated by producer 2 is  $D = a_2 - (q_1 + q_2)$ , and that he thinks producer 1 estimates the same. Here  $a_i, i = 1, 2$ , which are above 0.

Under such circumstances, we divide the Stackelberg Game Model into two stages as the conventional way. And apply it to backward induction method. First we consider the second stage, set the production of producer 1 is  $q_1$ , as producer 2 thinks the market demand function to be  $D = a_2 - (q_1 + q_2)$ . If producer 2 produces  $q_2$ , producer 2 thinks that his profit is expected to be  $\pi_2(q_1, q_2) = [a_2 - (q_1 + q_2) - c]q_2$ , from the first order condition, we can get the optimal reaction function of producer 2 is  $q_2 = R_2(q_1) = \frac{a_2 - q_1 - c}{2}$ .

Let's be back to the first stage, as producer 1 estimates the market demand function is  $D = a_1 - (q_1 + q_2)$ , and thinks that the production of producer 2 will be  $q_2$  in the second stage, then the profit is  $\pi_1(q_1, q_2) = [a_1 - (q_1 + q_2) - c]q_1$ . From the first order condition, we can get the optimal reaction function of producer 2 estimated by producer 1:  $q_2 = R_2(q_1) = \frac{a_2 - q_1 - c}{2}$ . In this case, producer 1 thinks

his profit is  $\pi_1(q_1, q_2) = [a_1 - (q_1 + R_2(q_1)) - c]q_1$ . From the first order condition, producer 1 thinks that his optimal production is  $q_1 = \frac{a_1 - c}{2}$ , in which  $q_1 = a_1 - c_2$ .

$$q_2 = R_2(q_1) = \frac{a_2 - q_1 - c}{2}$$

Let's be back to the second stage. As the production of producer 1 is, producer 2 thinks that his optimal reaction function is:

$$\text{Thus, producer 2 thinks that his optimal production is: } q_2 = \frac{a_2 - q_1 - c}{2} = \frac{2q_2 - a_1 - c}{2}$$

The actual market demand function is  $D = a - (q_1 + q_2)$ , so the market price will be:  $p = a - (q_1 + q_2) = a - [\frac{q_1 - c}{2} + \frac{2a_2 - a_1 - c}{4}] = \frac{4a - a_1 - 2a_2 + 3c}{4}$ .

$$\text{Therefore, the profits of two producers are: } \begin{cases} \pi_1 = (D - c)q_1 = \frac{(q_1 - c)(4a - a_1 - 2a_2 - c)}{8} \\ \pi_2 = (D - c)q_2 = \frac{(2a_2 - a_1 - c)(4a - a_1 - 2a_2 - c)}{16} \end{cases}$$

### 3.2 INVENTORY MODEL BASED ON (Q, r) STRATEGY

(Q, r) inventory strategy is a common inventory management strategy, in which the warehouse manager checks the inventory with continuity. When the existing inventory drops to the replenishment point  $r$ , the manger will order  $Q$  goods from suppliers in the upstream.

Those goods will arrive after the lead time  $L$ . The variation of inventory under  $(Q, r)$  strategy is shown in Figure 1.

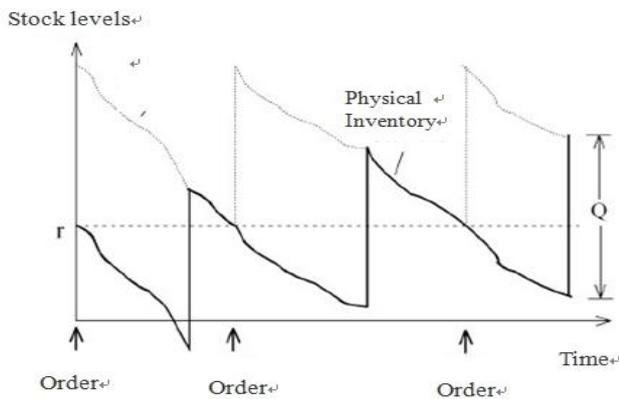


FIGURE 1 The variation of inventory

Here we discuss the replenishment cost with fixed lead time.  $D^L$  is the aggregate demand during the lead time, then function  $G(y)$  is expressed as:

$$G(y) = E[h(y - D^L)^+ + g(D^L - y)^+]$$

$D$  is the demand per unit time. It is a random variable.  $\lambda$  refers to the demand rate of market;  $R$  refers to fixed ordering cost;  $g$  / (piece·per unit time) refers to replenishment cost;  $h$  / (piece·per unit time) refers to holding cost; the replenishment lead time is  $L$ .  $\lambda$ ,  $R$ ,  $g$ ,  $h$ ,  $L$  are all constants.

When the demand is reached in the way that the demand grows stably and randomly and keeps such growth, under such condition, if inventory  $y$  is subject to the even distribution of  $(r, r + Q]$  (when the demand is discrete,  $y$  is subject to the even distribution of  $\{r, r + 1, \dots, r + Q\}$ )

When the demand is a continuous variable, the average cost per unit time is expressed as:

$$C(Q, r) = \frac{R\lambda + \int_r^{r+Q} G(y)dy}{Q}$$

When the demand is a discrete variable, the average cost per unit time is expressed as:

$$C(Q, r) = \frac{R\lambda + \sum_r^{r+Q} G(y)}{Q}$$

### 3.3 MODEL ESTABLISHMENT

We have studied the two-stage supply chain consisting of suppliers and retailers with single product. The ordering lead time  $L$  refers to the interval between ordering and receiving. It can be divided into  $n$  separated parts. Part  $i$  has the minimum interval  $a_i$  and the standard time  $b_j$ . If

$$L_o = \sum_{j=1}^n b_j$$

is used to express the initial ordering lead

time of the supply chain,  $L_i$  is to express the minimum length of ordering lead time of part  $1, 2, \dots, i$ , then there is:

$$L_i = \sum_{j=1}^i a_j + \sum_{j=i+1}^n b_j = \sum_{j=1}^i a_j + \sum_{j=i}^n b_j - \sum_{j=i}^i b_j = \sum_{j=1}^n b_j - \sum_{j=1}^i (b_j + a_j) = L_o - \sum_{j=1}^i (b_j + a_j) \tag{1}$$

The ordering lead time  $L$  can be shortened by adding some cost. Therefore, it is controllable. Suppose the cost for reducing a unit time for part  $i$  is  $C_i$ , and  $c_1 \leq c_2 \leq \dots \leq c_n$  refers to the cost of the ordering lead time as well as the cost for reducing the ordering lead time. Then, there is:

$$K(L) = c_i(L_{i-1} - L) + \sum_{j=1}^n C_j(b_j - a_j), L \in (L_i, L_{i-1}), \tag{2}$$

$$K(L_0) = 0, K(L_n) = \sum_{j=1}^n C_j(b_j - a_j). \tag{3}$$

The cost of suppliers and retailers and the cost of ordering constitute the cost of inventory. The cost of ordering and that of the lead time management are shouldered by suppliers. The aggregate ordering cost per unit time is  $\Pi_m(L, Q)$ . There is:

$$\Pi_m(L, Q) = \frac{D}{Q} C_m + \frac{Q}{2} P_m h + \theta \frac{D}{Q} K(L). \tag{4}$$

$D$  is the average demand per unit time;  $Q$  is the quantity of orders;  $h$  is the inventory holding cost rate;  $C_m$  is the cost of a single supplier;  $p$  is the purchasing price of the supplier;  $\theta$  is the ratio of cost for reducing the ordering lead time shouldered by suppliers. Given that  $(Q, r)$  strategy is very common in researches and actual situations, this paper also employs this strategy as the inventory strategy. Suppose retailers are faced with the natural distribution of demand during the lead time, averaging  $LD$  and the standard variation to be  $\sigma\sqrt{L}$ . Then we can know the average inventory level per unit time is  $I_p \cong \frac{Q}{2} + K\sigma\sqrt{L}$ .  $K$  is the inventory security coefficient. The ordering cost of suppliers consists of average inventory cost, ordering cost, and the lead time management cost. The aggregate ordering cost per unit time is  $\Pi_m(L, Q)$ . Then there is:

$$\Pi_r(L, Q) = \frac{D}{Q} C_r + (\frac{Q}{2} + K\sigma\sqrt{L})P_r h + (1 - \theta) \frac{D}{Q} K(L). \tag{5}$$

$C_r$  is the ordering cost of a single retailer.  $P_r$  is the purchasing price of retailers.  $(1-\theta)$  is the ratio of cost for reducing the lead time shouldered by retailers.

Usually, the ordering process is that the upstream decides the arrival time of goods and the downstream decides the quantity of goods. This paper compares two cases: one is that the upstream decides the ordering lead time first, the other is retailers decide the quantity of goods first.

**4 Comparison of different decision orders**

Usually, the ordering process is that the upstream decides the arrival time of goods and the downstream decides the quantity of goods. This paper compares two case: one is that the upstream decides the ordering lead time first, the other is retailers decide the quantity of goods first.

**4.1 THE OPTIMAL MODEL OF SUPPLIERS' PRIORITY DECISION**

Suppose suppliers are equipped with relevant information of retailers, such as the ordering cost and inventory cost. Suppliers have the priority to decide the ordering lead time and then leave it to retailers to decide the quantity of ordering. Obviously, in this Stackelberg Game, suppliers stand as leaders while retailers as the followers.

First, we consider the optimal ordering quantity of retailers under the condition that the ordering lead time  $L$  is given. The optimal decision mode of retailers is:

$$\min \Pi_r(L, Q) = \frac{D}{Q} C_r + \left(\frac{Q}{2} + K\sigma\sqrt{L}\right) P_r h + (1-\theta) \frac{D}{Q} K(L). \quad (6)$$

Calculate the derivatives of  $Q$  in the cost function and equal it to 0.

$$\frac{d\Pi_r(L, Q)}{dQ} = \frac{D}{Q^2} [C_r + (1-\theta) \frac{D}{Q} K(L)] + \frac{P_r h}{2} = 0, \quad (7)$$

$$\frac{d^2 \Pi_r(L, Q)}{dQ^2} = \frac{dD}{Q^3} [C_r + (1-\theta) K(L)] = 0. \quad (8)$$

As  $L \in (L_i, L_{i-1})$ ,  $\Pi_r(L, Q)$  is the concave function about  $Q$ , then the optimal quantity is:

$$Q^* = \sqrt{\frac{2D[C_r + (1-\theta)K(L)]}{P_r h}}, L \in (L_i, L_{i-1}), \quad (9)$$

$$\begin{aligned} \frac{dQ^*}{dL} &= -(1-\theta) \frac{DC_i}{P_r h} \left\{ \sqrt{\frac{2D[C_r + (1-\theta)K(L)]}{P_r h}} \right\} \\ &= -(1-\theta) \frac{DC_i}{P_r h Q^*} \end{aligned} \quad (10)$$

Thus,  $\frac{dQ^*}{dL} < 0$ . The ordering quantity of retailers will increase along with the reduction of the lead time. When suppliers can predict the quantity decided by retailers based on formula (9), there is:

$$\min \Pi_m(L, Q) = \frac{D}{Q} C_m + \frac{Q}{2} P_m h + \theta \frac{D}{Q^*} K(L). \quad (11)$$

For each  $(L_i, L_{i-1})$ , we can get the derivatives of  $\Pi_m(L, Q^*)$  to  $L$ .

$$\frac{d\Pi_m(L, Q)}{dL} = -\frac{D}{Q^2} [C_m + \theta K(L)] \frac{dQ^*}{dL} - \frac{DC_i}{Q^*} \theta + \frac{P_m h}{2} \frac{dQ^*}{dL}. \quad (12)$$

Substitute (9) to (12) and get:

$$\frac{d\Pi_m(L, Q)}{dL} = \left\{ -\frac{C_m + \theta K(L) P_r h}{2[C_r + (1-\theta)K(L)]} + \frac{\theta P_r}{1-\theta} + \frac{P_m h}{2} \right\} \frac{dQ^*}{dL}. \quad (13)$$

Because  $\frac{d\Pi_m(L, Q^*)}{dL} = 0$ , it is easy to get:

$$\frac{C_m + \theta K(L)}{C_r + (1-\theta)K(L)} = \frac{2\theta}{1-\theta} + \frac{P_m}{P_r}. \quad (14)$$

When  $K(L) > 0$ , we can get:

$$\frac{d\Pi_m(L, Q^*)}{dL} \Big|_{L=L_i} = -\frac{P_r C_i [(1-\theta)C_m - \theta C_r]}{2[C_r + (1-\theta)K(L)]^w} \frac{dQ^*}{dL} > 0. \quad (15)$$

Thus, when  $K(L_i) \in [0, K(L_n)]$ , the optimal ordering lead time of suppliers is  $L^* = L_i$ . If  $K(L_i) > K(L_n)$ , then  $L^* = L_n$ ; If  $K(L_i) < 0$ , then,  $L^* = L_0$ . Thus the optimal ordering quantity is

$$Q^* = \sqrt{\frac{2D[C_r + (1-\theta)K(L)]}{P_r h}}, L \in (L_i, L_{i-1}). L^* \text{ and } Q^* \text{ are}$$

the equilibrium of the Stackelberg Game with suppliers as leaders. Compare this result and the original ordering lead time and quantity, there is:  $\Pi_m(L^*, Q^*) \leq \Pi_m(L_0, Q_0)$ .

**4.2 THE OPTIMAL MODEL OF RETAILERS' DECISION PRIORITY**

Suppose retailers first decide the ordering quantity and leave it to suppliers to decide the ordering lead time. Obviously, such decision order is featured by retailers' decision priority in the Stackelberg Game.

First we assume that suppliers make the decision of the ordering lead time after they know about the ordering quantity  $Q$ . For them, there is the following model:

$$\min \Pi_m(L, Q) = \frac{D}{Q} C_m + \frac{Q}{2} p_r h + \theta \frac{D}{Q} K(L), \quad (16)$$

$$\frac{d\Pi_m(L, Q)}{dL} = -\frac{DC_i}{Q} \theta < 0. \quad (17)$$

Thus, in this case, the optimal ordering lead time is  $L_0$  and  $K(L_0) = 0$ . When retailers predict that the ordering lead time is  $L_0$  and when  $L \in (L_i, L_{i-1})$ ,  $\Pi_r(L, Q)$  is the concave function about  $Q$ . the optimal ordering quantity of retailers is:

$$Q_0 = \sqrt{\frac{2DC_r}{p_r h}}. \quad (18)$$

$L_0$  and  $Q_0$  are the equilibrium of the Stackelberg Game with retailers as leaders. Whatever the ordering quantity is, for suppliers, the ordering lead time is  $L_0$ . Retailers' decision does not affect the lead time.

Comparing the situations in which suppliers and retailers serve as leaders respectively. When suppliers are leaders, the ordering lead time  $L^*$  is smaller or equals to that when retailers are leaders. Thus, in the supply chain, suppliers' priority of decision helps reduce the ordering lead time. In the Stackelberg Game, suppliers have the right of priority of decision and the arrival time is:

$\Pi_m(L^*, Q^*) \leq \Pi_m(L_0, Q_0)$ . Compare to the situation in which retailers decide the ordering quantity first and suppliers decide the lead time later, suppliers are more willing to take the initiative to decide the length of the lead time and then leave it to retailers to decide the ordering quantity.

However, are retailers willing to decide the ordering quantity after suppliers' decision of the lead time? Suppose  $\Delta\Pi$  is to express the optimal cost respectively when suppliers and retailers serve as leaders, there is:

$$\begin{aligned} \Delta\Pi &= \Pi_m(L^*, Q^*) - \Pi_m(L_0, Q_0) \\ &= \left(\frac{D}{Q^*} - \frac{D}{Q_0}\right) C_r + \left(\frac{Q^*}{2} - \frac{Q_0}{2}\right) p_r h \\ &\quad + (\sqrt{L^*} - \sqrt{L_0}) p_r h k \sigma + \frac{R}{Q^*} (1 - \theta) K(L^*) \\ &= (Q^* - Q_0) p_r h + (\sqrt{L^*} - \sqrt{L_0}) p_r h k \end{aligned}, \quad (19)$$

$$\begin{cases} \frac{d\Delta\Pi}{dL_0} = -\frac{1}{2} L_0^{-\frac{1}{2}} p_r h k \sigma < 0 \\ \frac{d\Delta\Pi}{dL_0} = (\sqrt{L^*} - \sqrt{L_0}) p_r h k < 0, L^* \neq L_0 \\ \frac{d\Delta\Pi}{dk} = (\sqrt{L^*} - \sqrt{L_0}) p_r h k \sigma < 0, L^* \neq L_0 \end{cases} \quad (20)$$

With the increase of the standard deviation  $R$  of customer demand, the cost difference  $G$  of retailers of the original security coefficient  $k$  and the arrival time  $L_0$  will decrease for sure, no matter who is the leader. The fluctuation of the demand is more significant to retailers when suppliers take the priority to decide the lead time. Whether willing or not, under the original condition, the decision of suppliers will affect the fluctuation of customer demand, the service of retailers and the arrival time.

### 5 Data analysis

We calculate the aggregate ordering cost  $\Pi_m(L^*, Q^*)$  and  $\Pi_m(L_0, Q_0)$  under two different cost ratio  $\theta$ , and conclude that the cost of ordering lead time decided by suppliers first is smaller than or equals to the cost of ordering quantity decided by retailers. Thus, suppliers wish to take the priority to replenish the inventory.

With the increase of the standard deviation of customer demand faced by retailers, or to say, the uncertainty of the demand, the cost difference varies between the priority of decision of suppliers and retailers. This indicates that when the demand fluctuates, retailers are more willing to let suppliers decide the ordering lead time first. If the fluctuation is small, then it may not favour the retailers in that the cost function  $\Pi_r(L^*, Q^*)$  is bigger than the static Game cost function  $\Pi_r(L_0, Q_0)$  when retailers are exposed to full information.

When the cost sharing ratio  $\theta$  is given, the cost of suppliers' priority of decision is smaller than or equals to that of suppliers' following decision. That is to say, suppliers are prone to take the priority to decide the lead time.

Under different cost ratio  $\theta$ , we calculate the cost of suppliers, the cost of customers in the downstream, the aggregate cost of the supply chain and the cost of the optimal lead time. The results show than when the cost sharing ratio is  $\theta = 0.3$ , the cost of suppliers, the cost of customers in the downstream and the aggregate cost of the supply chain are smaller than that shouldered by retailers for reducing the lead time. This indicates that when the sharing cost is given, leaving suppliers to take the priority of decision and shoulder some cost for reducing the lead time is advantageous both to supplier and retailers. Suppliers can choose a proper ratio to decide the cost of the lead time and the operation.

**6 Conclusion**

This paper considers the cost sharing for reducing the lead time. It studies the decision order of the ordering lead time which is common but overlooked. Under the Stackelberg Game Model, suppliers take the initiative to decide the lead time and retailers, the quantity of goods. This paper analyses the ordering order in which retailers decide the quantity of goods after suppliers decide the lead time. It points out that this way helps to reduce the lead time in the two-stage supply chain as well as the cost of suppliers. If the demand is uncertain, then suppliers’

priority of decision on the lead time is advantageous to themselves, retailers and even to the whole supply chain. This paper provides a new idea to supply chain management.

**Acknowledgments**

The paper was supported by National Natural Science Foundation of China (61271413) and Academic Cultivation Project of Key Laboratory of Fluid and Power Machinery Engineering, Xihua University (Grant No. SBZDPY-11-10).

**References**

[1] Perry M, Sohal Amrik S 2001 Effective quick response practices in a supply chain partnership *International Journal of Operations & Production Management* **21**(5) 840-54

[2] Ben-Daya M, Hariga M 2003 Lead-time reduction in a stochastic inventory system with learning consideration *International Journal of Production Research* **41**(3) 571-9

[3] Zhang Chun-xiao, xie Jin- xing 2004 Optimization of a Two-level Distribution Inventory System with Random Leadtime and Stochastic Demand Progress *Mathematics In Practice And Theory* **34**(7), 1-8

[4] Pan J C, Yang J S 2002 A study of an integrated inventory with controllable lead time *International Journal of Production Research* **40**(5) 1263-73

[5] Ma Shi - hua, Lin Yong 2002 A Inventory Model Based on Stochastic Lead Time *Computer Integrated Manufacturing Systems* **8**(5) 396-8

[6] Li Yi-na, Xu Xue-jun 2009 Research on Stackelberg Model of Supply Chain Inventory Optimization with Controllable Lead Time and Fuzzy Circumstances *Operations Research And Management Science* **18**(1) 54-9

[7] Li Yi-na, Ye Fei, Xu Xue-jun 2009 Cost allocation model for optimizing supply chain inventory with controllable lead time *Journal Of System Sengineering* **24**(1) 9-17

[8] Tersine R J, Hummingbird E A 1995 Lead-time reduction: the search for competitive advantage *International Journal of Operations & Production Management* **15**(2) 8-18

[9] Song Hua-ming, Ma Shi-hua 2006 Pareto Optimization in Supply Chain under Lead-time Reduction *Control and Decision* **21**(7) 776-9

[10] Mingming Lenga, Mahmut Parlar 2009 Lead-time reduction in a two-level supply chain: Non-cooperative equilibria vs. coordination with a profit-sharing contract *International Journal of Production Economics* **118**(2) 521-44

[11] Wadhwa S, Rao K S, Chan F T S 2005 Flexibility-enabled lead-time reduction in flexible systems *International Journal of Production Research* **43** 3131-62

<b>Authors</b>	
	<p><b>Hao ran Shi, born on November 1, 1973, Da Zhou country, Si Chuan Province, China</b></p> <p><b>Current position, grades:</b> Xihua University, Master  <b>University studies:</b> Logistics Engineering  <b>Professional interests:</b> Operational Research</p>
	<p><b>Kejian Liu, born on June 1, 1974, Hubei, China</b></p> <p><b>Current position, grades:</b> associate professor  <b>University studies:</b> Xihua University  <b>Scientific interest:</b> Computer Network, Database and Information System.</p>

# Effect analysis of speed guidance on traffic demand and driver compliance

**Dashan Chen\***

*School of Transportation Engineering, Huaiyin Institute of Technology, Huai'an, China*

*Received 12 June 2014, www.tsi.lv*

---

## Abstract

Active traffic management is method of increasing capacity and smoothing traffic flows. As one of the most important methods, speed guidance control could be used to improve operational efficiency and reduce accident rates. This paper aims to consider a variety of factors to determine the best traffic management services. Driver obedience for the speed guidance value affected the effect directly. The effect might also lose even play a negative role when the traffic demand reached a certain level. Simulation was carried out depending on different traffic demand and driver compliance rate through abstract urban expressway model. Six kinds of traffic demand under different obedience level were analysed comparatively. Speed guidance control has the positive effect about safety and efficiency when the traffic demand is low or medium. When the traffic demand is high, the effect on safety and efficiency both are negative, and different driver compliance rate affected the effect level to some extent.

*Keywords:* Urban expressway, speed guidance, conflict, compliance rate, effect analysis

---

## 1 Introduction

As the backbone of the city road network, urban expressway shares large proportion of the traffic. In Beijing, major urban expressway accounts for only 8% of the total length, but carries nearly 50% of the traffic flow; only 5% bears more than 35% of the city traffic traveling in Shanghai [1]. Urban expressway plays a vital role in the urban road network, which gradually shifted from the large-scale infrastructure construction to refinement traffic management. With the traffic demand rapid growing, much more congestion and traffic accidents, integrated traffic management should be introduced in the background of coordination between road and vehicle. In recent years, with the development of Vehicles Infrastructure Integration system, taking active traffic management to expressway has become the research hotspots in this subject. The U.S. department of transportation five year its strategic research plan clearly pointed out that dynamic speed harmonization would be one of the main means to optimize traffic flow in urban transportation networks [2]. The ministry of science and technology of china has put intelligent traffic management system as one of the key research areas in the national science and technology plan. As an important part of the active traffic management, the speed guidance control has certain positive significance to improve the expressway capacity, reduce the accident risk and decrease even eliminates traffic congestion.

A traffic flow model of the expressway under the speed guidance condition was derived and an optimal coordination function for ramp metering and main road

speed guidance to maximize the expressway service capacity and minimize the ramp queue delay was built [3-6]. Without entrance ramp limit, combining vehicles state and the road state, and the fuzzy control method for highway was given for highway speed limitation. Incorporation the characteristics of urban expressway with a macroscopic dynamic traffic flow model, a variable speed limit control approach was proposed based on the principle of fuzzy logic. The control for speed limit on expressway was a nonlinear and time variable system, intelligent control method of neural network and neural-fuzzy network were proposed to solve the problem. Speed limited value was determined by the binary tree analysis considering traffic volume, vehicle speed and occupiers [7-9]. Shockwaves result in longer travel times and in sudden large variations in the speeds of the vehicles, which could lead to unsafe and dangerous situations, and variable speed limits could be used to eliminate or at least to reduce the effects of shock waves [10-14].

In fact, there have been several variable speed limits (VSL) applications in Germany, Netherlands, UK and other counties in Europe [15-17]. Some VSL experiment had been carried out and concluded that VSL experiment had a significant effect in reducing speed variation. These studies did not reveal the quantitative effect of VSL for the traffic safety [18-20]. Based on the accident data, Lee established a model, which was, used for comparative analysis the risk changes before and after the VSL controls. With the same model, Allaby analysed the road safety effects under different traffic demand by VSL controls. According to different traffic speed level,

---

\* *Corresponding author* e-mail: shyhits@126.com

Abdel-Aty founded two kinds of dynamic accident prediction models, which used in the simulation analysis for different VSL control strategies. The effects of VSL on traffic flow were studied using fundamental traffic model and shockwave theory [21-27]. In addition, some new models were involved in VSL research [28, 29]. Nevertheless, most of them were designed to smooth traffic flow without considering fuel consumption. Studies have shown that VSL has good control effect on greenhouse gas emissions [30, 31]. VSL control method has essential uncertain characteristics about speed limits value for drivers; they generally choose the driving speed among a range. For example, if the value of VSL is 80km/h, then vehicle speed not exceeding 80km/h is legitimate, may be 30km/h, 50km/h or 60km/h. This will cause the control not precise enough. However, this paper focused on speed guidance control, which provides the driver with a clear speed value.

Although speed guidance control can improve road efficiency and reduce the rate of accident, obedience of the driver directly affects the effect. Simultaneously, when the road traffic demand reaches a certain level, the speed guidance control may lose effect even may play a negative role. Traffic demand and driver obedience were adopted as the sensitivity analysis variable. The result provides theoretical support for speed guidance control in active traffic management.

**2 Macro dynamic traffic flow model**

Control-oriented macro dynamic traffic flow model describes the relationship among traffic flow over space and time even traffic control variables. LW model was proposed by British scholar Lighthill and Whitham in 1955. Against the defects, Payne proposed dynamic relationship between speed and density. The model was further extended considering off-ramp, on-ramp and lane change factors by Papageorgiou. Scholars between domestic and foreign have also proposed various models which are mostly around dynamic relationship between speed and density. Model proposed by Payne and Papageorgiou are widely used in practice. Urban expressway was divided into several segments. Each segment contained up to one entrance and one exit. Speed guidance control was embedded as a new variable in urban expressway. Dynamic speed change was affected not only by regular traffic flow parameters but also interfered by speed guidance control variable. When speed guidance control introduced as a control variable, Payne model was extended. To make the calculated optimal speed guidance value more realistic and accurate, dynamic traffic flow model are described as equation (1) to equation (3):

$$q_i(k) = \rho_i(k)v_i(k)\lambda_i, \tag{1}$$

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L_i\lambda_i}(\rho_{i-1}(k)v_{i-1}(k) - \rho_i(k)v_i(k) + r_i(k) - s_i(k)), \tag{2}$$

$$v_i(k+1) = v_i(k) + \frac{T}{\tau}(u_i(k) - v_i(k)) + \frac{T}{L_i}v_i(k)(v_{i-1}(k) - v_i(k)) - \frac{1}{\tau} \left( \frac{vT}{L_i} \frac{\rho_{i+1}(k) - \rho_i(k)}{\rho_i(k) + \kappa} \right), \tag{3}$$

where  $i$  is road segment index,  $k$  is time interval index,  $T$  is the time step used for data collection,  $L_i$  is length of road segment.  $\tau$  are the model parameters,  $\kappa$  is a time constant,  $\rho_i(k)$  is the anticipation constant and  $v_i(k)$  is model parameters which are equal at a segment.  $\rho_i(k)$  is the traffic density of road segment at time index,  $v_i(k)$  is the mean speed of road segment at time index,  $q_i(k)$  is the metering flow rate of road segment at time index,  $w_0(k)$  is the total off-ramp flow rate of road segment at time index,  $d_0(k)$  is the desired control speed of road segment at time index and  $n$  is the number of lanes of road segments.

The paper determines the origins flow and speed parameters of upstream mainline by the ramp length of queue of upstream cell segment. The models are listed as equation (4) to equation (6):

$$w_0(k+1) = w_0(k) + T(d_0(k) - q_0(k)), \tag{4}$$

$$q_0(k) = \min \left[ d_0(k) + w_0(k)/T, Q_0 \frac{\rho_{\max} - \rho_{\mu,1}(k)}{\rho_{\max} - \rho_{\text{crit},\mu}} \right], \tag{5}$$

$$v = \min \left[ (1 + \alpha)v_{\text{crit},m}(k), v_{\text{free}} \exp \left[ -\frac{1}{\alpha m} \left( \frac{\rho_{\mu,1}(k)}{\rho_{\text{crit},\mu}} \right)^{\alpha m} \right] \right], \tag{6}$$

where  $Q_0$  is the onramp flow capacity,  $\rho_{\max}$  is the maximum density of onramp,  $\rho_{\text{crit},\mu}$  is the critical density of onramp at which the traffic flow becomes unstable,  $\rho_{\mu,1}(k)$  is the density of mainline which segment the onramp linked,  $v_{\text{crit},m}$  is the critical speed of mainline, at which the traffic flow becomes unstable,  $v_{\text{free}}$  is the speed of freely traffic flow,  $\alpha m$  is the maximum number of vehicles stored in onramp and  $n$  is the onramp demand flow.

Based on consideration of traffic safety, the speed change should be smooth temporally and spatially. Therefore, the control speed variation should be less than 10km/h over time and distance interval, the following constraints are adopted in equation (7):

$$\bar{V}_{\min} \leq u_i(k) \leq \bar{V}_{\max}, \tag{7}$$

where  $\bar{V}_{\min}$  and  $\bar{V}_{\max}$  are the minimum and maximum control speed for speed guidance control respectively, taken as 20 km/h and taken as 80 km/h.

The extended macroscopic dynamic traffic flow models of urban expressway under speed guidance control listed as formulas (1) to (7) which were the foundations of sensitivity simulation analysis.

**3 Control strategy and simulation system**

The Speed guidance value has great influence on the effect of traffic operation, domestic and international

research results could be used as a sample. 85% speed value was usually used as the speed limit in some foreign country. The effects of safety and efficiency were determined by the speed dispersion of the vehicles on the road. When the roads are crowded, 85% speed as the speed limits seems not precise. To take full account of the parameter flexibility, dynamic control strategy of speed guidance value was shown in Figure 1.

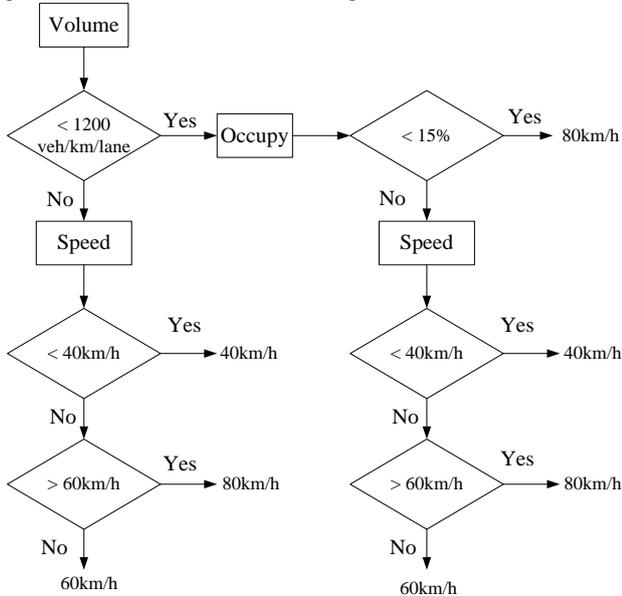


FIGURE 1 Strategy of speed guidance control

In the speed guidance control online simulation system, VISSIM software is used to simulate the real world traffic. The macro dynamic traffic flow model is established in MATLAB software. Data and control strategies were exchange through the API interface among VB.NET, VISSIM and MATLAB. According to the functional orientation, online simulation system was divided into four modules: simulation module, strategy module, interface module and database module. The structures of the simulation system as well as the relationship among the modules were shown in Figure 2.

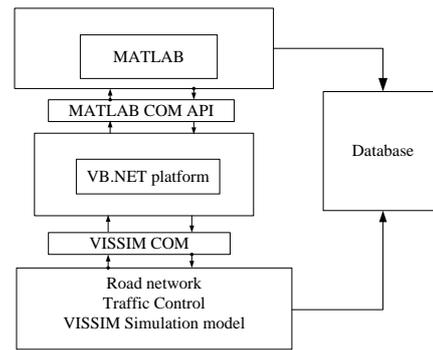


FIGURE 2 Framework of online simulation system

#### 4 Simulation analysis and results

Road safety assessment can be divided into direct and indirect methods. Direct assessment method was based on accident statistics, which was widely used in road traffic management. Though the method was simple to operate, it also had some disadvantages, such as relatively small of the road traffic accidents data, relatively long of the statistical period, great randomness of the accident and other problems.

Traffic conflict technique is the representative method of the indirect traffic safety evaluation, which has obvious advantages, relatively large number of conflict, short period and strong regularity. It was extremely widely applied in the field of road traffic safety. However, traffic conflict observed manually always was arbitrary, taking a lot of manpower and resources.

Therefore, SSAM (Surrogate Safety Assessment Model) which was developed by the United States Federal Highway Administration was used to analyse micro-simulation model output trajectory file. SSAM employed the simulation method to analyse the security. Figure 3 showed the test expressway. The road section used in sensitivity simulation analysis of speed guidance control was shown in Figure 4. It was the inner ring expressway in Shanghai about 5.2 km long. The research road section contained bottleneck caused by reducing lane, which provided a direct target object to analyse the speed guidance control effect. The section was divided into eight parts and each one affected by speed guidance control. The cycle of speed guidance control changed was 1 minute.



FIGURE 3 Test expressway overview

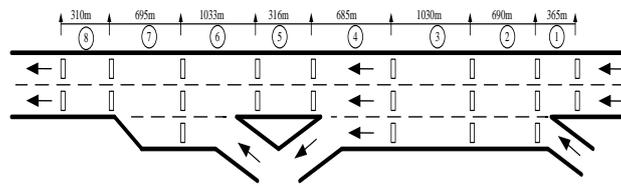


FIGURE 4 Simulation sections of expressway

The macroscopic traffic flow parameters characteristic is an important research content in traffic flow theory. As the backbone road network, traffic flow characteristic parameters have great value for the control and management of urban expressway. Parameters such as free-flow speed, capacity, blocking density, critical density and the critical speed described characteristic of traffic flow. Based on the real-time traffic date, basic parameters of the traffic flow were analysed and extracted, then the average length of the vehicle was statistical analysis to establish the relationship between

occupy and density. Four parameter structure model proposed by Van Aerde in 1995 which implements conversion between macroscopic and microscopic traffic flow model has been widely applied. Data was collecting from Shanghai expressway traffic information collection system. Raw data collection interval period was 20s, and a complete data record includes acquisition time, flow, speed, occupancy, headway, data validation, etc. The error data was removed and repaired. Macroscopic dynamic traffic flow model parameters were shown in Table 1.

TABLE 1 Model parameters

Parameter	$d_0(k)$	$w_0(k)$	$\rho_{max}$	$\rho_{crit,\mu}$	$Q_0$	$T$	$v_{free}$	$v_{crit,m}$	$\alpha$	$am$
Value	2000	10	150	38	1800	10	78	50	0.05	1.87

Driver compliance rate and traffic demand were considered in sensitivity simulation analysis of urban expressway under speed guidance control. Driver compliance rate rose up from 0% to 100% step by 10% increments. 0% represented no speed guidance control and 100% represented the driver fully complied the speed guidance control.

Traffic demand of the mainline entrance rose up from 1000veh/h to 3500veh/h step by 500 veh/h increments. Traffic demand of the upstream entrance ramp rose up from 500veh/h to 1750veh/h step by 250 veh/h increments. Traffic demand of the downstream entrance ramp rose up from 1000veh/h to 3500veh/h step by 500 veh/h increments.

According to the traffic demand of ramp and entrance, modes of traffic demand was divided into six categories as shown in Table 2. Mode A, B, C and D represented that the traffic demand was in low level. Mode E and F corresponded to high traffic demand. Different random seed was selected and 6 times simulation had been carried out. Then the average result was utilized to analyse and evaluate the program.

TABLE 2 Traffic demand pattern

Traffic demand pattern	Mainline traffic demand	Upstream ramp traffic demand	Downstream ramp traffic demand
A	1000veh/h	500veh/h	1000veh/h
B	1500veh/h	750veh/h	1500veh/h
C	2000veh/h	1000veh/h	2000veh/h
D	2500veh/h	1250veh/h	2500veh/h
E	3000veh/h	1500veh/h	3000veh/h
F	3500veh/h	1750veh/h	3500veh/h

Each mode contained eleven different kinds of driver compliance rates; simulation analysis was conducted in VISSIM software. One km part of the expressway was selected as the travel time monitoring segment.

The number of conflicts and monitored link travel time under mode A traffic demand were showed in Figure 5. Conflicts tended to reduce when the driver compliance rate increased. Compared to uncontrolled situation, the conflicts decreased 36% when all the vehicles were under speed guidance control. Travel time decreased significantly at first and when 40% vehicles under control time reached the minimum. Monitored link travel time decreased 3.4% if all the vehicles were under speed guidance control. In the case of traffic patterns A, the results indicated that speed guidance control not only reduce the number of conflicts but also reduce travel time, which demonstrated that safety and efficiency both had been improved.

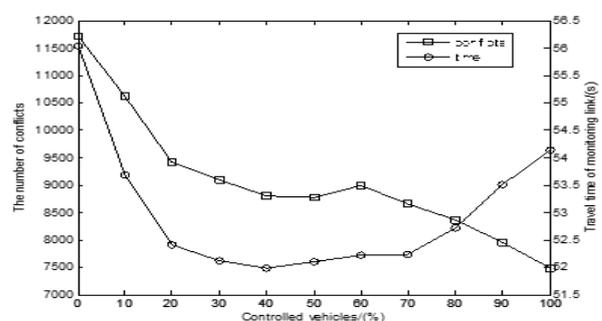


FIGURE 5 Conflicts and travel time under mode A

Figure 6 showed the number of conflicts and monitoring link travel time changes in the case of mode B under different driver compliance rate. Conflicts also

tended to reduce when the driver compliance rate increased. Compared to uncontrolled situation, the conflicts decreased 52% when all the vehicles were under control, which was as same as the mode A. Monitored link travel time decreased 5% if all the vehicles were under speed guidance control. In the case of traffic patterns B, the results also indicated that speed guidance control not only reduce the number of conflicts but also reduce travel time, safety and efficiency both had been improved, the effect was same as the mode.

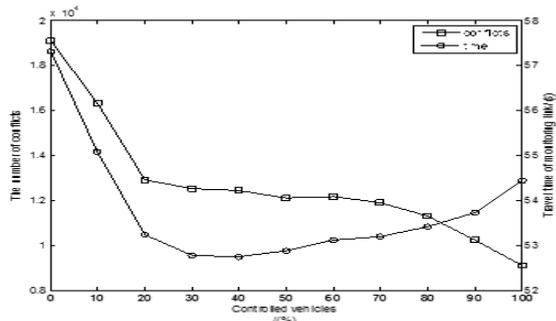


FIGURE 6 Conflicts and travel time under mode B

The number of conflicts and monitored link travel time under mode C were showed in Figure 7. Conflicts were not tended to reduce when the driver compliance rate increased. The number of conflicts slight rose when the driver's obedience was 50%. The travel time changes were different from mode A and B. With the increased of driver compliance rate, the travel time was in the course of a partial repetition. Because of the growing traffic flow, interactions between controlled vehicles and uncontrolled vehicles began to appear. Compared to uncontrolled situation, the conflicts decreased 64% when all the vehicles were under speed guidance control. Monitored link travel time decreased 10% if all the vehicles were under speed guidance control. In the case of traffic patterns C, the effect of reducing conflicts and travel time was more obvious.

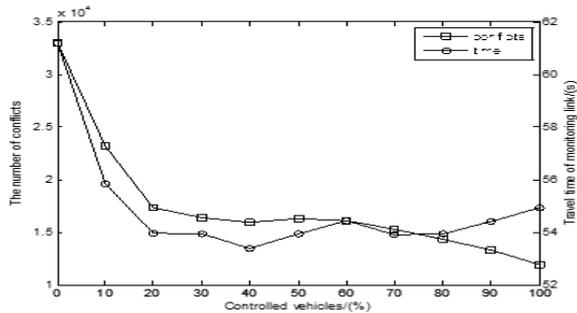


FIGURE 7 Conflicts and travel time under mode C

The number of conflicts and monitored link travel time under mode D were showed in Figure 8. Conflicts and time changed irregularly. Compared to uncontrolled situation, the conflicts decreased 74% when all the vehicles were under speed guidance control. Monitored link travel time decreased 33% when all the vehicles were under speed guidance control. In the case of traffic

patterns D, safety and efficiency both also had been improved.

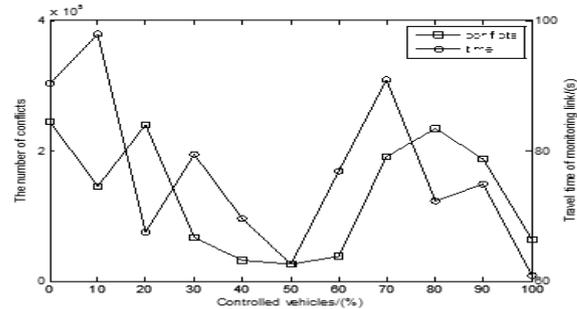


FIGURE 8 Conflicts and travel time under mode D

The number of conflicts and monitored link travel time under mode E were showed in Figure 9. Conflicts changed irregularly, but the travel time tended to rise when the driver compliance rate increased. Compared to uncontrolled situation, the conflicts increased 1.3% when all the vehicles were under control. Monitored link travel time increased 21% when the vehicles were under speed guidance control. Both conflicts and travel time were increased in traffic patterns E, which demonstrated that speed guidance control had negative impact on road safety and efficiency.

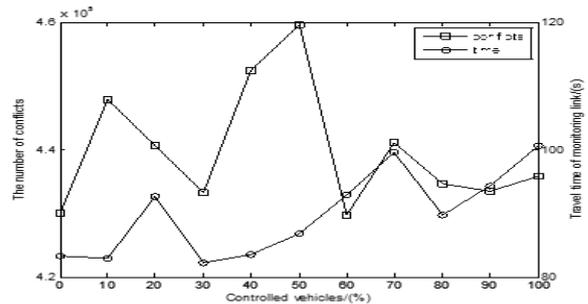


FIGURE 9 Conflicts and travel time under mode E

The number of conflicts and monitored link travel time under mode F were showed in Figure 10. Travel time changed irregularly, but the conflicts tended to gradually rise when the driver compliance rate increased. Compared to uncontrolled situation, the conflicts increased 3.4% when all the vehicles were under control. Monitored link travel time increased 11.4% when the vehicles were under control. Both conflicts and travel time were increased. Speed guidance control had negative impact on road safety and efficiency, which was the same as the mode E.

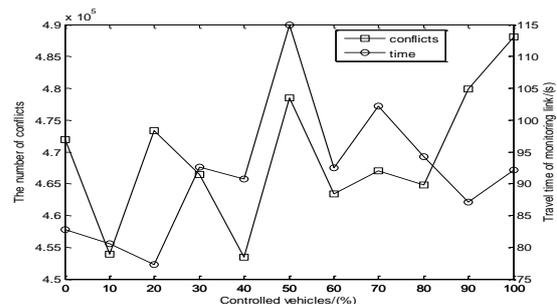


FIGURE 10 Conflicts and travel time under mode F

Comparative statistics analysis of speed guidance control effects were listed in Table 3. It can be seen that in traffic demand pattern A, B, C and D, the number of conflicts and travel time showed a downward trend. Especially in the case of traffic demand mode D both had the largest decline. In those four traffic demand modes, speed guidance control has positive effect in enhancing road traffic safety and improving operational efficiency. By contrast, conflicts and travel times were increased in mode E and F; speed guidance control had negative effects in such two kinds of mode. TRG files generated by simulation results were analysed through SSAM. Conflicts were contrastive processed and the results shown in Figure 11.

TABLE 3 Change of conflicts and travel time under speed guidance control

Traffic demand pattern	Change of the conflicts /%	Change of travel time /%
A	-36.1061	-3.40232
B	-52.1696	-5.00489
C	-63.7891	-10.2006
D	-74.1142	-32.6657
E	1.334282	20.65837
F	3.422591	11.35174

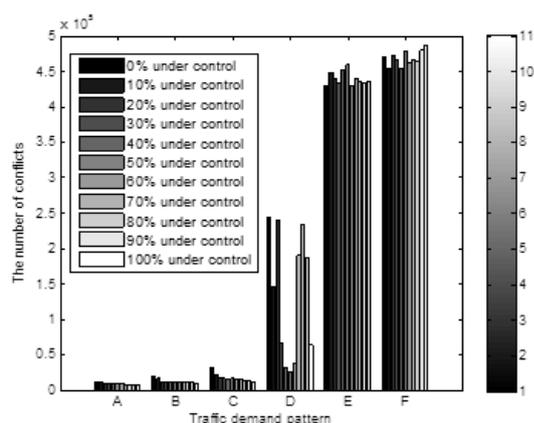


FIGURE 11 Conflicts under different traffic demand

The result of data analysis showed that travel time and the number of conflicts were reduced in low traffic demand when the expressway under speed guidance control. If the traffic demands in high level, both were increased immediately. The effect of speed guidance control were different, in some cases it might had a negative effect. When the speed guidance control applied in practice, detailed analysis about its boundary conditions should be considered to maximize the positive effect.

References

[1] Zhao N L, Yu L, Zhao H, et al 2009 Analysis of Traffic Flow Characteristics on Ring Road Expressways in Beijing Using Floating Car Data and Remote Traffic Microwave Sensor Data *Transportation Research Record* 2124(1) 178-85

5 Conclusions

Simulation analysis platform of urban expressway was established by VB.NET. Data and control strategy were exchange between VISSIM and MATLAB. Simulation was carried out depending on different traffic demand and driver compliance rate through abstract urban expressway model. Six kinds of traffic demand under different obedience level were analysed comparatively. Speed guidance control has the positive effect about safety and efficiency when the traffic demand is low or medium. When the traffic demand is high, the effect on safety and efficiency both are negative. Different driver compliance rate affected the effect level to some extent. Concrete measures should be designed based on the specific circumstances when the speed guidance control involved in the management and control of urban expressway.

Nevertheless, there are remains rooms that needed future endeavours. More factors should be taken into account in the research. The length of the road section was decided considering the reference researches, which pay close attention to the physical properties of the expressway. Dividing the length of the road into different sections will inevitably lead to different results, and then the effect of speed guidance control was diversification. Optimal control model based on length of the road needs to establish in the future. Though efficiency and security had been taking into account in the research, environmental benefits were ignored. As a very important direction of future research, speed guidance control based on ecological driving is required to establish the models and strategies.

Sharp originals (*not transparencies or slides*) should be submitted close to the size expected in the publication. Charges for the processing and printing of colour will be passed on to the author(s) of the paper. As the costs involved are per page, care should be taken in the selection of size and shape so that two or more illustrations may be fitted together on one page. When your paper has been received and logged by the publisher you will be contacted re: prices for colour figures.

Acknowledgments

The author would like to thank the Department of Traffic Engineering, Tongji University and Sino-German Transportation Research Centre for their support in the research work. This research was supported by the science and technology projects of MOHURD (No. 2013-K5-3).

[2] Hadiuzzaman M, Qiu T Z 2013 Cell Transmission Model Based Variable Speed Limit Control for Freeways *Journal of Civil Engineering* 40(1) 46-56  
 [3] Carlson R C, Papamichail I, Papageorgiou M, et al 2010 Optimal Motorway Traffic flow Control Involving Variable Speed Limits and Ramp Metering *Transportation Science* 44(2) 238-53  
 [4] Hegyi A, Schutter B, Hellendoorn H 2005 Model Predictive Control for Optimal Coordination of Ramp Metering and Variable

- Speed Limits *Transportation Research Part C: Emerging Technologies* **13**(1) 185-209
- [5] Frejo J D, Camacho E F 2012 Global Versus Local MPC Algorithms in Freeway Traffic Control with Ramp Metering and Variable Speed Limits *IEEE Transactions on Intelligent Transportation Systems* **13**(4) 1556-65
- [6] Lu X Y, Varaiya P, Horowitz R 2011 Novel Freeway Traffic Control with Variable Speed Limit and Coordinated Ramp Metering *Transportation Research Record* **2229**(1) 55-65
- [7] Allaby P, Hellinga B, Bullock M 2007 Variable Speed Limits: Safety and Operational Impacts of a Candidate Control Strategy for Freeway Applications *IEEE Transactions on Intelligent Transportation Systems* **8**(4) 671-80
- [8] Yang Y, Lu H, Yin Y, et al 2013 Optimization of Variable Speed Limits for Efficient, Safe, and Sustainable Mobility *Transportation Research Record* **2333**(1) 37-45
- [9] Jo Y, Kim Y, Jung I 2012 Variable Speed Limit to Improve Safety Near Traffic Congestion on Urban Freeways *International Journal of Fuzzy Systems* **14**(2) 278-88
- [10] Hegyi A, Schutter B, Hellendoorn J 2005 Optimal Coordination of Variable Speed Limits to Suppress Shock Waves *IEEE Transactions on Intelligent Transportation Systems* **6**(1) 102-12
- [11] Chen D S, Sun J, Li K P 2012 Speed Guidance Predictive Control Model on Urban Expressway *Journal of Traffic and Transportation Engineering* **12**(1) 102-7
- [12] Hadiuzzaman M, Qiu T Z, Lu X Y 2012 Variable Speed Limit Control Design for Relieving Congestion Caused by Active Bottlenecks *Journal of Transportation Engineering* **139**(4) 358-70
- [13] Pu Y, Hu L, Jiang Y S, et al 2012 Variable Speed-Limit Control Before Expressway Mainline Toll Station *Journal of Traffic and Transportation Engineering* **12**(5) 119-26
- [14] Fudala N J, Fontaine M D 2010 Interaction Between System Design and Operations of Variable Speed Limit Systems *In Work Zones, Transportation Research Record* **2169**(1) 1-10
- [15] Buddemeyer J, Young R K, Dorsey-Spitz B 2010 Rural Variable Speed Limit System for Southeast Wyoming *Transportation Research Record* **2189**(1) 37-44
- [16] Sabawat V, Young R K 2013 Control Strategy for Rural Variable Speed Limit Corridor *Transportation Research Record* **2329**(1) 31-44
- [17] Bertini R L, Boice S, Bogenberger K 2006 Dynamics of a Variable Speed Limit System Surrounding a Bottleneck on German Autobahn *Transportation Research Record* **1978**(1) 149-59
- [18] Lee C, Hellinga B, Saccomanno F 2006 Evaluation of Variable Speed Limits to Improve Traffic Safety *Transportation Research Part C* **14**(3) 213-28
- [19] Islam M T, Hadiuzzaman M, Fang J, et al 2013 Assessing Mobility and Safety Impacts of a Variable Speed Limit Control Strategy *Transportation Research Record* **2364**(1) 1-11
- [20] Abdel-Aty M, Cunningham R J, Gayah V 2008 Dynamic Variable Speed Limit Strategies for Real-Time Crash Risk Reduction on Freeways *Transportation Research Record* **2078**(1) 108-16
- [21] Hassan H M, Abdel-Aty M A, Choi K, et al 2012 Driver Behavior and Preferences for Changeable Message Signs and Variable Speed Limits in Reduced Visibility Conditions *Journal of Intelligent Transportation Systems* **16**(3) 132-46
- [22] Lee C, Abdel-Aty M 2008 Testing Effects of Warning Messages and Variable Speed Limits on Driver Behavior Using Driving Simulator *Transportation Research Record* **2069**(1) 55-64
- [23] Cho H, Kim Y 2012 Analysis of Traffic Flow with Variable Speed Limit on Highways *Journal of Civil Engineering* **16**(6) 1048-56
- [24] Papageorgiou M, Kosmatopoulos E, Papamichail I 2008 Effects of Variable Speed Limits on Motorway Traffic Flow *Transportation Research Record* **2047**(1) 37-48
- [25] Heydecker B G, Addison J D 2011 Analysis and Modelling of Traffic Flow under Variable Speed Limits *Transportation Research Part C: Emerging Technologies* **19**(2) 206-17
- [26] Kononov J, Durso C, Reeves D 2012 Relationship between Traffic Density, Speed, and Safety and Its Implications for Setting Variable Speed Limits on Freeways *Transportation Research Record* **2280**(1) 1-9
- [27] Zhang J J, Pang M B, Ren S S 2012 Characteristic Analysis of Traffic Flow in Variable Speed Limit Section of Freeway Based on Cellular Automaton Model *Acta Physica Sinica* **61**(24) 503-13
- [28] Carlson R C, Papamichail I, Papageorgiou M 2012 Local Feedback-Based Mainstream Traffic Flow Control on Motorways Using Variable Speed Limits *IEEE Transactions on Intelligent Transportation Systems* **12**(4) 1261-76
- [29] Wang Y, Ioannou P A 2011 New Model for Variable Speed Limits *Transportation Research Record* **2249**(1) 38-43
- [30] Bel G, Rosell J 2013 Effects of the 80km/h and Variable Speed Limits on Air Pollution in The Metropolitan Area of Barcelona *Transportation Research Part D: Transport and Environment* **23**(1) 90-7
- [31] Liu B, Ghosal D, Chuah C N, et al 2012 Reducing Greenhouse Effects via Fuel Consumption-Aware Variable Speed Limit *IEEE Transactions on Vehicular Technology* **61**(1) 111-22

## Authors



**Dashan Chen, born on April 28, 1983, Huai'an, Jiangsu**

**Current position, grades:** Lecturer with the school transportation engineering, Huaiyin Institute of Technology.

**University studies:** B.S. and M.S. degrees from Chang'an University, Xi'an, China, in 2004 and 2009, respectively. He received the Ph.D. degrees from Tongji University, Shanghai, China, in 2012.

**Scientific interests:** traffic safety and intelligent transportation systems.

# Simulating a crowd with dynamic emotional transmission based on Hidden Markov Model

**Xueling Jiang<sup>1, 2</sup>, Shuijie Qin<sup>1\*</sup>**

<sup>1</sup>Laboratory for Photoelectric Technology and Application, Guizhou University, Guiyang, Guizhou, 550025, China

<sup>2</sup>College of Computer and Information Engineering, Guangxi Teachers Education University, Nanning, Guangxi, 530023, China

Received 12 June 2014, www.tsi.lv

---

## Abstract

Crowd simulation has been widely applied in computer animation and graphics rendering technology. However, the social communication and emotional characteristics are often unrecognized in crowd simulation. In psychology, there are two kinds of emotional factors for humans: the internal one from the individual, and the external one from the neighbours. To this end, in this work, we propose simulate a crowd using affective computing with dynamic emotional transmission. Specifically, we use Hidden Markov Model (HMM) to model the emotions for individuals with consideration of personality, and to capture the internal emotion state transfer. Besides, we abstract a two-layer transmission process to quantify the impact from highly active neighbours. In addition, we conduct some simulation experiments to evaluate our proposed model.

*Keywords:* Crowd simulation, Hidden Markov Model, Affective computing

---

## 1 Introduction

The process of reproducing real world systems based on specific model, and investigating the characteristics of the systems through experiments over the model, is called simulation [1]. With the rapid development of information technology and relevant techniques, simulation has become one of the important tools of improving the planning, designing and operating in various areas. For example, crowd animation simulation [2] is a significant direction in computer animation and graphics rendering technology, and it has been widely used in movies and games. Indeed in real life, the population of cities, the social communication and the crowd activities are all increasingly growing, which post a challenge in public security. Therefore, creating an effective crowd simulation model is necessary.

Existing research on crowd simulation can be grouped into three categories. (1) Macro models focus on the global perspective of the whole systems, such as speed, flow and density. However, macro models fail to describe the movements and activities of individuals in the crowd in details. (2) Micro models study on the individuals in the crowd, such as the characteristics (sex, age, psychology, etc.) as well as the environmental factors. (3) Meso models are defined between above two types. Meso model can describe the whole structure of the crowd, and meanwhile preserve the core data of micro model.

However, above existing efforts mainly investigate the path or movements of the crowd or individuals, without considering the emotional factor of each individual. Therefore, the simulation cannot precisely

represent the process of decision making of the crowd. To this end, we propose to embrace affective computing technique to integrate emotional affects. The intuitive is that not only external environmental factors are important, but also the emotional ones. One reason is that only a small fraction of population is directly affected by the environment, while others are actually affected by other individual in the crowd. Another reason might be that some active people are more likely to affect others.

In this work, we propose to model the crowd simulation using affective computing to take into consideration of emotional factors. Specifically, we focus on the emotional transmission among crowd. Specifically, there are three issues included in emotional transmission. (1) How to model emotions of each individual? (2) How to capture the emotions transmitted by highly active individuals? (3) Will those transmitted emotions be accepted, given various personalities?

Generally, the emotional transmission is considered two-way in this work: (1) emotional state transfer among individuals, represented by the HMM model; and (2) emotional impact from other highly active individuals, simulated by the two-layer transfer process. Specifically, we propose the following solutions to solve the questions above. First, we try to model emotions using Hidden Markov Model (HMM) to capture the dynamic features of affective crowd. Second, we simulate emotional transmission as a two-layer transfer process, and quantify the amount of emotion transmission. Third, to differentiate the influences on individuals, we modify our emotion model by considering the personality of each individual.

---

\* *Corresponding author* e-mail: shuijie@gmail.com

The rest of this paper is organized as follows. In Section 2 we provide some related work. Section 3 presents the individual affective model based on HMM, and Section 4 proposes the two-layer emotional transmission model. Simulation experiments are conducted in Section 5. Finally, the paper is concluded in Section 6.

## 2 Related work

Related work can be categorized into three groups: modelling the virtual environment of the crowd, simulating the crowd behaviour and synthetic crowd movement. Now we provide some literature review as follows.

There are basically two methods in virtual environment modelling is applied: geometrical and non-geometrical methods. The common spatial data structures for geometrical models include: bounding volume hierarchy [3-4], trees with binary space partitioning [5-8], and octree [9-10]. Non-geometrical method is typically employed in two fields: computer animation and robots. In the field of computer animation, Lamarche et al. proposed a topological structuring of the geometric environment to allow fast path finding as well as an efficient reactive navigation algorithm for virtual humans evolving inside a crowd [11]. Noser et al. employed synthetic vision for navigation by a digital actor [12]. Sung et al. proposed to combine the behaviour model of virtual agents with the environment model [13]. In the robots area, typical models include: Grid map proposed by Elfes et al [14], topological method [15], and Featuremap [16], etc.

Many efforts have been made in simulating behaviour of crowd. Reynolds et al pioneered in this field by simulating the behaviour of flocks and herds using a particle system [17]. Later agent theory has been applied to explore the crowd behaviour. For example, Alyett et al developed a sequence distribution graph using intelligent agents [18].

The third group of related work is about synthetic crowd movement. Some methods are based on graph structure [19-21], which represents the set of motion segments as a database with graph structure, and displacement mapping is performed to assemble the motion. Another method is procedural motion synthesis [22-23], which is more designed for real-time scenarios. Motion blending is also employed to motion synthesis [24-25], which could continuously simulate the motion routes of virtual human.

Unlike existing works, in this work, we focus on simulate a crowd with dynamic emotional transmission. Specifically, we capture that there are two types of transmission among the crowd. One is emotional state changing between individuals, and the other is emotional transmission among the whole crowd triggered by highly active individuals.

## 3 Individual affective model

In this section, we discuss how to model affection of each individual. A popular method is to construct emotions in a state space with transitions among different states based on Markov state model [26-27]. However, even though the existing model describes the probability of each emotional state, it fails to present the exact emotional changing, which is largely related to personality. To this end, we propose a HMM based emotion model with consideration of personality in this section.

### 3.1 EMOTION AND EMOTION SPACE

Emotions of human beings, such as angry, surprised, neutral, sad, happy, etc., makes up the emotion space, notated as  $S_E = \{S_{E_i} | i=1, 2, \dots, N\}$  where  $S_{E_i}$  is the  $i$ -th emotion state, and  $N$  is the size of basic emotions. Let  $X$  be a random variable, and the probability of is notates as  $P_i$ , and:

$$\sum_{i=1}^N P_i = 1, 0 \leq P_i \leq 1 (i=1, 2, \dots, N). \quad (1)$$

Therefore, the probability space model of emotion space can be represented as:

$$\begin{pmatrix} S_E \\ P \end{pmatrix} = \begin{pmatrix} S_{E_1} & S_{E_2} & \dots & S_{E_N} \\ P_1 & P_2 & \dots & P_N \end{pmatrix}. \quad (2)$$

### 3.2 PERSONALITY AND PERSONLAITY SPACE

Emotional response is greatly dependent on the personality of individuals. For example, a person with an optimistic personality is more likely to be bright, open minded and stepped forward, while a reserved person might be apathetic to everything. From above observations, we percept that with different personalities, the acceptance of emotional contagion originated from highly active people is different.

In the field of psychology, the Big Five personality traits explore five dimensions of personality to describe human personality, and the theory is called the Five Factor Model (FFM) [28]. According to FFM, there are five factors to represent the personality space, i.e., openness, conscientiousness, extraversion, agreeableness, and neuroticism. Therefore, any personality can be represented with above five dimensions:

$$S_P = (p_o, p_c, p_e, p_a, p_n), \quad (3)$$

where  $p_o, p_c, p_e, p_a, p_n \in [-1, 1]$  denote the value of openness, conscientiousness, extraversion, agreeableness, and neuroticism respectively. Note that we define values

within [0,1] as positive personality, and [-1,0] as negative personality.

3.3 HMM MODEL WITH EMOTION AND PERSONALITY

Before constructing a HMM model with emotion and personality, we need to build a unified representation of emotion and personality. The PAD emotional state model [29] is an established method to describe and measure emotional states. Thus, in this work, we employ PAD model to map both emotions and personalities into a unified space.

There are three dimensions in PAD: (1) the Pleasure-Displeasure Scale measures how pleasant an emotion may be, (2) the Arousal-Nonarousal Scale measures the intensity of the emotion, and (3) the Dominance-Submissiveness Scale represents the controlling and dominant nature of the emotion. Every emotion is represented as a point in the 3-dimensional space, notated

as  $e = (e_p, e_a, e_d)$ , where  $e_p, e_a, e_d \in [-1,1]$ . For example, Table 1 shows the PAD representation of eight basic emotions [30]

TABLE 1 Relationship between basic emotion and PAD space

Emotion	Pleasure	Arousal	Dominance	PAD subspace
Fear	-0.64	0.60	-0.43	Anxious
Angry	-0.51	0.59	0.25	Hostile
Happy	0.40	0.20	0.15	Exuberant
Bored	-0.65	-0.62	0.33	Disdainful
Curious	0.22	0.62	-0.01	Dependent
Sleepy	0.20	0.70	-0.44	Docile
Dignified	0.55	0.22	0.61	Exuberant
Elated	0.50	0.42	0.23	Exuberant

Then we consider mapping the 5-dimensional personality based on FFM model into 3-dimensional PAD space. The method we use in this study is inspired by [31]. The mapping from personality space to PAD space is performed as Table 2.

TABLE 2 Mapping from personality space to PAD space

	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Pleasure	0	0	0.21	0.59	0.19
Arousal	0.15	0	0	0.3	-0.57
Dominance	0.25	0.17	0.6	-0.32	0

Therefore, the personality in Equation (3) can be represented as follows:

$$\begin{bmatrix} e_p \\ e_a \\ e_d \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0.21 & 0.59 & 0.19 \\ 0.15 & 0 & 0 & 0.3 & -0.57 \\ 0.25 & 0.17 & 0.60 & -0.32 & 0 \end{bmatrix} \begin{bmatrix} p_o \\ p_c \\ p_e \\ p_a \\ p_n \end{bmatrix} \quad (4)$$

Now we have mapped all emotions and personalities into a 3-dimension coordinate space. Suppose  $E^{(i)} = [e_p^{(i)}, e_a^{(i)}, e_d^{(i)}]$  is the probability distribution of emotional states at time  $i$ , and  $E^{(0)} = [e_p^{(0)}, e_a^{(0)}, e_d^{(0)}]$  is the initialized probability distribution of emotional states.

Suppose  $\lambda = (E, S, \pi, A)$ , where  $X$  is the emotional space represented in PAD model, i.e.,  $\{(e_p, e_a, e_d)\}$ ,  $S$  is the set of external stimulations, and  $\pi$  is the initialized probability distribution of emotional states, i.e.,  $\pi = E^{(0)} = [e_p^{(0)}, e_a^{(0)}, e_d^{(0)}]$ .  $A$  is the transition probability matrix, where element  $a_{ij}$  means the probability of emotional state  $e^{(i)}$  transferred to  $e^{(j)}$ .

The basic equations of Markov model are as follows:

$$e_i^{(k+1)} = \sum_{j=1}^N e_j^{(k)} a_{ji} \quad (5)$$

$$e_i^{(k)} = \sum_{j=1}^N e_j^{(0)} a_{ji}^{(k)} \quad (6)$$

where  $e_i, e_j$  denote the emotion in the 3-dimensional emotion space. The corresponding matrix representation is:

$$E^{(k+1)} = E^{(k)} \bar{A} \quad (7)$$

$$E^{(k)} = E^{(0)} \bar{A}^{(k)} \quad (8)$$

Suppose the limit probability of  $\bar{A}$  can be represented as  $\bar{\pi}^* = [\bar{\pi}_1^* \bar{\pi}_2^* \bar{\pi}_3^*] = [1/3 \ 1/3 \ 1/3]$ . Let

$$\bar{A} = \begin{bmatrix} \frac{L-2}{L} & \frac{1}{L} & \frac{1}{L} \\ \frac{1}{L} & \frac{L-2}{L} & \frac{1}{L} \\ \frac{1}{L} & \frac{1}{L} & \frac{L-2}{L} \end{bmatrix} \quad (9)$$

where  $L = \theta \bar{\pi}_i^* = \frac{\theta}{3} \geq 2$ . Solve the equation

$$|A(\lambda)| = |\lambda I - A| = (\lambda - 1) \left( \lambda - \frac{L-3}{L} \right)^2 = 0, \text{ we get the}$$

following characteristic roots:  $\lambda_1 = 1, \lambda_2 = \lambda_3 = \frac{L-3}{L} = \delta$ .

We can calculate  $\bar{A}$  as follows:

$$a_{ij}^{(k)} = \pi_j + \Delta q_{ij}, \pi_j^* = \frac{\lambda^k A_{ji}(1)}{a_0(1)},$$

$$\Delta q_{ij} = \sum_{\lambda} \frac{1}{(m_i - 1)!} D_{\lambda}^{m_i - 1} \left[ \frac{\lambda^k A_{ji}(\lambda)}{a_i(\lambda)} \right]_{\lambda = \lambda_i} \quad (10)$$

Then we have the results:

$$\Delta q_{ij} = \begin{cases} \frac{2}{3} \delta^k = \frac{2}{3} \left(\frac{L-3}{L}\right)^k, & i = j, \\ -\frac{1}{3} \delta^k = -\frac{1}{3} \left(\frac{L-3}{L}\right)^k, & i \neq j. \end{cases} \quad (11)$$

And

$$\bar{A}^k = \begin{bmatrix} \frac{1}{3} + \frac{2}{3} \delta^k & \frac{1}{3} - \frac{1}{3} \delta^k & \frac{1}{3} - \frac{1}{3} \delta^k \\ \frac{1}{3} - \frac{1}{3} \delta^k & \frac{1}{3} + \frac{2}{3} \delta^k & \frac{1}{3} - \frac{1}{3} \delta^k \\ \frac{1}{3} - \frac{1}{3} \delta^k & \frac{1}{3} - \frac{1}{3} \delta^k & \frac{1}{3} + \frac{2}{3} \delta^k \end{bmatrix} \quad (12)$$

Substitute into Equation (5), we have

$$e_i^{(k)} = \frac{1}{3} + (e_i^{(0)} - \frac{1}{3}) \delta^k, \text{ and}$$

$$e_{i\Delta}^{(k)} = e_{i\Delta}^{(0)} \delta^k, e_{i\Delta}^{(k)} = e_i^{(k)} - \frac{1}{3}, e_{i\Delta}^{(0)} = e_i^{(0)} - \frac{1}{3}, \quad (13)$$

where  $e_{i\Delta}^{(k)}, e_{i\Delta}^{(0)}$  denote the strength of  $i^{th}$  emotion at current and initialized state.

In this section, we have modelled the affection of individuals based on HMM with consideration of both emotion and personality. In next section, we will discuss the transmission among the crowd.

**4 Dynamic emotional transmission model**

In this study, we assume that the emotions could be transmitted among individuals, as proved in existing social psychology efforts [32]. Therefore, we consider the interactive emotional influences among each other by simulating the process as a two-layer transfer process. The basic assumption is that the crowd is dynamic and contagious. Specifically, the emotional transmission is conducted in a two-layer manner. We abstract the emotional transmission as a transmission layer, and the individuals with certain emotional state forms the crowd layer. First, individuals are sampled from the crowd

layer. Then, internal emotional state transmission is performed among the individuals. After that, the changes of emotional state are returned to the crowd layer for updates. The process is illustrated in Figure 1.

We define the perception scope of emotional transmission at time  $t$  as  $d(t)$ . For individual  $b$ , the emotional influence sent by  $b$  can be calculated as  $S = E_b^t / d(t)$ , where  $E_b^t$  denotes the emotional value. If individual  $a$  gets affected, then:

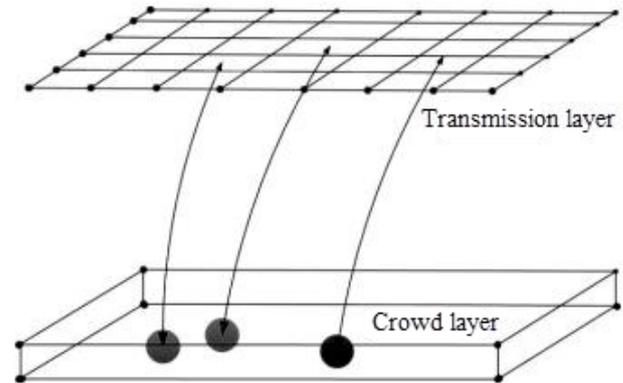


FIGURE 1 Illustration of the two-layer transfer process

$$E_a^{t+1} = \begin{cases} (1 - \alpha_a) E_a^t + \alpha_a S_a^t, & \text{if } a, b \text{ have the} \\ & \text{same sign in the} \\ & \text{pleasure scale,} \\ (1 - \alpha_a) E_a^t - \alpha_a S_a^t, & \text{else.} \end{cases} \quad (14)$$

Therefore, the emotional transmission equation is:

$$\frac{\partial E}{\partial t} = K \frac{\partial^2 E}{\partial^2 U} + E_e, \quad (15)$$

where  $K$  is the damping coefficient, and  $E_e$  is the external influence.

After transmission, the emotion state of individual is updates as follows:

$$E^t = \sum_{i=1}^{\theta} \frac{\sum_{i=1}^{\theta} (|c - c_i|) - |c - c_i|}{(\theta - 1) \sum_{i=1}^{\theta} (|c - c_i|)} M_i^t, \quad (16)$$

where  $\theta$  is the number of individuals within the influence scope,  $c_i$  is the position of neighbour individual, and  $M_i^t$  is emotional state of the transmission layer.

In summary, our proposed method can simulate a crowd with dynamic emotional transmission. Specifically, the affective model of each individual is built upon HMM model with consideration of both emotion and personality. After that, there are two types of emotional transmission: (1) inherent emotion state

changes over time; and (2) impact from others such as highly active neighbours. Correspondingly, our solutions are: (1) probability transfer matrix in HMM, and (2) two-layer transmission model. Figure 2 summarizes the overall process flow of our method.

### 5 Experiment

In this section, we conduct several simulations to evaluate the effectiveness of our proposed model.

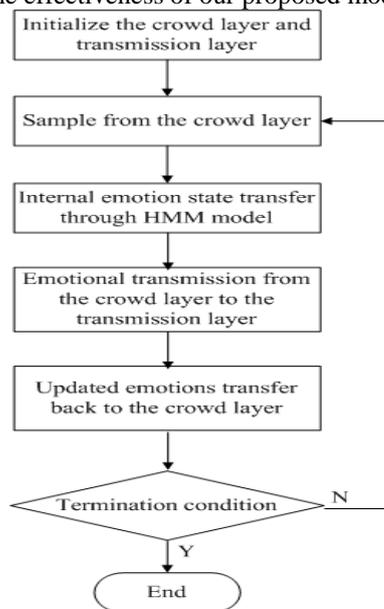


FIGURE 2 Illustration of the two-layer emotional transmission process

First of all, we simulate the influence of emotional transmission upon crowd. As shown in Figure 3, there are four groups of people formed at some time in the left figure, and then after a while, it turns into two groups in the right figure. This transition is caused by the dynamic emotional transmission among the crowd. Figure 4 shows a 3D example of emotional contagion. We can observe that, due the emotional transmission process, people transfer from one place to another.

We believe that, it is the highly active person who triggers the emotional transmission process. Figure 5 provides a crowd evacuation simulation with different percentage of highly active persons. From the left figure with many highly active persons, notated as circles, we can see that the crowd is evenly spread all over the space. However, as shown in the right figure, where the number of active persons is relatively small, the crowd seems to be grouped together towards the highly active centres. Therefore, we can observe that the impact of highly active people is significant and could influence the overall crowd.

Figure 6 illustrates the influence of personality in the process of emotional transmission, where the height denotes the emotional activity of each individual. The upper figures in Figure6 show the dynamic evolution of the crowd without personality consideration, while the lower figures are the changing with personality included.

From those figures, we have the following observations. (1) If no personality considered, the emotion of the crowd is greatly dependent on the highly active person (notated as the green point in upper figures); and the transmission is approximately equal to all the neighbours. (2) When taking personality into consideration for emotional transmission, the overall activity of the crowd is increased, and the emotional state for each individual varies.

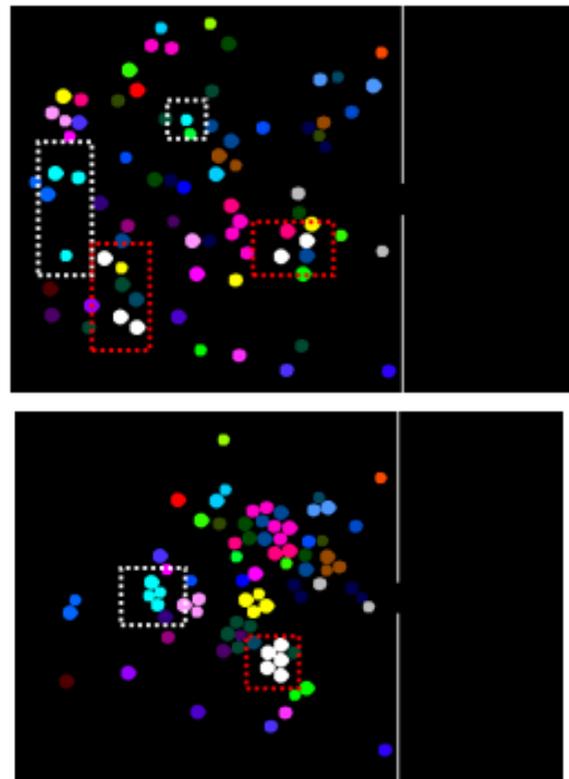


FIGURE 3 Influence of emotional transmission upon crowd

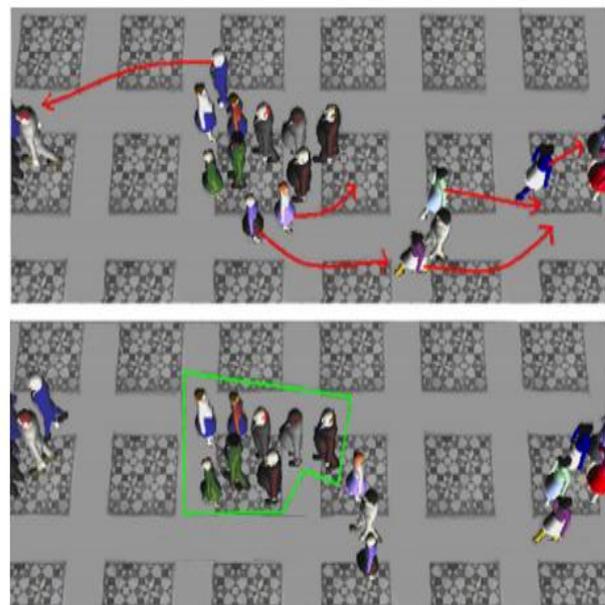


FIGURE 4 3D illustration of emotional transmission among crowd

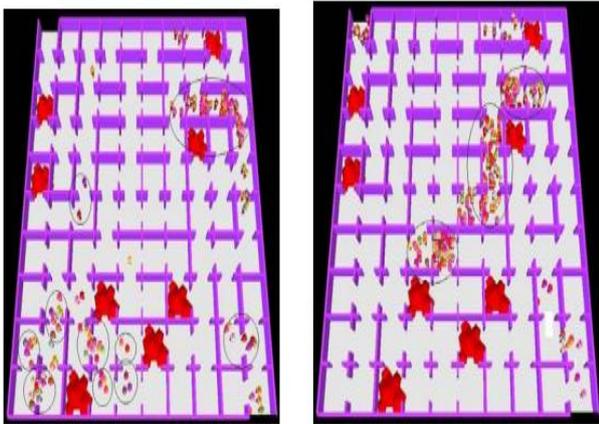


FIGURE 5 Crowd evacuation with high vs. low percentage of highly active persons

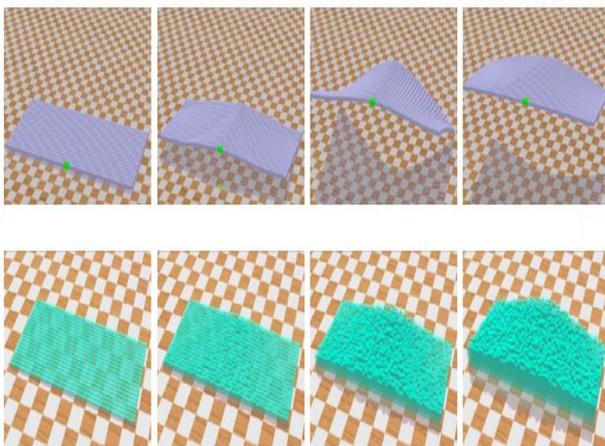


FIGURE 6 Influence of personality in emotional transmission

## References

- [1] Law Averill M, Kelton W D 1991 Simulation modelling and analysis 2 New York: McGraw-Hill
- [2] Thalmann D 2007 *Crowd simulation* John Wiley & Sons, Inc.
- [3] Weyrich M, Drews P 1999 An interactive environment for virtual manufacturing: the virtual workbench *Computers in industry* **38**(1) 5-15
- [4] Klosowski J T, et al 1998 Efficient collision detection using bounding volume hierarchies of k-DOPs *Visualization and Computer Graphics, IEEE Transactions on* **4**(1) 21-36
- [5] Akenine-Möller T, Haines E, Hoffman N 2011 Real-time rendering CRC Press
- [6] Bradshaw G, O'Sullivan C 2002 Sphere-tree construction using dynamic medial axis approximation *Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation. ACM*
- [7] Gordon D, Chen S 1991 Front-to-back display of BSP trees *IEEE Computer Graphics and Applications* **11**(5) 79-85
- [8] James A 1999 *Binary space partitioning for accelerated hidden surface removal and rendering of static environments* Diss. University of East Anglia
- [9] Watt A H, Policarpo F 2001 *3D games: real-time rendering and software technology* 1 Addison-Wesley
- [10] Ulrich T 2000 Loose octrees *Game Programming Gems* **1** 434-42
- [11] Lamarche F, Donikian S 2004 Crowd of virtual humans: a new approach for real time navigation in complex and structured environments *Computer Graphics Forum* **23**(3) Blackwell Publishing, Inc
- [12] Noser H, et al. 1995 Navigation for digital actors based on synthetic vision, memory, and learning *Computers & graphics* **19**(1) 7-19
- [13] Sung M, Gleicher M, Cheney S 2004 Scalable behaviours for crowd simulation *Computer Graphics Forum* **23**(3) Blackwell Publishing, Inc
- [14] Elfes A 1987 Sonar-based real-world mapping and navigation *Robotics and Automation, IEEE Journal of* **3**(3) 249-65
- [15] Kuipers B, Byun Yung-Tai 1991 A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations *Robotics and autonomous systems* **8**(1) 47-63
- [16] Leonard J J, Durrant-Whyte H F 1991 Mobile robot localization by tracking geometric beacons *Robotics and Automation, IEEE Transactions on* **7**(3) 376-82
- [17] Reynolds C W 1987 Flocks, herds and schools: A distributed behavioral model *ACM SIGGRAPH Computer Graphics* **21**(4)
- [18] Luck M, Aylett R 2000 Applying artificial intelligence to virtual reality: Intelligent virtual environments *Applied Artificial Intelligence* **14**(1) 3-32
- [19] Arikan O, Forsyth D A 2002 Interactive motion generation from examples *ACM Transactions on Graphics (TOG)* **21**(3)
- [20] Arikan O, Forsyth D A, O'Brien J F 2003 Motion synthesis from annotations *ACM Transactions on Graphics (TOG)* **22**(3) 402-8
- [21] Gleicher M, et al 2008 Snap-together motion: assembling run-time animations *ACM SIGGRAPH 2008 classes*
- [22] Bruderlin A, Calvert T W 1989 Goal-directed, dynamic animation of human walking *ACM SIGGRAPH Computer Graphics* **23**(3) 233-42
- [23] Sun H C, Metaxas D N 2001 Automating gait generation *Proceedings of the 28th annual conference on Computer graphics and interactive techniques. ACM*
- [24] Park Sang Il, Hyun Joon Shin, Sung Yong Shin 2002 On-line locomotion generation based on motion blending *Proceedings of*

## 6 Conclusion

In this work, we focus on the problem of simulating a crowd with emotion and personality, as well as the emotional transmission among the individuals. Specifically, we first model the emotion and personality of each individual using HMM model. Also, the internal emotional changes of individual is naturally captured by the HMM transfer matrix. Then, we design a two-layer transmission model to simulate the process of emotional transmission among the crowd. Our simulation results prove the effectiveness of our method.

## Acknowledgments

This work was supported by the Scientific Research Fund of Guangxi Education Department of China. (2007), Natural Science Foundation of Guizhou Province of China (Grant Nos. (2009) 2219), National Natural Science Foundation of China (Grant Nos. 61363074), Natural Science Foundation of Guangxi Province of China (Grant Nos.2013GXNSFAA019346), Scientific, Research Fund of Guangxi Education Department of China. (Grant Nos.2013YB148).

- the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation. ACM*
- [25] Pettré J, Laumond J-P, Siméon T 2003 A 2-stages locomotion planner for digital actors *Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation. Eurographics Association*
- [26] Cohen I, Garg A, Huang T S 2000 Emotion recognition from facial expressions using multilevel HMM *Neural information processing systems 2*
- [27] Nogueiras A, et al 2001 Speech emotion recognition using hidden Markov models *INTERSPEECH*
- [28] Digman J M 1990 Personality structure: Emergence of the five-factor model *Annual review of psychology* 41(1) 417-40
- [29] Mehrabian A 1995 *Framework for a comprehensive description and measurement of emotional states* Genetic, social, and general psychology monographs
- [30] Mehrabian A 1996 Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament *Current Psychology* 14(4) 261-92
- [31] Mehrabian A 1996 Analysis of the Big- five Personality Factors in Terms of the PAD Temperament Model *Australian Journal of Psychology* 48(2) 86-92
- [32] Hatfield E, Cacioppo J T 1994 *Emotional contagion* Cambridge university press

Authors	
	<p><b>Xueling Jiang</b>, born on August 9, 1971, Xinjiang province, China</p> <p><b>Current position, grades:</b> doctoral candidates  <b>University studies:</b> Gui Zhou university  <b>Scientific interest:</b> Computer Simulation and Modelling  <b>Publications:</b>            [1] XueLing Jiang, ChaoYun Long, Shuijie Qin, Liping Wang, Jianghui Dong Pedestrian Evacuation Simulation Based on Dynamic Parameter model with Friction. <i>Applied Mechanics and Materials</i> Vols. 543-547 (2014) pp 1876-1879.            [2] XueLing Jiang, ChaoYun Long, Shuijie Qin Solution of Dirac equation with the time-dependent linear potential in non-commutative phase space. <i>Journal of Modern Physics (JMP)</i>            [3] Xiuguang Ge, Xueling Jiang, Lihui Chen and Jianhong Liao. Key Technology of the Denture CAD/CAM System. ISVME2013. <i>Applied Mechanics and Materials</i>. <i>Applied Mechanics and Materials</i> Vols. 494-495 (2014) pp 637-640            [4] Liping Wang, Yi Guo, Xueling Jiang, Jianghui Dong, Long Wang. Study on three-dimensional surgical simulation and face prediction of the individualized maxillofacial soft and hard tissue. <i>Applied Mechanics and Materials</i> Vols. 543-547 (2014) pp 1892-1895            [5] Ying Pan, Tianjiang Wang, Xueling Jiang. Ontology-based Intelligent Information Retrieval System [J]. <i>Journal of Computational Information Systems</i>. 2008, 4(1): 91-96.  <b>Experience:</b> She received her Me.Sc. in Computer Application (2006), now she is a PhD candidate in Mechanical and Automation at GuiZhou University in China. She work at College of Computer and Information Engineering, Guangxi Teachers Education University, Her current research interests include different aspects of Artificial Intelligence and Affective computing.</p>
	<p><b>Shuijie Qin</b>, born in October, 1963, Guangxi province, China</p> <p><b>Current position, grades:</b> professor  <b>University studies:</b> Guizhou university  <b>Scientific interest:</b> MEMS, Computer Simulation and Modelling  <b>Publications:</b>            1. S. J. Qin and Wen J. Li, "Formation Mechanism Analysis to Laser-Induced Splitting Nano Channels in Quartz Cubes", Accepted by <i>Acta Mechanica Sinica</i> (2003).            2. S. J. Qin and Wen J. Li, "Fabrication of Submicron Channels in Quartz Cubes Using Nd:YAG Laser", <i>International Journal of Non-linear Science and Simulation</i>, Vol.3, NOS.3-4, 2002, pp763-768.            3. S. J. Qin and Wen J. Li, "Fabrication of nano channels using laser-induced substrate splitting", <i>Proc. of, IEEE-Nano 2001, Hawaii, USA, October 2001</i>, pp.233-237.            4. S. J. Qin and Wen J, "Fabrication of Micro Channels Using Laser-induced Plasma Ablation of Quartz With Q-Switched Nd:YAG laser", <i>Piezoelectrics &amp; Acoustooptics</i>, Vol.23, 5(2001), pp.139-143.(National conference on Micro and Nano Systems, Chongqing, China, October26-28, 2001).            5. S. J. Qin and Wen J. Li, "Micromachining of complex channel systems in 3D quartz substrates using Q-switched Nd:YAG laser", <i>Applied Physics A, Materials Science &amp; Processing</i>, Vol.74 (2002) 6, 773-777.            6. S. J. Qin and Wen J. Li, "Process characterization of fabricating 3D micro channel systems by laser micromachining", <i>Sensors and Actuators A: Physical</i>, Vol. 97-98, (2002), pp. 749-757.  <b>Experience:</b> She received her M.Sc. in Optical instruments (1989) and PhD in MEMS (2002) from Chinese University Hong Kong CN.</p>

# Uncertain demand of farming-enterprise supply chain coordination based on the option contract

Xiangyang Ren\*, Yu Ren, Qingmei Li

*School of Economics and Management, Hebei University of Engineering, Guangming South Str.199, 056038 Handan, P.R.China*

*Received 12 June 2014, www.tsi.lv*

---

## Abstract

The paper sets up a single-cycle and two-level supply chain model of single rural cooperative and multiple retail enterprises, based on subsidies of option price to combined contract, with the profits maximization of the whole farming-enterprises' supply chain and each member as the goal. It gives the best order quantity and the profits of the whole supply chain and each member under two circumstances, centralized decision and decentralized decision. It also obtains conditions to achieve supply chain coordination. Finally, the paper verifies the model through an example.

*Keywords:* Farming-enterprise supply chain, Option contract, Price subsidy contract, Supply chain coordination

---

## 1 Introduction

In recent years, agricultural industrialization operation mode revealed gradually, farming-enterprise supply chain consisting of rural cooperative and enterprise is the most basic one. This can reduce the circulation, shorten the circulation time, thereby reducing the procurement costs of enterprises, meanwhile to address slow selling problem of the rural cooperative's agricultural product. However, because the industrialization of agriculture is still relatively low, various of systems are not perfect and the contract lack of flexibility and others. That led to the phenomenon of default occur frequently. This will not only damage the interests of the rural cooperative, but also damage the interests of enterprises even coordination is far from discernible.

In order to make the profits of the whole supply chain and two sides of farming-enterprise have increasing and then realize supply chain coordination, many researchers research it from different aspects in different ways. Foreign scholars research it in the following areas, Research on the agricultural products supply chain coordination, such as: Gigler [1] has defined agricultural products' two features in appearance and quality. And have optimization study on the agricultural supply chain using dynamic programming. Cai [2] have coordinated research on agricultural supply chain in the business forms of remote transport's FOB price. Philip [3] consider it can prevent the risk of agriculture by using financial derivative instruments, and seeking for the optimal solution combination of contract farming using time sequence and dynamic multi-period model. Research on options contracts in the supply chain coordination applications, such as: Cachon [4] discussed channels coordination problems of optimal options

contracts and managing wholesale price contract in single suppliers and single retailers system, assuming that suppliers' cost is public information while the demand is the retailer's private information. Wang [5] proposed another type of option contract, in this contract assuming the option execution amount in the second phase can be less than or more than the purchase volume of the first stage of the option, and obtained the optimal strategy from the options buyer's view. Fugate [6] has coordinated multi-stage supply chain related to the market demand forecast constantly updated using quantity flexibility contract, and has discussed the effect from the forecast quality and the flexible level to optimal decision. Research on one to many supply chain coordination, such as: Cachon [7] built a supply chain model containing  $n$  retailers, respectively analysed coordinating role of the revenue-sharing contract for the supply chain when  $n$  retailers compete only on the order quantity, compete only on price, or compete both on order quantity and price, and got that revenue sharing contract can coordinate the supply chain in the first two competition but it cannot make the supply chain coordination in case of the third competition. Domestic scholars research in the following aspects, research on the agricultural products supply chain coordination, such as: Bai [8] introduced the option mechanism in the agricultural products supply chain, established the decision-making mode of each decision-making body under the supply chain coordination that achieve the coordination of agricultural supply chain. Zheng [9] respectively researched that applied option contract to two-level supply chain model of single rural cooperative and single enterprise under asymmetric information and government participation. Research on options contracts in the supply chain coordination applications, such as: Hu

---

\* *Corresponding author* e-mail: boyrenxy@126.com

[10] introduced one-way and two-way option under the premise of market demand is uncertain and analysed the influence of two kinds of options to two-level supply chain flexibility contract consists of single seller and single suppliers, and got the sellers and suppliers optimal decisions under two options contract. Zhu [11] constructed a three-level supply chain model of a single suppliers, single retailers and a single sellers under the premise of demand is uncertain and coordinate the model by applying option repurchase combined contract and gives the optimal decision of supply chain and each member of the supply chain. Research on the coordination of one to many supply chain, such as: Sun [12], constructed a two- level supply chain model of single suppliers and multiple retailers under the premise of demand is uncertain, characterized competition between retailers by the use of Cournot competition, and proposed a linear contract and achieved the coordinate of supply chain by the application of this contract. Wu [13], constructed a two- level supply chain model of single suppliers and multi retailers by the premise of unexpected events lead to changes in market demand and production costs and analysed optimal decision of supply chain and each member of the supply chain under unexpected events, and revenue sharing contract has been improved so that it has anti-burst. This paper research the two- level supply chain model of single rural cooperative and multiple retail enterprise under the premise of demand is uncertain by the application of option price subsidies combined contract on the base of previous studies on agricultural supply chain coordinator, application of options contract coordination and one to many supply chain model.

**2 Model symbols and basic assumptions**

**2.1 MODEL SYMBOL**

$w^d$  : Under the Option Contract, wholesale price of per unit of product of each retail enterprises given by rural cooperative;

$c$  : Rural cooperative production costs of per unit of product;

$c'$  : Rural cooperative processing costs of units remaining products;

$c_o$  : The purchase price of per unit of product options;

$\alpha M_i c_o$  : The interest cost of retail enterprises by buying the option  $i = 1, 2, \dots, n$ ;

$c_e$  : Execution price of per unit of product options;

$p_i$  : Retail enterprise  $i$ 's retail price of unit of product,  $i = 1, 2, \dots, n$ ;

$Q_i$  : Retail enterprise  $i$ 's fixed order quantity,  $i = 1, 2, \dots, n$ ;

$M_i$  : The purchase amount of option of retail enterprise  $i$ ,  $i = 1, 2, \dots, n$ ;

$c'_i$  : Retail enterprise  $i$ 's processing costs of the remaining units of the product,  $i = 1, 2, \dots, n$ ;

$g$  : Retail companies' out of stock losses of per unit of product

$m$  : The price subsidies of remaining products in unit that rural cooperative offered to retail enterprises;

Demand function for the retail enterprises:

$$D_i = D_o - \alpha p_i + \sum_{j \neq i}^n \beta_j p_j, \quad i = 1, 2, \dots, n, \quad \beta_j > 0, \alpha > \beta_j.$$

$D_o$  : The respective maximum demand of n retail enterprises. It is a random variable;

$F(x)$  : Distribution function of the largest market demand  $D_o$ ;

$f(x)$  : Probability density function of the largest market demand  $D_o$ ;

$F(x)$  : Differentiable, continuous incremented, and

$$F(0) = 0, \quad \mu = \int_0^\infty xf(x)dx;$$

$\mu$  : Mean of the largest market demand  $D_o$ ;

$\alpha$  : Consumer's price sensitivity coefficient;

$\beta_j$  : Influence coefficient of market demand from retail enterprises  $j$  to retail enterprises  $i$ ;

If the order quantity of retailer  $i$  is  $Q_i$ , when  $Q_i \leq D_i$ ,

$$Q_i \leq x - \alpha p_i + \sum_{j \neq i}^n \beta_j p_j, \quad x \geq Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j \text{ then the}$$

sales volume of retail enterprises  $i$  is  $Q_i$ ; when  $Q_i > D_i$ ,

$$Q_i > x - \alpha p_i + \sum_{j \neq i}^n \beta_j p_j, \quad x < Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j \text{ then the}$$

sales volume of retail enterprises  $i$  is  $D_i$ .

So the expectations sales volume of retail enterprise  $i$ 's products is:

$$\begin{aligned} S(Q_i) &= E[\min(Q_i, D_i)] \\ &= \int_0^{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} xf(x)dx + \int_{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j}^\infty Q_i f(x)dx \\ &= Q_i - \int_0^{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x)dx \end{aligned}$$

Expectations remaining amount of retail enterprise  $i$ 's products is:

$$\begin{aligned} I(Q_i) &= E[\max(Q_i - D_i, 0)] \\ &= E\left[\max\left(Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j - x, 0\right)\right] \\ &= \int_0^{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} \left(Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j - x\right) f(x)dx \\ &= \int_0^{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x)dx = Q_i - S(Q_i) \end{aligned}$$

Expectations shortage of retail enterprise  $i$ 's products  
 $L(Q_i) = E[\max(D_i - Q_i, 0)]$

$$= E \left[ \max \left( x - \alpha p_i + \sum_{j \neq i}^n \beta_j p_j - Q_i, 0 \right) \right]$$

is:

$$= \int_{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j}^{\infty} \left( x - \alpha p_i + \sum_{j \neq i}^n \beta_j p_j - Q_i \right) f(x) dx$$

$$= \mu - S(Q_i)$$

2.2 THE BASIC ASSUMOTION

Make the following assumptions to the relationship between parameters:

(1)  $c < w^c < p_i$ ,  $c < w^d < p_i$ , rural cooperatives wholesale price is higher than its cost of production to ensure that the profits of rural cooperatives, retail enterprises retail price is higher than its wholesale price that is to ensure it is profitable to retail business.

(2)  $(1 + \tau)c_o + c_e > w^d$ , the purchase price of the option plus the exercise price is more than the wholesale price that making when retail companies get the elasticity

$$\begin{aligned} \Pi^{iD} &= \sum_{i=1}^n \{ p_i S(Q_i + M_i) - c'_i I(Q_i) - c'[I(Q_i + M_i) - I(Q_i)] - gL(Q_i + M_i) - c(Q_i + M_i) \} \\ &= \sum_{i=1}^n [ (p_i + c' + g)S(Q_i + M_i) + (c'_i - c')S(Q_i) - (c + c'_i)Q_i - (c + c')M_i ] \\ &= \sum_{i=1}^n \left[ (p_i + g - c)(Q_i + M_i) - (p_i + g + c') \int_0^{Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx - (c'_i - c') \int_0^{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx - g\mu \right] \end{aligned} \tag{1}$$

Formula (1): The first item represents: sales revenue of agricultural products. The second item represents: the cost of processing of the retail enterprises' remaining agricultural. The third item represents: cost of processing of rural cooperative's remaining agricultural. The fourth represents losses of out of stock. The fifth represents: production costs.

Solving the first derivative to  $Q_i + M_i$  in formula (1):

$$\frac{\partial \Pi^{iD}}{\partial (Q_i + M_i)} = (p_i + g - c) - (p_i + g + c') F \left( Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j \right)$$

Solving the second derivative to  $Q_i + M_i$  in formula

$$(1): \frac{\partial^2 \Pi^{iD}}{\partial (Q_i + M_i)^2} = -(p_i + g + c') f \left( Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j \right) \leq 0$$

Thus  $\Pi^{iD}$  is concave function to  $Q_i + M_i$ , so:

$$(Q_i + M_i)^* = F^{-1} \left( \frac{p_i + g - c}{p_i + g + c'} \right) - \alpha p_i + \sum_{j \neq i}^n \beta_j p_j \tag{2}$$

of supply by options the same time that was encouraged using wholesale prices to order products.

(3)  $(1 + \tau)c_o + m < w^d + c'_i$ , to ensure the validity of the option.

(4)  $m < w^d$ , retail companies cannot benefit from price subsidies directly.

(5) Assuming rural cooperatives and each retail enterprises are risk-neutral and completely rational and both of them are making decision according to their own expectations profit maximization principle.

3 Centralized decision-making farming-enterprises supply chain model under the option contract

Centralized decision-making model is researching the farming-enterprises supply chain as a whole. It is assumed that a centralization unit for unified management to all members of the supply chain and decision makers to arrange the rural cooperative's product and each retail enterprises order with the objective that the whole supply chain to maximize profits.

Under the centralized decision-making, the whole profit expectation of farming-enterprise is:

4 Decentralized decision farming-enterprise supply chain model based on option contracts

In decentralized decision supply chain, retailers often determine the order with their own profit-maximization as a criterion. The order cannot achieve overall optimization of agricultural products supply chain. In order to realize the coordination of agricultural products supply chain. As a leader, rural cooperative must enact terms of options that have incentive effect to the retailers, in order to make the optimal order quantity of retailers' and agricultural products supply chain equal.

Expectations of exercise options:

$$\begin{aligned} & \int_{Q_i}^{Q_i + M_i} (D_i - Q_i) f(D_i) dD_i + \int_{Q_i + M_i}^{+\infty} M_i f(D_i) dD_i \\ &= M_i - \int_{Q_i}^{Q_i + M_i} F(D_i) dD_i = M_i - \int_{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j}^{Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx \\ &= M_i - \left[ \int_0^{Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx - \int_0^{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx \right] \\ &= S(Q_i + M_i) - S(Q_i) \end{aligned}$$

In decentralized decision supply chain, profit model of retailers:

$$\begin{aligned}
 \Pi_{Ri}^D &= p_i S(Q_i + M_i) + mI(Q_i) - w^d Q_i - (1 + \tau)M_i c_o - c_e [S(Q_i + M_i) - S(Q_i)] - gL(Q_i + M_i) - c'_i I(Q_i) \\
 &= (p_i + g - c_e)S(Q_i + M_i) - (m - c_e - c'_i)S(Q_i) + (m - w^d - c'_i)Q_i - (1 + \tau)M_i c_o - g\mu \\
 &= (p_i + g - c_e) \left[ Q_i + M_i - \int_0^{Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx \right] + (m - w^d - c'_i)Q_i - \\
 &(m - c_e - c'_i) \left[ Q_i - \int_0^{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx \right] - (1 + \tau)M_i c_o - g\mu \\
 &= (p_i + g - w^d)Q_i - (p_i + g - c_e) \int_0^{Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx + \\
 &[p_i + g - c_e - (1 + \tau)c_o]M_i + (m - c_e - c'_i) \int_0^{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx - g\mu
 \end{aligned} \tag{3}$$

In (3), the first one: agricultural products sales revenue of retail enterprise *i*; the second one: subsidy of the residual agricultural products from rural cooperative to retail enterprise *i*; the third one: fixed ordering cost of retail enterprise *i*; the fourth one: option purchase cost and interest costs of retail enterprise *i*; the fifth one: option

execution cost of retail enterprise *i*; the sixth one: shortage cost of retail enterprise *i*; the seventh one: processing cost of the residual agricultural products of retail enterprise *i*.

Expectation profit of rural cooperative:

$$\begin{aligned}
 \Pi_N^D &= \sum_{i=1}^n \{w^d Q_i + (1 + \tau)M_i c_o + c_e [S(Q_i + M_i) - S(Q_i)] - c'_i [I(Q_i + M_i) - I(Q_i)] - c(Q_i + M_i) - mI(Q_i)\} \\
 &= \sum_{i=1}^n \{ (w^d - c - m)Q_i + [(1 + \tau)c_o - c' - c]M_i + (c_e + c')S(Q_i + M_i) + (m - c_e - c')S(Q_i) \} \\
 &= \sum_{i=1}^n \{ (w^d - c - m)Q_i + [(1 + \tau)c_o - c' - c]M_i + (c_e + c')S(Q_i + M_i) + (m - c_e - c')S(Q_i) \} \\
 &= \sum_{i=1}^n \left\{ (w^d - c)Q_i + [(1 + \tau)c_o + c_e - c]M_i - (c_e + c') \int_0^{Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx + (c_e + c' - m) \int_0^{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx \right\}
 \end{aligned} \tag{4}$$

In (4), the first one: sales revenue of rural cooperative benefited from retail enterprises fixed order; the second one: sales revenue of put option and interest costs of rural cooperative; the third one: sales revenue of rural cooperative benefited from retail enterprises option execution; the fourth one: processing cost of the residual

agricultural products of rural cooperative; the fifth one: production cost of rural cooperative; the sixth one: subsidy from rural cooperative to retail enterprises.

Calculate the first and the second order partial derivative of  $Q_i$  and  $M_i$  in (3):

$$\begin{aligned}
 \frac{\partial \Pi_{Ri}^D}{\partial Q_i} &= (p_i + g - w^d) - (p_i + g - c_e)F \left( Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j \right) + (m - c_e - c'_i)F \left( Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j \right) \\
 \frac{\partial \Pi_{Ri}^D}{\partial M_i} &= p_i + g - c_e - (1 + \tau)c_o - (p_i + g - c_e)F \left( Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j \right) \\
 \frac{\partial^2 \Pi_{Ri}^D}{\partial Q_i^2} &= -(p_i + g - c_e)f \left( Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j \right) - (c_e + c'_i - m)f \left( Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j \right) \leq 0 \\
 \frac{\partial^2 \Pi_{Ri}^D}{\partial M_i^2} &= -(p_i + g - c_e)f \left( Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j \right) \leq 0
 \end{aligned}$$

It is clearly that  $\Pi_{Ri}^D$  is concave function about  $Q_i$  and  $M_i$ .

$$Q_i^D = F^{-1}\left(\frac{c_e + (1+\tau)c_o - w^d}{c_e + c'_i - m}\right) - \alpha p_i + \sum_{j \neq i}^n \beta_j p_j, \quad (5)$$

$$M_i^D = F^{-1}\left(\frac{p_i + g - c_e - (1+\tau)c_o}{p_i + g - c_e}\right) - F^{-1}\left(\frac{c_e + (1+\tau)c_o - w^d}{c_e + c'_i - m}\right). \quad (6)$$

The whole expectation profit of farming-enterprise supply chain:

$$\Pi^D = \sum_{i=1}^n \Pi_{Ri}^D + \Pi_N^D. \quad (7)$$

**5 The coordination of farming-enterprise supply chain**

To achieve the coordination of farming-enterprise supply chain, rural cooperative must make price policy that has incentive effect to the retailers. So it can make the optimal decision of retail enterprises and the decision of supply chain coordination corresponding.

$$\frac{p_i + g - c}{p_i + g + c'} = \frac{p_i + g - c_e - (1+\tau)c_o}{p_i + g - c_e}. \quad (8)$$

From (8), the coordination of farming-enterprise supply chain must satisfy:

$$c_o = \frac{(p_i + g - c_e)(c + c')}{(1+\tau)(p_i + g + c')}. \quad (9)$$

To make the supply chain coordinating, retail prices from all retail enterprises must be equal (at this time, all the retail enterprises accept the same contract).

To make the supply chain coordinating completely, it must adjust contract parameter to realize arbitrary allocation of supply chain profit between rural cooperative and retail enterprises.

Suppose that:

$$p_i + g - w^d = \lambda_i(p_i + g - c), \quad (10)$$

$$p_i + g - c_e - (1+\tau)c_o = \lambda_i(p_i + g - c), \quad (11)$$

$$p_i + g - c_e = \lambda_i(p_i + g + c'), \quad (12)$$

$$c'_i + c_e - m = \lambda_i(c'_i - c'). \quad (13)$$

It can be known from the hypothesis relationship among parameters:  $0 \leq \lambda_i \leq 1$ .

From (1) and (10) to (13), (3) can be expressed as:

$$\begin{aligned} \Pi_{Ri}^D &= \lambda_i(p_i + g - c)(Q_i + M_i) \\ &- \lambda_i(p_i + g + c') \int_0^{Q_i + M_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx \\ &+ \lambda_i(c'_i - c') \int_0^{Q_i + \alpha p_i - \sum_{j \neq i}^n \beta_j p_j} F(x) dx - g\mu \\ &= \lambda_i \Pi_i'^D - (1 - \lambda_i)g\mu \end{aligned} \quad (14)$$

So the profit of rural cooperative  $\Pi_N^D$  can be expressed as:

$$\Pi_N^D = \sum_{i=1}^n [(1 - \lambda_i)\Pi_i'^D + (1 - \lambda_i)g\mu]. \quad (15)$$

**6 Empirical analysis**

The paper takes a supply chain composed of a rural cooperative and two retail enterprises from Handan city for example to prove. With a lot of surveys, the trading process of the rural cooperative and the two retail enterprises is as follows: Before the selling season, rural cooperative gives the wholesale price of the product  $w^c$  (without option contract),  $w^d$  (with option contract), purchase price of option  $c_o$ , exercise price of option  $c_e$  and subsidy of residual products from retail enterprise  $m$ . The two retail enterprises decide the fixed-order quantity  $Q_i$  and option purchase quantity  $M_i$ , according to profit maximization. In the selling season, retail enterprises decide whether to implement option or not, and the quantity of exercise option. After the selling season, if retail enterprises have residual products, the rural cooperative will give some subsidy.

The parameters and the parameter values, as Table 1 shows:

TABLE 1 The parameters and the parameter values

Parameter	$w^c$	$w^d$	$c$	$c'$	$\tau$	$p_1$	$p_2$
Value	19	20	12	0.5	0.1	32	32
Parameter	$c'_1$	$c'_2$	$g$	$m$	$\alpha$	$\beta_1$	$\beta_2$
Value	2	2	2	10	2	0.5	0.5

Suppose market demand is  $x \sim N(100,30)$  normal distribution,  $c_o = 2$ , solve the model by using Matlab.

(1) In centralized decision farming-enterprise supply chain based on options contracts, from (1) and (2), the optimal order quantity of the retail enterprises and the optimal profit of the farming-enterprise supply chain are:

$$(Q_1 + M_1)^* = 62.57 \quad ; \quad (Q_2 + M_2)^* = 62.57 \quad ; \quad \Pi'^D = 2050.34 .$$

(2) In option contract coordination farming-enterprise supply chain, the optimal fixed order quantity, the optimal option purchase quantity of retail enterprises  $i (i=1,2)$ , the optimal profit of the whole supply chain and the profit of each members can be calculated according to (8) ~ (15):  $Q_1^D = 52.55$  ;  $Q_2^D = 52.55$  ;  $M_1^D = 10.02$  ;  $M_2^D = 10.02$  ;  $\Pi_{R1}^D = 603.08$  ;  $\Pi_{R2}^D = 603.08$  ;  $\Pi_N^D = 422.09 + 422.09 = 844.18$  ;  $\Pi^D = 2050.34$ .

It is clearly that option contract can make farming-enterprise supply chain coordinated.

**7 Conclusions**

The paper established a two-stage supply chain model of one rural cooperative and many retail enterprises, taking into account several factors, such as the product cost of rural cooperative, the processing cost of the remaining

product and shortage loss of retail enterprises. It realized the coordination of supply chain, by studying the model from centralized decision and decentralized decision, with option price subsidy joint contract. Some meaningful conclusions are arrived with the guidance of this conclusion,, the two sides of the supply chain can improve the efficiency of supply chain operation, reduce supply chain cost, realize the sustainable development and cooperation for a long time .They can finally share risks and profits, achieve the ideal level of supply chain management.

**Acknowledgements**

This work benefited from National Natural Science Foundation of China (NSFC) (61240050), Hebei Soft Science Research Program (13455406D) and Social Sciences in Colleges and Universities in Hebei Province in 2013 Annual Fund Project (SY13101).

**References**

[1] Giger J K 2002 On optimization of agricultural chain by dynamic programming *European Journal of Operational Research* **139** 613-25  
 [2] Cai X Q, Chen J, Xiao Y B 2006 Optimal decisions of the manufacturer and distributor in a fresh product supply chain involving long distance transportation *Working Paper, School of Economics and Management, Tsinghua University*  
 [3] Philip G 2004 A selected review of agricultural commodity futures and options markets *European Review of Agricultural Economics* **31**(3) 61-70  
 [4] Cachon G P, M A Laribiere 2002 Supply chain coordination with contracts *University of Pennsylvania Working Paper*  
 [5] Wang Q Z, Tsao D 2006 Supply contract with bidirectional options: The buyer's perspective *International Journal of Production Economics* **41**(1) 30-52  
 [6] Fugate B, Sahin F, Mentzer J T 2006 Supply chain management coordination mechanisms *Journal of Business Logistics* **27**(2) 129-61  
 [7] Cachon G P, Lariviere M A 2005 Supply chain coordination with revenue-sharing contracts: strengths and limitations *Management Science* **51**(1) 30-44  
 [8] Bai S Z, Jiang L H 2008 Research on option contract value in the coordination of agricultural product SC *Logistic Technology* (11) 93-5  
 [9] Zheng T T 2009 Research of the agricultural products supply chain coordination based on the option *Sichuan: Southwest Jiaotong University*  
 [10] Hu B Y, Wang X Y, Peng Q Y 2009 Supply chain flexibility contract and the selection of unilateral & bidirectional option style *Journal of Systems & Management* **4**(7) 52-7  
 [11] Zhu Z, Zhu Y L, Shen H 2012 Coordinating a three-level supply chain with core-enterprise under demand uncertainty *Operations Research and Management Science* **21**(1) 88-95  
 [12] Sun R T, Sun L Y, Li G 2010 Supply chain coordination with competed retailers under uncertain demand *Industrial Engineering and Management* **15**(1) 49-52  
 [13] Wu X J, Yang Z H 2010 Coordinating multiple retailers supply chains when cost and demand are disrupted *Journal of Hunan University (Natural Sciences)* **37**(5) 88-92

Authors	
	<p><b>Xiangyang Ren, born on February 11, 1979, Leting County, Hebei Province, P.R.China</b></p> <p><b>Current position, grades:</b> an associate professor.  <b>University studies:</b> Ph.D in Management Science and Engineering (2009) from China University of Mining and Technology, Beijing, P.R.China.  <b>Scientific interests:</b> different aspects of artificial intelligence and supply chain management. <b>Publications:</b> Enterprise Emergency Logistics Capability Index System Construction Based on Rough Set Theory, ICIC Express Letters, Part B: Applications, 2013; Study on Evaluation of Enterprise Emergency Logistics Capacity Based on Rough Set, Logistics Technology, 2013; Study on Benefit Allotment of Agricultural Produce Supply Chain Based on Game Model, Logistics Technology, 2013.</p>
	<p><b>Yu Ren, born on January 25, 1990, Handan city, Hebei Province, P.R.China</b></p> <p><b>Current position, grades:</b> a graduate student of Logistics Engineering at School of Economics and Management, Hebei University of Engineering, Han Dan, Hebei Province, P.R.China.  <b>University studies:</b> Bachelors Degree in Management (2013) from Hebei University of Engineering, Handan, P.R.China.  <b>Scientific interests:</b> different aspects of logistics and supply chain management.</p>
	<p><b>Qingmei Li, born on May 18, 1984, Shijiazhuang city, Hebei Province, P.R.China</b></p> <p><b>Current position, grades:</b> graduate student of Logistics Engineering at School of Economics and Management, Hebei University of Engineering, Han Dan, Hebei Province, P.R.China.  <b>University studies:</b> Bachelors Degree in Management (2010) from Hebei University of Engineering, Handan, P.R.China.  <b>Scientific interests:</b> different aspects of logistics and supply chain management.</p>

# Traffic driven epidemic spreading in weighted networks with different routes

Fei Shao<sup>1, 2\*</sup>, BingHua Cheng<sup>2</sup>

<sup>1</sup> Jiangsu Information Analysis Engineering Laboratory, Nanjing 211169, China

<sup>2</sup> School of Information Technology, Jinling Institute of Technology, Nanjing 211169, China

Received 12 June 2014, www.tsi.lv

---

## Abstract

How does traffic processes in weighted networks impact on the dynamics of epidemic spreading have attracted increasing attention. It is of great importance to reduce the epidemic spreading velocity and increase the critical epidemic threshold of those real world networks. In this paper, the traffic driven epidemic spreading behaviour in BBV weighted networks is investigated. Formulas to describe the infected density and the epidemic threshold of weighted networks are derived and validated by simulations. The infected density and the epidemic threshold are found to undergo a corresponding change when packets are forwarded through different routes according to the different tuneable parameter. By simulations, the optimal route is explored which is better to control the epidemic spreading.

*Keywords:* weighted network, BBV network, epidemic spreading, traffic flow

---

## 1 Introduction

Since the seminal study of small-world networks by Watts and Strogatz [1] and on scale-free networks by Barabási and Albert [2], the complex networks have attracted the dramatically increasing interest in the past few years. A great deal of real networks can be viewed as complex networks while nodes representing individuals and edges representing the relationships between them. The previous studies on networks have been principally focused on those unweighted networks, edges between nodes are either present or not, represented as binary states. However, lots of social, biological, and communication systems such as the mobile networks [3], the scientific collaboration networks [4], the cellular metabolism [5], the world-wide airport network [6] and the Internet [7] have presented that real networks are specified not only by their topology but also by the weight of the edges. Lots of models [8-15] have been presented to describe those weighted networks, among which the BBV model [12], take into account the coupled dynamical evolution of topology and weights, is the most widely used.

The past decade has witnessed lots of large-scale international epidemics among human, animal, and plant which caused an enormous amount of damage or loss. Those disease outbreaks in real systems can be viewed as epidemic spreading on complex networks while individuals are modelled as nodes and possible contacts between individuals are connected by edges. It is of great significance to inspect how to control the epidemic spreading taking place in complex networks. Many

models have been proposed to investigate the feature of epidemic spreading such as SIS [16, 17], SIR [18], and SI [17, 19, 20] models. In these models, a node is classified in three states: susceptible (which will not infect others but may be infected), infected (which is infective) and recovered (which has recovered from the disease and has immunity). The propagation of epidemic from one node to another is assumed to be a reaction process, that is, an infected node can infect any of its neighbourhood nodes with a fixed probability  $\nu$  at each time step and the recovering rate of infected ones is  $\psi$ . Hence the effective spreading rate  $\lambda$  is defined as  $\lambda = \nu/\psi$ . Without lack of generality,  $\psi$  is set to 1, since it only affects at the definition of the time scale.

Recently, it was found that a susceptible node is more likely to be infected if it receives more packets from infected neighbours [21]. And in many real systems, propagation of the epidemic will not occur unless there is a packet interaction on the network that can physically transfer the epidemic from one node to another. The probability that the epidemic spreads from infected to susceptible nodes mainly depends on the traffic flow. A novel approach, which called traffic driven epidemic spreading is introduced to investigate the outcome of the epidemic spreading process driven by traffic flow [21-24]. However, these studies are focused on unweighted networks, the important issue of how the traffic-driven epidemic spreads in weighted networks has not been considered.

In this paper, we probe the traffic driven epidemic spreading behaviour in BBV weighted networks to obtain the formulas of the infected density and the epidemic

---

\* *Corresponding author* e-mail: shaofei@jit.edu.cn

threshold of weighted networks. While the packets are specified to transfer through the path based on the weight of the edges with a tuneable parameter  $\alpha$ , the optimal route is explored which is better to control the epidemic spreading by simulations.

This paper is organized as follows. In section 2 we describe the models and the formulas, followed by the experimental evaluations on BBV weighted networks and real world network in section 3. The conclusions are given in section 4.

## 2 Models

### 2.1 NETWORK MODEL

In BBV weighted networks, the topological as well as weighted properties can be described by a weighted adjacency matrix  $\mathbf{W}$ , whose elements  $w_{ij}$  denote the weight of the edge between node  $i$  and  $j$ . The generation of BBV weighted networks is based on two coupled mechanisms:[6, 11]

(i) Growth. Starting from an initial number of  $N_0$  nodes fully connected by edges with assigned weight  $w_0$ , a new node will be added at each time step. The newly added node is connected to  $m$  different previously existing nodes with equal weight  $w_0$  for every edge and chooses preferentially nodes with large strength according to the probability  $\prod_{n \rightarrow i} = s_i / \sum_l s_l$ , where  $s_i$  is the node strength described as  $s_i = \sum_j w_{ij}$ .

(ii) Weight dynamics. The weight of each newly added edge is initially set to a given value  $w_0$  which is often set to 1 for simplicity. But the adding of edge connecting to node  $i$  will result in increasing the weight of the other edges linked to node  $i$  which is proportional to the edge weights. If the total increase is  $\delta$  (we will focus on the simplest form:  $\delta_i = \delta$ ), we can get

$$w_{ij} = w_{ij} + \Delta w_{ij} = w_{ij} + \delta * \frac{w_{ij}}{s_i} . \tag{1}$$

This will yield the strength increase of node  $i$  as:

$$s_i = s_i + \delta + w_0 . \tag{2}$$

The degree distribution of BBV network  $P(k) \propto k^{-\gamma_k}$  and the strength distribution  $P(s) \propto s^{-\gamma_s}$  yield scale-free properties with the same exponent [6, 11, 12, 14]:

$$\gamma_k = \gamma_s = \frac{4\delta + 3}{2\delta + 1} = 2 + \frac{1}{2\delta + 1} . \tag{3}$$

### 2.2 TRAFFIC MODEL

Our traffic model can be described as follows:

1) All the nodes can create packets with addresses of destination, receive packets from other nodes, and forward packets to their destinations.

2) At each time step  $t$ , an information packet is generated at every node with probability  $\beta$  with randomly chosen sources and destinations.

3) At each time step  $t$ , each node forwards all packets in its queue one step according to the specified route at the same time. In our model, each node has unbounded packet delivery capability for simplicity which means traffic congestion will not occur.

4) A packet, upon reaching its destination, is removed from the system.

Denote  $P_{i \rightarrow j}$  as the path between node  $i$  and  $j$ , which pass through the nodes sequence  $x_0 (= i), x_1, x_2, \dots, x_{n-1}, x_n (= j)$ , we define

$$F(P_{i \rightarrow j}, \alpha) = \sum_{i=0}^{n-1} w_{ij}^\alpha . \tag{4}$$

In our routing strategy, we specify the route between  $i$  and  $j$  as the one makes  $F(P_{i \rightarrow j}, \alpha)$  minimum under a tunable parameter  $\alpha$ . When  $\alpha$  is 0, the specified route is the same as the traditional the shortest path route [25].

### 2.3 EPIDEMIC MODEL

While investigating the dynamical behaviours in the very early stage of epidemic outbreaks, this case corresponds to the simplified SI [20] model, for which infected nodes remain always infective and spread the infection to susceptible neighbours with spreading rate  $\lambda$ .

With the average density of infected nodes of degree  $k$  defined as  $i_k(t)$ , in BBV weighted networks we have

$$\frac{di_k(t)}{dt} = \lambda * \beta * n * b_k * (1 - i_k(t)) * \Theta(t) . \tag{5}$$

The right-hand side takes into account the probability that a node with  $k$  neighbours belongs to the susceptible class represented by  $(1 - i_k(t))$  and gets the infection via packets travelling from infected nodes. The latter process is determined by the spreading probability  $\lambda$ , the number of packets that a node of degree  $k$  receives at each time step  $\beta * n * b_k$  ( $n$  is the node number and  $b_k$  is the so-called algorithmic betweenness [26]), and the probability  $\Theta(t)$  that a packet travels through a link pointing to an infected node.

Assume that the network is uncorrelated,  $\Theta(t)$  takes the form

$$\Theta(t) = \frac{\sum_k b_k P(k) i_k(t)}{\sum_k b_k P(k)} = \frac{\sum_k b_k P(k) i_k(t)}{\langle b \rangle} . \tag{6}$$

When the epidemic begins spreading, the density of infected nodes  $i_k(t)$  is very small. With the initial condition  $i(t)_{t=0} = i_0$ , we gain

$$i(t) = i_0 \left( \frac{\langle b \rangle^2}{\langle b^2 \rangle} (e^{\tau} - 1) + 1 \right). \tag{7}$$

And the epidemic outbreaks time scale of the BBV networks is

$$\tau = \langle b \rangle / (\lambda * \beta * n * \langle b^2 \rangle). \tag{8}$$

The classical SIR [18] model, which is often used for these in which the infected nodes will get recovered and will not return to the susceptible state again, and thus nodes run stochastically through the cycle susceptible  $\rightarrow$  infected  $\rightarrow$  recovered. It is generally used to study epidemics leading to endemic states with a stationary average density of infected nodes. With the effective spreading rate is defined as  $\lambda$  (which means the recovering rate  $\psi=1$ ), we can obtain

$$\begin{cases} \frac{ds_k(t)}{dt} = -\lambda * \beta * n * b_k * s_k(t) * \Theta(t) \\ \frac{di_k(t)}{dt} = -i_k(t) + \lambda * \beta * n * b_k * s_k(t) * \Theta(t) \\ \frac{dr_k(t)}{dt} = i_k(t) \end{cases} \tag{9}$$

In the second row of formula (9),  $-i_k(t)$  means the infected nodes are recovered with  $\psi=1$  and  $s_k(t)$  means the average density of susceptible nodes of degree  $k$  which is replaced by  $(1-i_k(t))$  in formula (5).  $r_k(t)$  means the average density of recovered nodes of degree  $k$ . The other symbols in formula (9) have the same meaning as in formula (5).

By imposing  $di_k(t)/dt=0$  and  $s_k(t) + i_k(t) + r_k(t) = 1$ , we get

$$i_k(t) = \frac{\lambda * \beta * n * b_k * \Theta}{1 + \lambda * \beta * n * b_k * \Theta}. \tag{10}$$

Substituting formula (6) into formula (10), we get

$$\Theta = \frac{\sum_k \frac{b_k^2 * P(k) * \lambda * \beta * n * \Theta}{1 + b_k * \lambda * \beta * n * \Theta}}{\sum_k b_k * P(k)}. \tag{11}$$

The value  $\Theta = 0$  is always a solution. In order to have a non-zero solution, the condition must be filled.

$$\frac{d}{d\Theta} \left( \frac{\sum_k \frac{b_k^2 * P(k) * \lambda * \beta * n * \Theta}{1 + b_k * \lambda * \beta * n * \Theta}}{\sum_k b_k * P(k)} \right) \Big|_{\Theta=0} \geq 1. \tag{12}$$

Therefore,

$$\frac{d}{d\Theta} \left( \frac{\sum_k \frac{b_k^2 * P(k) * \lambda * \beta * n * \Theta}{1 + b_k * \lambda * \beta * n * \Theta}}{\langle b \rangle} \right) \Big|_{\Theta=0} \geq 1, \tag{13}$$

$$\frac{1}{\langle b \rangle} * \sum_k b_k^2 * P(k) * \lambda * \beta * n \geq 1, \tag{14}$$

$$\frac{1}{\langle b \rangle} * \langle b^2 \rangle * \lambda * \beta * n \geq 1. \tag{15}$$

We can obtain the epidemic threshold of traffic driven SIR epidemic model in BBV weighted networks:

$$\lambda_c = \frac{\langle b \rangle}{\langle b^2 \rangle * \beta * n}. \tag{16}$$

### 3 Simulations and analysis

In figure 1, we plot the infected density  $i(t)$  versus time  $t$  with different parameter  $\alpha$  in a BBV weighted network with  $n=1000$ ,  $\delta=6$ ,  $m=6$  and  $\omega_0=1$ . (For every network, 20 instances are generated and for each instance, we run 20 simulations. The results are the average over all the simulations.)

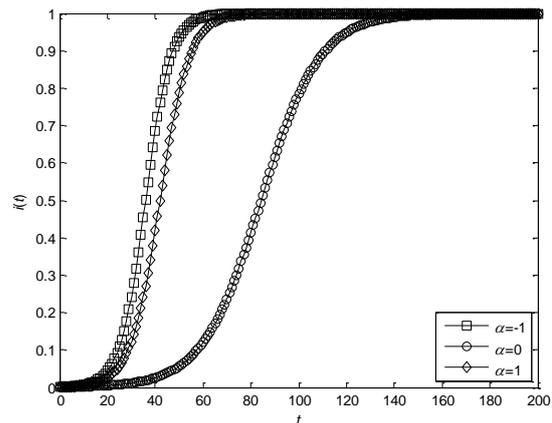


FIGURE 1  $i(t)$  VS  $t$ . BBV network with  $\lambda=0.01$ ,  $\beta=1$ ,  $n=1000$ ,  $\delta=6$ ,  $m=6$  and  $\omega_0=1$

From figure 1, we can discover that the infected density  $i(t)$  varies with the tuneable parameter  $\alpha$ . Figure 1 shows that when the parameter  $t$  is the same, the infected density  $i(t)$  is the smallest in the case that the packets are forwarded through the route path with tuneable parameter

$\alpha=0$ , which means the velocity of epidemic spreading is the lowest.

Then we check the impact of the spreading rate  $\lambda$  and the packet generation rate  $\beta$  on the infected density. Simulation results are shown in figure 2.

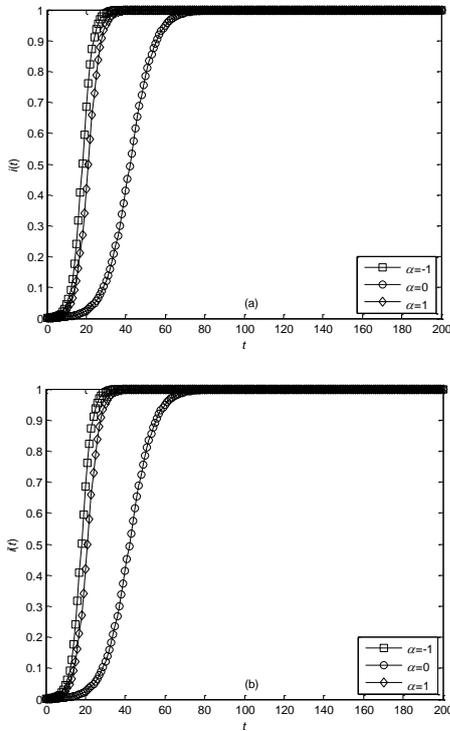


FIGURE 2  $i(t)$  VS  $t$ . BBV network with  $n=1000$ ,  $\delta=6$ ,  $m=6$  and  $\omega_0=1$ . (a)  $\lambda=0.02$ ,  $\beta=1$  (b)  $\lambda=0.01$ ,  $\beta=2$

As is shown in figure 2(a), when the spreading rate  $\lambda$  is doubled, the velocity of epidemic spreading is increased correspondingly which means the spreading velocity is apparently related to the characteristics of epidemic. In figure 2(b), the packet generation rate  $\beta$  is doubled which means there are more packets and more traffic flow in the network, the spreading velocity is greater consequently. Compare figure 2(a) with figure 2(b), we can obtain that in a given network, when the product of the spreading rate and the packet generation rate is fixed, the velocity of epidemic spreading is almost constant. The accuracy of the formula (7) is proved to be correct.

Then we check the impact of the BBV parameter  $\delta$  on the velocity of epidemic spreading. We set  $\delta=3$  and 12 to obtain different simulation results in figure 3.

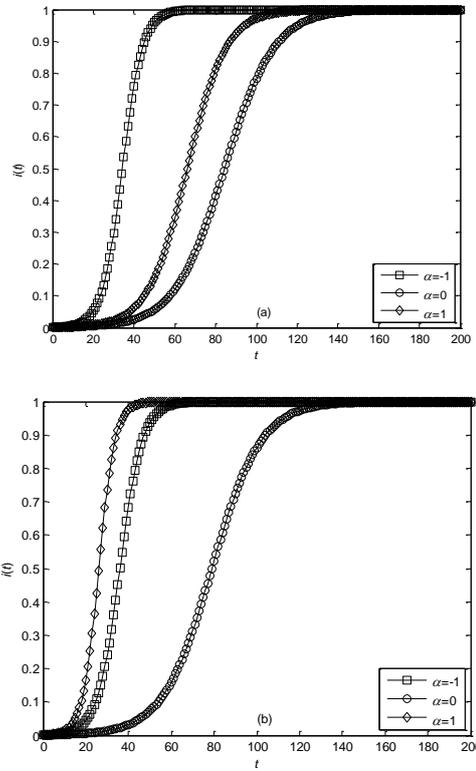


FIGURE 3  $i(t)$  VS  $t$ . BBV network with  $\lambda=0.01$ ,  $\beta=1$ ,  $n=1000$ ,  $m=6$  and  $\omega_0=1$ . (a)  $\delta=3$  (b)  $\delta=12$

From figure 3(a) and 3(b), we can figure that even though the BBV parameter  $\delta$  is changed, the velocity of epidemic spreading with  $\alpha=0$  is the slowest. And the BBV parameter  $\delta$  will have a great influence on the velocity of epidemic spreading.

Then we check the influence of the newly added edge number  $m$  and the total node number  $n$ . Simulation results are shown in figure 4.

From figure 4, we can come to the conclusion that both the newly added edge number  $m$  and the total node number  $n$  have a certain effect on the velocity of epidemic spreading. And the velocity of epidemic spreading with  $\alpha=0$  is also the slowest.

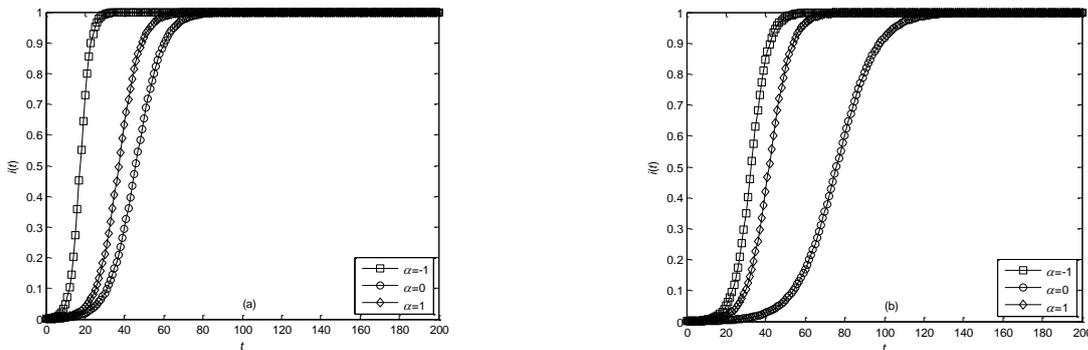


FIGURE 4  $i(t)$  VS  $t$ . BBV network with  $\lambda=0.01$ ,  $\beta=1$ ,  $\delta=6$  and  $\omega_0=1$ . (a)  $n=1000$ ,  $m=3$  (b)  $n=2000$ ,  $m=6$

Then we extend to the SIR model. 20% of nodes are infected in the initial status and the spreading rate is

increased step by step to check the infected density  $i(t)$  of the endemic state.

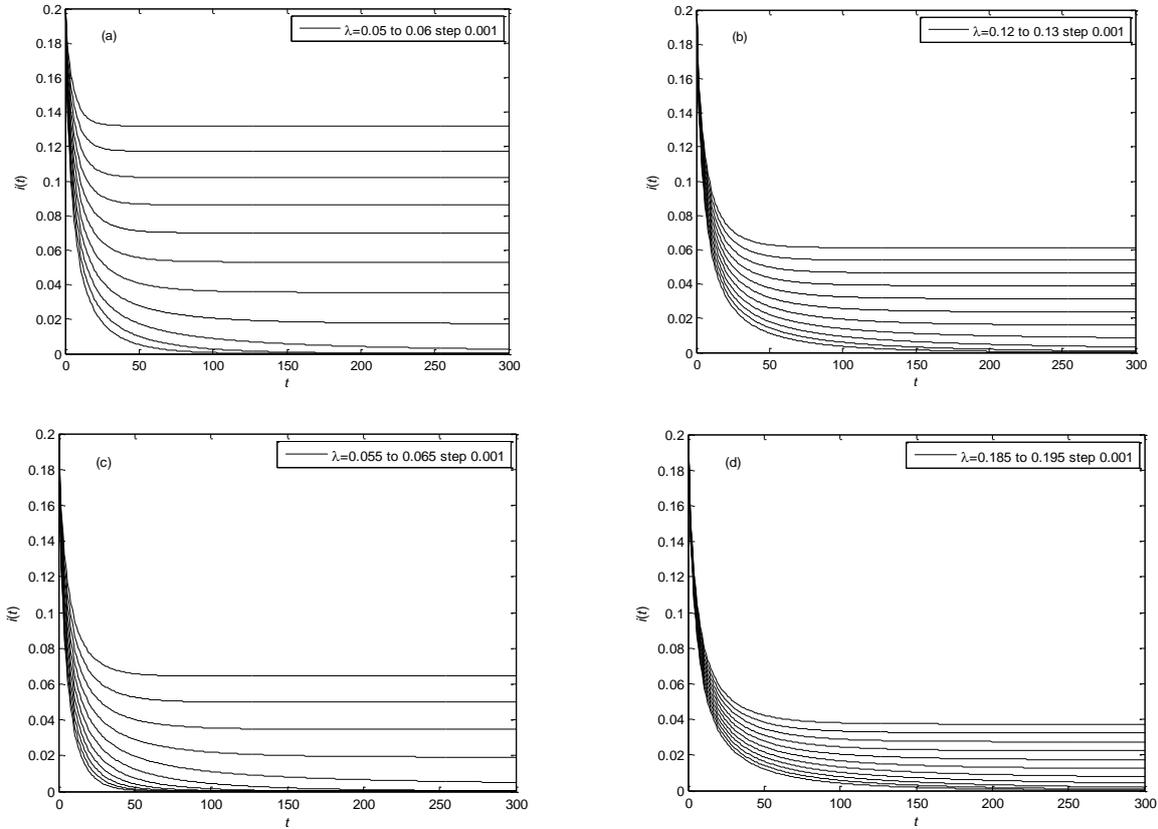


FIGURE 5  $i(t)$  VS  $t$ . BBV network with  $\beta=1, n=1000, \delta=6, m=6$  and  $\omega_0=1$ . The results denote different spreading rate  $\lambda$  (from bottom to top) as labelled in the figures. (a)  $\alpha=-1$  (b)  $\alpha=0$  (c)  $\alpha=1$  (d)  $\alpha=0.2$

Figure 5(a), 5(b), 5(c) and 5(d) exhibit the epidemic threshold  $\lambda_c$  of a BBV weighted network under differ route paths. In figure 5(a), when the spreading rate  $\lambda$  is lower than 0.052, the infected nodes disappear gradually. And while it is up to 0.053, the infections can proliferate on the network. It is in good agreement with analytical finding of the formula (16),  $\lambda_c=0.0521$ .

And the predication of formula (16) for  $\alpha=0, \alpha=1$  and  $\alpha=0.2$  is 0.1220, 0.0608 and 0.1877 consequently. One can see clearly from figure 5(b), 5(c) and 5(d) that the simulation results also agree very well with the analytic results.

To explore the exact optimal value of the tuneable parameter  $\alpha$ , which can produce the largest epidemic threshold, we enlarge the tuneable parameter  $\alpha$  step by step to achieve the corresponding epidemic threshold.

Figure 6 demonstrates that the epidemic threshold varies with the tuneable parameter, that is to say, the different route path. Moreover, the epidemic threshold reaches the peak when  $\alpha$  is around 0.2. It means it is the most effective way to restrain the traffic driven epidemic spreading when the packets are forwarded through the route path specified by  $\alpha=0.2$ .

To find the influence of the packet generation rate  $\beta$  on the epidemic threshold, we change  $\beta$  from 1 to 2 to get the simulation result as presented in figure 7.

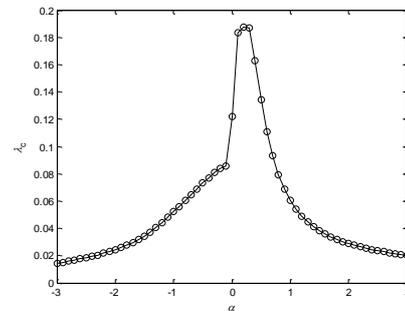


FIGURE 6  $\lambda_c$  VS  $\alpha$ . BBV network with  $\beta=1, n=1000, \delta=6, m=6$  and  $\omega_0=1$

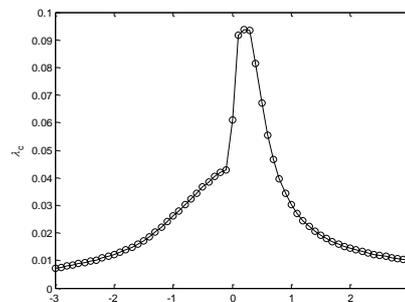


FIGURE 7  $\lambda_c$  VS  $\alpha$ . BBV network with  $\beta=2, n=1000, \delta=6, m=6$  and  $\omega_0=1$

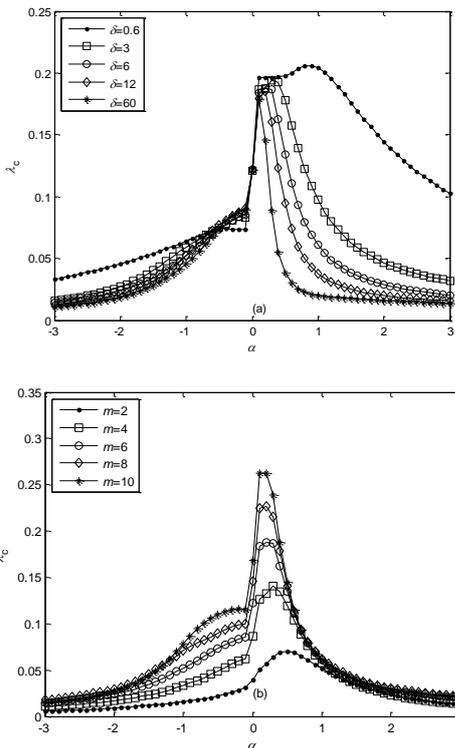


FIGURE 8  $\lambda_c$  VS  $\alpha$ . BBV network with  $\beta=1$ ,  $n=1000$  and  $\omega_0=1$ . (a)  $m=6$  (b)  $\delta=6$

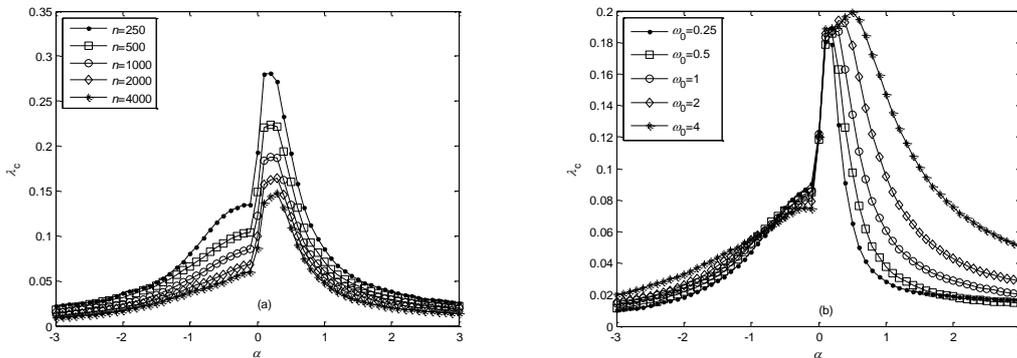


FIGURE 9  $\lambda_c$  VS  $\alpha$ . BBV network with  $\beta=1$ ,  $\delta=6$  and  $m=6$ . (a)  $\omega_0=1$  (b)  $n=1000$

From figure 8(b) and figure 9(a), we can notice that the newly added edge number  $m$  and the total node number  $n$  have a little effect on the impact of tuneable parameter  $\alpha$  on the epidemic threshold. They only affect the absolute value of the epidemic threshold. As shown in figure 9(b), the assigned weight  $w_0$  also affect the tuneable parameter  $\alpha$  to obtain the highest epidemic threshold. As presented in formula (2), when  $w_0$  is considerably high, the minor variation of the parameter  $\delta$  may be passed over. In other words, all edges have almost the same weight. The smaller the parameter  $\delta$  is, the flatter the curve of the epidemic threshold is. Actually, when  $\delta$  is set to 0, it is an unweighted network where the epidemic threshold is fixed no matter how the tuneable parameter  $\alpha$  changes.

Finally, we choose the scientific collaboration network [27] which has a giant component of 5835 nodes

When the packet generation rate is doubled, the epidemic threshold decreases by almost half which also agree well with the formula (16). The epidemic threshold also reaches the maximum value when  $\alpha$  is around 0.2.

Then we investigate how the BBV parameter  $\delta$ , the newly added edge number  $m$ , the total node number  $n$  and the assigned weight  $w_0$  affect the epidemic threshold. The simulation results are depicted in figure 8 and figure 9.

As shown in figure 8(a), the highest epidemic threshold is achieved at different tuneable parameter  $\alpha$  because of different parameter  $\delta$ . The smaller the parameter  $\delta$  is, the flatter the curve of the epidemic threshold is. Actually, when  $\delta$  is set to 0, it is an unweighted network where the epidemic threshold is fixed no matter how the tuneable parameter  $\alpha$  changes. From formula (3) we can discover that both the degree distribution and the strength distribution of BBV network depend on the parameter  $\delta$ . When the parameter  $\delta$  is increased, the distributions become broader which result the maximum value of the epidemic threshold is obtained at different tuneable parameter  $\alpha$ .

to check whether our conclusions are tenable in real world network. Simulation result is shown in figure 10.

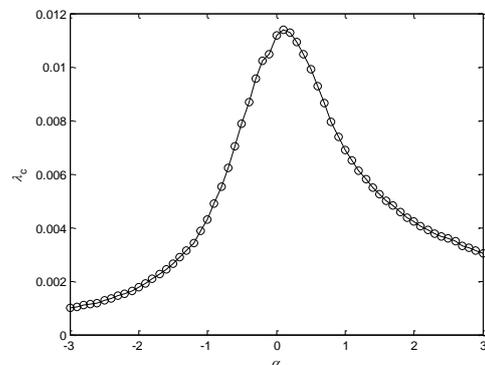


FIGURE 10  $\lambda_c$  VS  $\alpha$ . Real world network

From figure 10 we can discover that the epidemic threshold also varies with the tuneable parameter in the scientific collaboration network which has the maximum epidemic threshold when the tuneable parameter  $\alpha$  is near 0.2. It means our conclusions also work well in the real world network.

#### 4 Conclusions

Considering the traffic driven epidemic spreading behaviour in BBV weighted networks, this paper has deduced the formulas to describe the infected density and the epidemic threshold of BBV weighted networks. The infected density and the epidemic threshold of BBV weighted networks are found to be proportional to the ratio between the first and the second moments of the node betweenness distribution as well as the scale free networks. The validity of the presented formulas is

demonstrated by simulations. The infected density and the epidemic threshold vary accordingly when packets are forwarded through different routes according to the tuneable parameter. In most cases, the epidemic threshold reaches the maximum when the optimal value of the tuneable parameter is 0.2. It is worth mentioning that in some weighted networks the optimal values fluctuate around the mentioned value. At last, we use the scientific collaboration network to show the validity of our conclusions on real world networks.

#### Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (Grant No. 61373136), the Natural Science Foundation of Jiangsu Province, China (Grant No. BK2012082) and sponsored by Qing Lan Project.

#### References

- [1] Watts D J, Strogatz S H 1998 *Nature* **393** 440
- [2] Barabási A L, Albert R 1999 *Science* **286** 509
- [3] Onnela J P, Saramäki J, Hyvönen J, Szabó G, Lazer D, Kaski K, Kertész J, Barabási A L 2007 *Proc. Natl. Acad. Sci. USA* **104** 7332
- [4] Newman M E J 2001 *Phys. Rev. E* **64** 016132
- [5] Almaas E, Kovacs B, Vicsek T, Oltvai Z N, Barabási A L 2004 *Nature* **427** 839
- [6] Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A 2004 *Proc. Natl. Acad. Sci. USA* **101** 3747
- [7] Pastor-Satorras R, Vespignani A 2007 *Evolution and structure of the Internet: A statistical physics approach* (Cambridge: Cambridge University Press)
- [8] Yook S H, Jeong H, Barabási A L, Tu Y 2001 *Phys. Rev. Lett.* **86** 5835
- [9] Zheng D F, Trimmer S, Zheng B, Hui P M 2003 *Phys. Rev. E* **67** 040102
- [10] Dorogovtsev S N, Mendes J F F 2004 arXiv:cond-mat/0408343v2
- [11] Barrat A, Barthélemy M, Vespignani A 2004 *Phys. Rev. E* **70** 066149
- [12] Barrat A, Barthélemy M, Vespignani A 2004 *Phys. Rev. Lett.* **92** 228701
- [13] Antal T, Krapivsky P L 2005 *Phys. Rev. E* **71** 026103
- [14] Barthélemy M, Barrat A, Pastor-Satorras R, Vespignani A 2005 *Physica A* **346** 34
- [15] Wang W X, Wang B H, Hu B, Yan G, Ou Q 2005 *Phys. Rev. Lett.* **94** 188702
- [16] Bailey N T J 1975 *The mathematical theory of infectious diseases and its applications* (New York: Hafner Press)
- [17] Anderson R M, May R M, Anderson B 1992 *Infectious diseases of humans: dynamics and control* (Oxford: Oxford University Press)
- [18] Hethcote H W 2000 *SIAM Review* **42** 599
- [19] Murray J D 2002 *Mathematical biology: I. An introduction* (New York: Springer)
- [20] Barthélemy M, Barrat A, Pastor-Satorras R, Vespignani A 2004 *Phys. Rev. Lett.* **92** 178701
- [21] Meloni S, Arenas A, Moreno Y 2009 *Proc. Natl. Acad. Sci. USA* **106** 16897
- [22] Wang Y Q, Jiang G P 2011 *Acta Phys. Sin.* **60** 080510
- [23] Yang H X, Wang W X, Lai Y C 2012 *CHAOS* **22** 043146
- [24] Yang H X, Wang W X, Lai Y C, Wang B H 2012 *Europhys. Lett.* **98** 68003
- [25] Zhou T 2008 *Physica A* **387** 3025
- [26] Guimerà R, Diaz-Guilera A, Vega-Redondo F, Cabrales A, Arenas A 2002 *Phys. Rev. Lett.* **89** 248701
- [27] Newman M E J 2001 *Proc. Natl. Acad. Sci. USA* **98** 40

#### Authors



**Fei Shao, born in December, 1978, Jiangsu China**

**Current position, grades:** From 2010 to now, he is an associate professor at the School of Information Technology, Jinling Institute of Technology and working for Jiangsu Information Analysis Engineering Laboratory.

**University studies:** Bachelor Degree from Jiangsu University of Science and Technology in 1998, Master degree in Computer Application Technology from Nanjing University of Technology, Jiangsu, China in 2003, and Ph. D Degree in the School of Computer Science & Technology at Nanjing University of Posts & Telecommunications.

**Scientific interests:** information security and complex dynamical networks.

**Publications:** 20 papers, 4 of them are indexed by SCI, 9 of them are indexed by EI.



**Binghua Cheng, born in December, 1977, Jiangsu China**

**Current position, grades:** From 2006 to now, he is a lecturer at the School of Information Technology, Jinling Institute of Technology.

**University studies:** Bachelor Degree from Huaihai Institute of Technology in 2000. Since 2009, he has been pursuing his Master. Degree in Hohai University.

**Scientific interests:** information security and complex dynamical networks.

**Publications:** 3 papers

# Planning and scheduling model of production process in iron and steel enterprise

**Guozhang Jiang<sup>1</sup>, Chongwu Lei<sup>1</sup>, Honghai Liu<sup>2</sup>, Gongfa Li<sup>1, 2\*</sup>**

<sup>1</sup>College of Machinery and Automation, Wuhan University of Science and Technology, Wuhan Postal 430081, Hubei, China

<sup>2</sup>Intelligent Systems and Biomedical Robotics Group, School of Creative Technologies, University of Portsmouth, PO1 2DJ, United Kingdom

Received 1 March 2014, www.tsi.lv

---

## Abstract

According to the analysis of the characteristics of production process in iron and steel enterprise, production planning and scheduling model of multi-stage hybrid procedure and match of production scheduling mode and time were studied. Then the production planning model of production process in iron and steel enterprise was set up. Simulation analysis of production planning and scheduling was carried out using the production procedure of an iron and steel enterprise as an example. The good simulation results show that the established models are correspond to the actual conditions.

*Keywords:* production planning and scheduling, iron and steel enterprise, production process, multi-stage, simulation

---

## 1 Introduction

Iron and steel enterprises need dynamic, timely, ordered and integrated production strategy of production planning and scheduling urgently in order to realize the improvement of product structure and the increase of the production efficiency. It's quite necessary to pay high attention to production planning and scheduling of enterprises.

Steel production process contains a complete process flow, including raw materials, sintering, ore tank, blast furnace tapping, metal adding, LD tapping, refining, continuous casting, slab yard (slab transportation), heating furnace, rolling-coiling and shipping department. The temperature variation with time plays a significant role in each of the steel production process, influencing directly the realization of production planning and scheduling.

## 2 Production scheduling of multi-stage hybrid procedure

The scheduling of multi-stage hybrid procedure, dependent of the major production plan, uses process level scheduling as a method to realize the optimal configuration of production resources and to make sure of the smooth production progress on condition that the delivery date is guaranteed. Thus, to formulate a reasonable and effective job shop scheduling is an important routine for all manufacturing enterprises.

However it is usually difficult to formulate a scheduling which can satisfy each objective due to the complexity, dynamic randomness and multi-targets characters of real production.

The scheduling of production system can be represented by a model which can reflect the actual production status concisely and completely. The flow shop scheduling problem of hybrid procedure uses methods of mathematics to abstract and describe the production activities and resources (materials, devices, working crew and so on) of enterprises, with the help of computers to carry out statistics, operation and analysis of the system to create fine conditions for a timely and feasible production scheduling. Currently, studies on production scheduling mainly focus on two types, discrete ones and flow ones [1] gives investigations about a typical discrete-typed flow shop scheduling model of an enterprise. Few literatures are found on hybrid process because of the actual difficulty in modelling these kinds of scheduling problems. This study aims to build the mathematical model of the production scheduling of multi-stage hybrid procedure, using the real production of steel enterprises as examples.

Assuming that there are  $N$  batches of tasks, each one of which goes through  $M$  work centres and completes  $M$  procedures. Interruption is allowed between one procedure and the following. All tasks have similar process routes. This type of production scheduling is listed as production scheduling of multi-stage hybrid procedure. The following are assumed to be true in this discussion. 1) Every procedure must be completed in the specified work centre; 2) Every procedure is allowed to start only when the previous procedure is finished; 3) No more than one batch of task is allowed to go at any moment and in any work centre; 4) Once one procedure begins, no interruption is allowed. The optimized target is to find a scheduling to minimize the total cost of earliness/tardiness.

---

\* *Corresponding author* e-mail: ligongfa@wust.edu.cn

The problems of the production scheduling of multi-stage hybrid procedure is already an NP problem. Combined with the nonlinearity, randomness and uncertainty of real production process, the large-scale problem can sometimes meet with the so-called combination explosion problems, which makes it difficult to obtain the optimal solution. The essence of production scheduling problems lies in optimization. That is, on the condition of satisfying the constraints, to choose the best or better solution from a series of feasible scheduling plans to assign the production resources and time to the work. Of course the constraints are complicated. It has been an effective way to solve the production scheduling problems by using computer techniques with the development of computer technology. The main means are to search for the optimal solution or the near optimal solution. The current searching approaches include enumeration method, analytical method, random method and so on. Genetic algorithm is a type of random searching method. Genetic algorithm has been widely applied to fields like engineering design, job scheduling, automatic control machine learning, image processing and artificial life for its obvious advantages of flexibility, glottal search performance, universality and robustness. The traditional management way of mankind suffers from the disability of timely production arrangement according to order changes, which has been the main bottleneck in restricting the development of modern enterprises. In order to reduce cost and promote competitiveness, enterprises need to improve production management by means of information integration, process optimization and resource optimization.

Steel production is a multi-process, multi-unit and multi-stage hybrid procedure and also shows combination of discrete and continuous modes. The converter and refining equipments work discretely. The coticaster works continuously within the working life of the tundish to improve work efficiency and reduce set number and production costs. Hot rolling and cold rolling again are discrete production procedure. The entire production process is iron-making, hot metal pretreatment, steel-making, refining, continuous casting and hot rolling. It contains many production processes, each of which requires many kinds of equipments. The mode of the production procedure is multi-stage hybrid. Figure 1 shows the steel-making system model of multi-stage hybrid procedure.

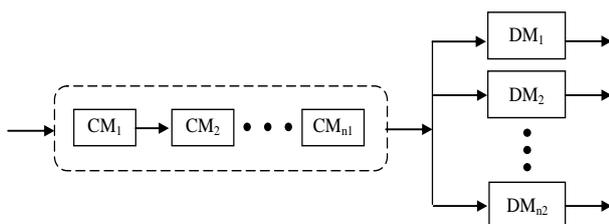


FIGURE 1 The steel-making system model of multi-stage hybrid procedure

### 3 Match of production scheduling mode and time

The continuous steel making and rolling production line consists of one converter, one refining equipment, one coticaster and tandem mill. The working processes are linked end-to-end. The match among time, temperature and logistics is quite tight, which makes the traditional artificial scheduling unsatisfactory. Thus, the integrative continuous production line of shortened process usually adopts automatic control system, which uses computers to process information and scheduling to optimize the production management.

However, experiences are lacked when it comes to the entire steel-making continuous production line scheduling. Thus, it's essential to study and build the scheduling management model of the process of the continuous production line.

#### 3.1 PRODUCTION PROGRESS AND EQUIPMENT

The production process flow of steel-making production lines is as follows. First, the melted iron are transferred from the converter to the steel tank and go through refining in LF furnace. After processing, the steel grades which don't need to be dehydrated are sent direct to the coticaster for casting. The billets completed in the coticaster are cut into one certain dimensions and transferred by billet shifting machine and hot delivery table into walking beam furnace. After heated, the billets are rolled into different kinds of lumber. Under normal circumstances, the billet rolling ability of the tandem mill accords with the billet providing ability of the coticaster to ensure synchronized production. If there is a short failure or pause of rolling, the billets will be sent to the stepping buffer unit. When the rolling mill gets right, the billets waiting in the buffer unit will then be transferred to heating furnace. Then the rolling machine will go full steam ahead. In that the compact rolling ability of the rolling mill is much bigger than the billets providing ability of the coticaster, and that the billets in the buffer units have played a fine part of buffering, the rolling mill will recover to maintain synchronized production with the coticaster. The process of cold billet knockout from the tandem mill amounts to that in the converter steelshop. The transfer of the cold billet into the heating furnace amounts to the production process in the tandem rolling workshop.

The continuous production line of steel-making actually contains several units, i.e. converter, LF furnace, coticaster, heating furnace, hot billet buffering device and tandem mill. Take some steel-making enterprise as an example, the main performance parameters of each unit are as follows.

1) Converter. The nominal capacity is 120t. The production of steel water is 120t. The tap-to-tap cycle is 42min.

2) LF furnace. The nominal capacity is 120t. The production of steel water is 120t. The usual process cycle is 35~55min, dependent on the types of steel. Varying the

power input will change the average ramp up rate, which can be used to justify the actual process cycle of LF furnace and coordinate the match between the converter and tandem mill as a buffer.

3) Conticaster. The pattern is continuous straightening, two-strand. The types of steel include low carbon steel and high strength low alloy steel. The casting section contains two types, 240mm×240mm and 180mm×180mm. The restranding time is 36min/casting. The output time of the billet is 32min.

4) Heating furnace. The waling beam type furnace is adopted. The specification of the billets is 240mm×240mm×6000mm. The production capacity of hot billets is 132t/h. The production capacity of cold billets is 100t/h. The output time relates to the billet section and reheating schedule.

5) Hot billet buffering device. The buffering time is 30~36min.

6) Tandem mill. The time from rolling termination to gate change is about 10min. The time from rolling termination to frame change is about 20min.

7) Lifting time of steel ladle. The ladle is lifted from converter to LF furnace, and then to the rotary table. After casting, the time of deslagging is respectively 6min, 5min and 6min. The performance parameters and operating time are the basis in of carrying out the logistics control of the production line, which should be tested, statistically detected, analyzed and confirmed. Only when the performance parameters, operating time and fluctuation range of every process unit are possessed, can it be possible to establish the realistic scheduling management.

### 3.2 CHOOSING OF PRODUCTION MODE

Steel-making and continuous casting are key links in steel production procedure, especially continuous casting. Because on the one hand, continuous casting is a phase-transition in which steel water turns from liquid water to solid billet. On the other hand, there is an inevitable preparation time before casting resumption after a certain amount of continuous casting, during which time it's impossible to cast steel water and also there is no billet output. In continuous production line, both the steel-making process and steel rolling process must adapt to this characteristic, i.e. to provide qualified steel water according to the pace of continuous casting [2, 3]. Under normal circumstances, the hot continuous casting billets outputted from the conticaster will be rolled 100% in the afterwards rolling mill. If the converter and continuous casting don't match well and interrupt happens, dummy bar has to be installed again, which will not only cost time and make the normal order of production disturbed, but also affect the product quality, cause yield loss and improve the consumption of fire resistant materials. Thus, adjust and control should be carried out with continuous casting as the centre in continuous steel-making production line. The control target starts with the optimal

choice of key parameters of the conticaster. The steel type, section, casting speed, the continuous casting heats and preparation time should be determined first. And then it will be possible to determine the production capacity of the conticaster. Specific requirements for the flow and timing of converter and rolling mill are given to make sure the continuity and stability of the production line. Figure 2 shows the production paths for a hybrid steel-making system.

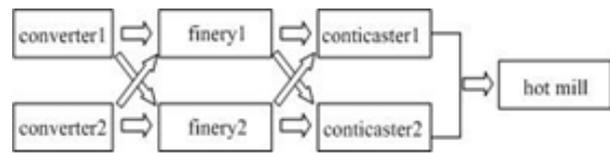


FIGURE 2 The production paths for a hybrid steel-making system

The production modes can be divided into 3 kinds due to the production match type between the conticaster and the converter [4].

#### 3.2.1 Large cycle fit mode

This kind of production mode based on the cycle of continuous casting. If the number of continuous casts is "n". Then, the aim is to make the sum of n casting and the preparation time between two casts (i.e. the cycle of continuous casting) equal to the sum of the refining time of n furnaces of steel water. Because this mode is based on the casting period and that the period is larger than the refining period, so it's called large cycle fit mode.

If the sum of the "i"th steel water processing time in the LF furnace and the lifting time of the ladle from converter to the LF furnace is equal to or larger than the refining period of the converter, it will be impossible for the ladle to reach the LF station on time. The ladle has to wait outside the LF station for a while. The LF station needs to be allocated with several ladle carts. However, the ladle carts will interfere with each other. Also, LF can't play buffering performance between the converter and conticaster. Of course, all of these are unreasonable. Take into account that the buffering ability of LF furnace  $t = t_{converter} - t_{crane}$  reduces along with the shortening of the refining period and the lengthening of LF's shortest processing period. If  $t = t_{converter} - t_{crane} = 0$  happens, the LF furnace will lose its buffering performance. If the amount of planned casts is less and the preparation time is longer, the casting speed allowed will not be able to satisfy the restriction of large cycle fit mode for single tank casting time.

#### 3.2.2 Small cycle fit mode

This production mode based on the refining period of the converter. That is to maintain the consistency between single tank casting time and the refining time. Similarly, the refining period is smaller than the casting cycle. This is why it is called small cycle fit mode. This mode is able

to give the best match of furnace and machines. Also, the processing cycle of the LF furnace remains unchanged, which shows great repeatability and reproducibility. It is easier to be understood and the buffering action of the furnace will perform better. However, the converter could not do continuous production. In one casting, the converter need a timeout according to the preparation time of continuous casting. This timeout can be used for hot fixing or equipment maintenance. Also, the timeout could be used to lower the input power intentionally, or to carry out heat preservation after adding the steel scrap in order to prolong the refining cycle of the steel in the first furnace to make it match with the conticaster.

3.2.3 Large and small cycle combination fit mode

This fit mode is between the large cycle fit mode and the small cycle fit mode. This mode starts with the key parameters of the conticaster. If the best casting speed chosen according to the steel type and steel section, based on which the single tank casting time is chosen, then the single tank casting time can be able to either longer or shorter than the refining time as long as the two times are close. The imbalance between the converter and continuous casting is justified by LF furnace. When the single tank time is shorter than the refining time, it's not necessary for the conticaster to lower intentionally the casting speed to maintain the accordance of single tank casting time and refining time as in small cycle fit mode. What's more, the timeout after finishing one casting is shorter than that in the small cycle fit mode, which can improve the utilization rate of factory equipments and production capacity.

Considering the multi-stage scheduling and match problems, this study will choose the first kind of mode, i.e. large cycle fit mode, according to the analysis on the 3 production modes.

In the large cycle fit mode, the following relations exist.

$$nt_{CC} + t_P = nt_{LD}, \tag{1}$$

where,  $t_{CC}$  is the single casting time,  $t_{LD}$  is the refining time in the converter,  $t_P$  is the preparation time for one casting,  $n$  is the amount of casts.

In this mode, the converter goes continuously with the same refining cycle. But the processing period of the LF furnace in one casting will vary along with the difference of processing furnace, calculated from

$$t_{LFi} = t_{LFO} + (n-i)t_P/n, \tag{2}$$

where  $t_{LFi}$  is the processing cycle of the  $i$ -th steel water in the LF furnace,  $t_{LFO}$  is the minimal processing cycle which can meet the technical requirements,  $i$  is the sequence number of the LF processing and  $i \leq n$ .

Now there are 2 converters, 2 LF furnaces, 2 conticasters and 1 rolling mill, among which, 1 converter

faces 1 LF furnace and 1 conticaster. Assuming that there are  $k$  rolling units, each of which has  $w$  slabs. Also, there are  $m$  casts in each production line and each casts will cast  $n$  furnaces.

The starting time of each station of the  $i$ -th steel water in each production line is

$$T_{LDi} = (i-1) \times t_{LD} \tag{3}$$

$$T_{LFi} = i \times t_{LD} + i \times t_{LD \rightarrow LF} + \sum_{l=0}^{i-1} t_{LFl} \tag{4}$$

$$T_{CCi} = i \times t_{LD} + i \times t_{LD \rightarrow LF} + \sum_{l=1}^i t_{LFl} + i \times t_{LF \rightarrow CC} + T_P + (i-1) \times t_{CC} \tag{5}$$

where  $T_P = \left[ \frac{i-1}{n} \right] \times t_P$  is the preparation time of the cast

( $\left[ \frac{i-1}{n} \right]$  is a rounding function),  $T_{LDi}, T_{LFi}, T_{CCi}$  are the processing starting moment of the  $i$ -th steel water in the converter, the LF furnace and the continuous casting station, respectively,  $t_{LD \rightarrow LF}, t_{LF \rightarrow CC}$  are the transfer time from the converter to the LF furnace and from the LF furnace to the conticaster, respectively.

The starting time of the  $j$ -th slab in every procedure process is calculated from

$$T_{ROj} = T_{CCi} + t_{CC} + j \times t_{CC \rightarrow RO} + (j-1) \times t_{RO} + T_Q, \tag{6}$$

where,  $T_Q = \left[ \frac{j-1}{w} \right] \times t_Q$  is the preparation time needed

for the rolling unit ( $\left[ \frac{j-1}{w} \right]$  is a rounding integration)

where,  $t_Q$  is the preparation time of the rolling unit,  $t_{CC \rightarrow RO}$  is the transfer time from the conticaster to the tandem mill.

The total time of the production process is

$$T = mn \times (t_{LD} + t_{LD \rightarrow LF} + t_{LF \rightarrow CC} + t_{CC}) + \sum_{l=1}^{mn} t_{LFl} + (m-1) \times t_P + kw \times (t_{CC \rightarrow RO} + t_{RO}) + (k-1) \times t_P \tag{7}$$

4 Production scheduling simulation analysis

The production plan and scheduling in steel-making enterprises are transmitted to every flow shop according to the order contracts and inventories. Then, the flow shop determine the refining, continuous casting and fit mode to arrange the production scheduling time according to global optimization principle based on the performance parameters of the units and some other principles. These are the basic control of all kinds of operation targets in the production process. The continuous production flow process of steel begins from

charging to finished products of hot rolling [5]. An realistic production scheduling of the steel enterprises is shown in Figure 3. The scheduling results of the actual working procedure time and production is shown in Figure 4. Based on this, the ideal production scheduling model is built on basis of the production plan and the performance parameters of each unit [3, 4]. Here, take an realistic steel enterprise as an example to illustrate the general process. Make production plans according to the steel products order contracts and inventories. Determine the steel type, casting section and number of casting furnaces. Choose the production fit mode and put it into the database. The computer will arrange the ideal working procedure time and production scheduling sheets, i.e. formulate an ideal production plan and scheduling model, according to the procedure time sequences of each unit, based on the starting time of the converter refining. Working instructions are decomposed, that is to set out the production states of the converter, LF, conticaster, heating furnace, hot slab buffering devices and tandem mill at each moment, based on which the control of all kinds of operating targets is carried out. For example, to make grade steel according to the production plan. The casting slab is 200mm×1600mm, 4 casts are permitted and corresponding melting numbers are given. Put these data into the database to determine the production mode. If large cycle fit mode is chosen, the preparation time for casting of the 200mm×1600mm slab stored in the computer is 58min. The single tank casting time is 36min based on the model. Then compute the casting speed of the production line based on Equations (1) – (7), with combination of single tank casting time, refining time and hot rolling. Then the operating time sequence of each flow procedure can be calculated. Figure 5 is the simulation chart of the production scheduling. That is, Figure. 5 gives an ideal scheduling mode of a continuous production process concluding steel-making and steel rolling.

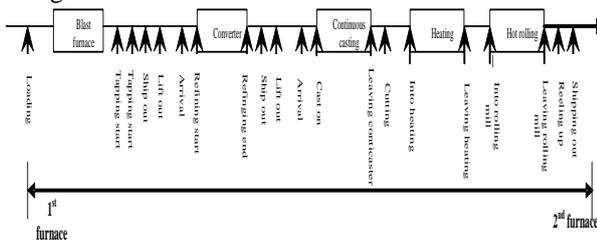


FIGURE 3 scheduling arrangement of the steel making production procedure

ID	Task	6:00	7:00	8:00	9:00	10:00	11:00	12:00
1	converter			█	█	█	█	
2	shipping			█	█	█	█	
3	finery			█	█	█	█	
4	shipping			█	█	█	█	
5	conticaster			█	█	█	█	
6	shipping			█	█	█	█	
7	hot mill			█	█	█	█	

FIGURE 4 The real working procedure time and production scheduling

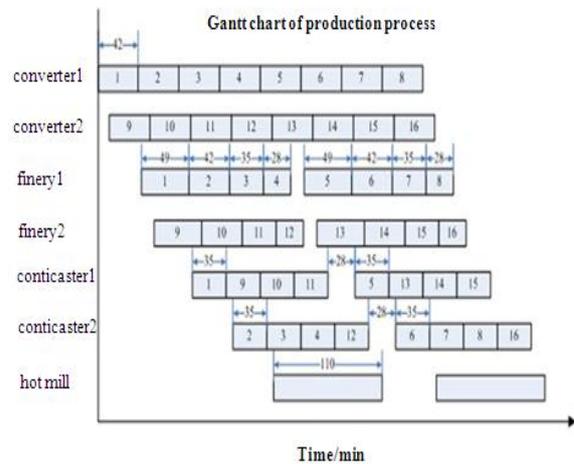


FIGURE 5 The ideal table of working procedure time and production scheduling

It can be seen from Figure 5 that:

1) The steel water production capacity of the converter is 12t. The refining cycle is 42min. There is a timeout of 18min after the continuous production of 4 converters of steel water, which could be used to carry out equipment maintenance and repair.

2) The Lf processing cycle reduces with the increase of the processing numbers. The minimal processing cycle is 28min.

3) The single tank casting time of the conticaster is 35min. The preparation time between two casts is 28min. The output time from casting beginning to send the first billet into the heating furnace or the hot buffering devices is 30min.

4) The transfer time of the ladle from the converter to LF, from LF to the conticaster are 4~6min. The dispatch table gives the dislagging working time of the crane.

5) 15 minutes after the rolling mill stops working, the inventories of billets in the hot buffering devices reduces gradually, which is the same when it comes to the inventories after the rolling mill restarts.

6) There is an accordance between the rolling capacity of the rolling mill and the billets providing capacity of the conticaster. Both are 80t/h.

In continuous production line, every working procedure time is constrained by its process condition and some external conditions. The variation of one working procedure time will cause changes in the adjacent working procedure, which will result in disturbances in the entire flow process. Thus, real-time control is needed. The database system needs two data sheets. One is a permanent data storage sheet about the real-time simulation of the working procedure time and production scheduling, with the refining number as the only symbol. The other one is an ideal data sheet for the current casting, which is an ideal scheduling model. The latter one gives the production pattern, effects range and

processing mode of refining and continuous casting. If the combination of the computers and experiences is adopted, then whether to continue casting will be determined after one casting. If the answer is yes, then the requirements of service life of the equipment should be satisfied, i.e. the water gap of the tundish and the conditions of the fire resistance materials. In the meantime, there is requirements on time plan, mainly the plan of the ladle continuous casting. Also, there is another problem which needs to be paid attention to. That is, after LF processing and lifting, whether the link is available after the last furnace of steel water finishes its casting.

**References**

[1] Haijun Niu, Jianhui Ma, Weiping Miao 2004 Research on Optimal Scheduling of Hybrid Production Processes *Journal of Xidian University (Natural Science)* 31(1) 9-12  
 [2] Weixin Hu 2002 Dispatch Management Model for Steelmaking and Rolling Continuous Production Line(1) *Steelmaking* 18(15) 21-6  
 [3] Yongliang Zhou, Ping He, Liu Liu 2003 Analysis on Production Mode and Schedule Model Design of LD Steel-making *Steelmaking* 19(5) 57-61

**5 Conclusions**

Gantt chart is obtained after the application of the production scheduling model of the steel-making process. Production plan and scheduling of the enterprise are realized.

**Acknowledgments**

This research reported in the paper is supported by National Natural Science Foundation of China (71271160) and China Scholarship Council (CSC). This support is greatly acknowledged.

[4] Hua Zhang 2006 Study on the System and Model of Knowledge Webs to Continuous Casting-Continuous Rolling in Iron and Steel Enterprises Master Dissertation of Wuhan University of Science and Technology (*in Chinese*)  
 [5] Guozhang Jiang, Yuesheng Gu, Jianyi Kong, Liangxi Xie 2011 Product Line Production Planning Model Based on Genetic Algorithm *International Review on Computers and Software* 6(6) 1023-7

Authors	
	<p><b>Guozhang Jiang, born on December 15, 1965, Tianmen, China</b></p> <p><b>Current position, grades:</b> Professor of Industrial Engineering, and the Assistant Dean of the college of machinery and automation, Wuhan University of Science and Technology.  <b>University studies:</b> the Ph.D. degree in mechanical design and theory from Wuhan University of Science and Technology, China, 2007.  <b>Scientific interest:</b> computer aided engineering, mechanical CAD/CAE and industrial engineering and management system.  <b>Publications:</b> 120.</p>
	<p><b>Chongwu Lei, born in 1988, Hubei China</b></p> <p><b>Current position, grades:</b> student M.S. degree in mechanical manufacture and automation at Wuhan University of Science and Technology.  <b>University studies:</b> B.S. degree in mechanical design manufacturing and automation from Jiangcheng College, China University of Geosciences, Wuhan, 2012.  <b>Scientific interest:</b> production planning and scheduling, industrial engineering and decision support system.</p>
	<p><b>Honghai Liu, born in 1973, China</b></p> <p><b>Current position, grades:</b> Professor in Intelligent Systems, Head of Intelligent Systems and Biomedical Robotics, University of Portsmouth.  <b>University studies:</b> PhD in Intelligent Robotics in 2003 from Kings College, University of London, UK.  <b>Scientific interest:</b> approximate computation, pattern recognition, multi-sensor based information fusion and analytics, human machine systems, advanced control, intelligent robotics and their practical applications.  <b>Publications:</b> 300.  <b>Experience:</b> research appointments at King's College London, University of Aberdeen, and project leader appointments in large-scale industrial control and system integration industry.</p>
	<p><b>Gongfa Li, born in 1979, Hubei China</b></p> <p><b>Current position, grades:</b> Associate professor, college of Machinery and Automation, Wuhan University of Science and Technology.  <b>University studies:</b> Ph.D. degree in mechanical design and theory from Wuhan University of Science and Technology in China.  <b>Scientific interest:</b> modeling and optimal control of complex industrial process.  <b>Publications:</b> nearly twenty papers in related journals.</p>

# Production procedure optimization in iron and steel enterprise

**Gongfa Li<sup>1, 2\*</sup>, Yikun Zhang<sup>1</sup>, Guozhang Jiang<sup>1</sup>, Honghai Liu<sup>2</sup>, Jia Liu<sup>1</sup>**

<sup>1</sup>College of Machinery and Automation, Wuhan University of Science and Technology, China

<sup>2</sup>Intelligent Systems and Biomedical Robotics Group, School of Creative Technologies, University of Portsmouth, PO1 2DJ, United Kingdom

Received 1 March 2014, www.tsi.lv

## Abstract

Green manufacturing is an effective way of realizing the sustainable development strategy. From the view of evolution of production chain and goods value, energy-saving, cleaner production and green manufacturing of iron and steel industry is discussed and the importance of system optimization of steel manufacturing process is stressed. Connotation of green manufacturing for the iron and steel was explained, the function of steel production process for green manufacturing was discussed and the content system of implementation of green manufacturing for the iron and steel enterprise was established. Finally, the steel production process was optimized. The function of iron and steel manufacturing procedure are broaden-manufacturing function of steel product, function of energy conversion and function of waste recycling, which will enhance enterprise's competitiveness and sustainability.

*Keywords:* production process optimization, iron and steel enterprise, green manufacturing, resource, environment protection

## 1 Introduction

Iron and steel industry is a raw material manufacturing industry, which produces iron and steel, belongs to essential industry in national economy. The iron and steel products are very important structural material and the biggest functional material in output up until now. It is the foundations of industry, agriculture, communication and transportation business and national defence industry. As an industrial family of our country, iron and steel enterprise is not merely the consumption rich and influential families of energy and resource, a large number of particulate matters, SOX, NOX, greenhouse and wastewater are released in its production process, so it becomes environmental heavy polluter. With the issue of ISO14000 environment management serial standards, OHSAS18000 occupational health and security sanitary standard series, green products authentication mark, the connotation of the market competitiveness of iron and steel enterprise is developing constantly, is not merely the competitions of quality, price now, not even the competition for centre of quality and variety, but reflect the green degree of the products, that is to say comprehensive competitions of such multifactor in many aspects as price, cost, quality, variety, capital retrieves time, product supply time, course service, environmental friendship. Iron and steel enterprise, in order to get the market competitiveness, one of the most important proposition (fundamental proposition) facing the new century is taking sustainable development using advanced green manufacturing technology.

Green manufacturing is a modern manufacturing mode that synthesisly considers environmental impact and resource consumption, its essence is embodiment of a

sustainable development strategy of human society in the modern manufacturing industry [1]. A key way to improve environmental friendship of iron and steel production process is the source control strategy consumed and polluted, namely green manufacturing mode is adopt to manage steel production process, controlling and technological innovation.

## 2 Connotation of green manufacturing for the iron and steel enterprise

Green of iron and steel industry is concrete embodiment of the green manufacturing concept in the steel and iron industry. It is not merely cleaner production, reflects the thoughts of ecological industry and recycling economy yet, namely reduction, utilizing and recycling. The following respects are reflected concretely [2].

1) Raw materials: use little iron ore and other natural mineral resource, use many renewable resources, develop the new energy with little use non-renewable energy, use the new water and fresh water resource little.

2) Production process: fully utilize resource and energy, discharge few disposals, pollutant and include the poisonous substance quality, do not use the noxious substance.

3) Products: low environmental load of the products, low pollution or non-polluting environment, increasement the service life and service efficiency of the products, reduce the pollution load to the environment of products, easy recovery and circulation use after the products scrap.

(4) Relationship with other trades and the society: offer remaining energy and by product to the society; dissolve the social offal, for instance waste steel and

\* Corresponding author e-mail: ligongfa@wust.edu.cn

waste plastics, form the ecological chain of industry with other industry effectively. Please read through the following sections for more information on preparing your paper. However, if you use the template you do not have to worry about setting margins, page size, and column size etc. as the template already has the correct dimensions.

The iron and steel enterprise belongs to a procedure manufacturing industry. Its essence of iron and steel manufacturing process is an integrative product

manufacturing system effectively and orderly collecting conversion of matter state, control of matter nature and control and management of material flow. There are influences in different levels and varying degrees on environment from a large number of materials, products flow, energy conversion processes, and diversified forms of discharge, disposal and waste materials. The input and output of the production process of iron and steel enterprise is shown as Figure 1.

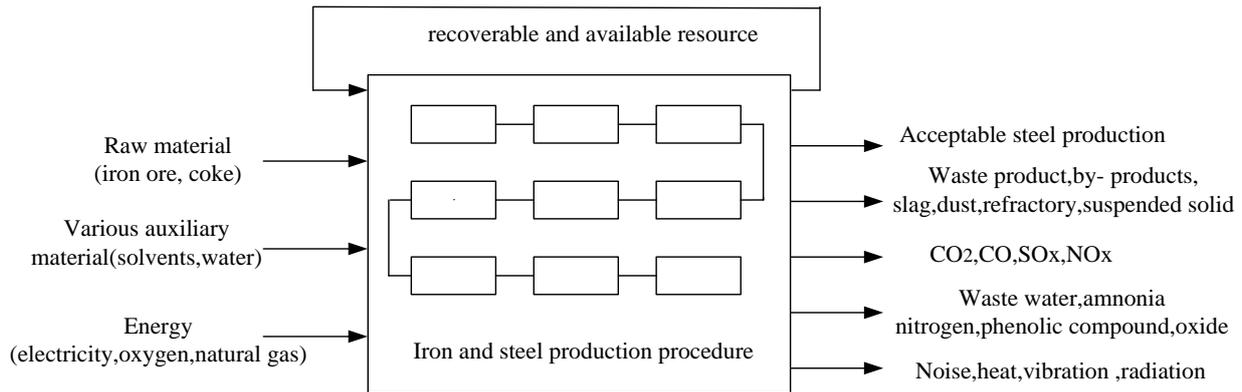


FIGURE 1 Input and output of iron and steel production process

Production of iron and steel enterprise has relatively strong dependence on mineral resource and energy, consumes a large number of energy and material and emits various kinds of incomplete material and wastes side by side. Influence degree on the environment is multi-level that all kinds of discharge materials produced in the steel production process. Generally speaking, the influence surface that the steel production process gas discharges involves the global range, such as CO<sub>2</sub>, SO<sub>x</sub>, NO<sub>x</sub>, these materials are presented the accumulating on quite great degree in the atmosphere, have enormous influence such as the greenhouse effects, acid rain. It is influenced that the liquid taking place in every process discharges and grows river system basin, the ocean or the underground aquatic products in the production process. And such factors as all kinds of slag, dust, noise and vibration, influence the life or healthy of attendants and residents of relevant communities on quite big degree.

It is from the viewpoint of ecological industry chain combining effectively ecology with economics, the steel materials are studied from mining ,manufacturing , use until discarding , retrieve , the recycled impact on ecological environment in the whole course of green manufacturing in iron and steel enterprise, production method , technological route , technological process are recognized and evaluated further, all kinds of production procedure technologies , discharge course and control technology in production process, a series of professional technique environmental protection from harmful discharge [3]. It is shown in Figure 2. The goal of green steel and iron industry should be rationalization of utilization of resources, few quantization disposal,

pollution-free or low pollution to the environment, form a ring of the ecological chain of social industry finally.

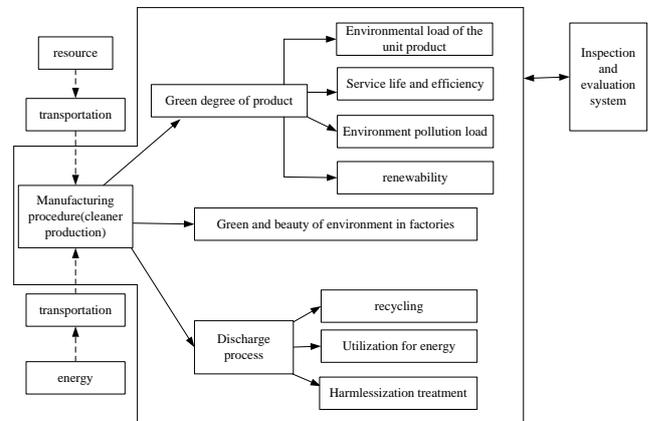


FIGURE 2 Connotation of environmental protection of iron and steel enterprise

### 3 Function and system of iron and steel production process on the basis of green manufacturing

The steel and iron industry in the 21st century can be interpreted as coordination and optimization production system considering social whole energy-conservation, reducing social environment load Green steel life cycle system [4], shown as Figure 3.

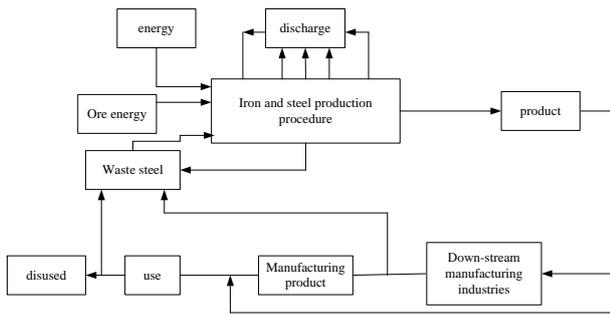


FIGURE 3 Life cycle of green steel

Therefore, the green steel production procedure should possess the following 3 kinds of functions.

1) Manufacturing function of iron and steel product. This is basic function of iron and steel produce, namely turning raw materials and energy into steel with high quality, low costs, little environmental and meeting user's demand.

2) High-efficient energy conversion function. Energy flow input has not been totally used by the iron element flow in the steel production process, should mean its recycling too through all kinds of conversion technologies that has reduced the environmental load of the society correspondingly.

3) Large social disposal treatment and disseverment function. A large number of social disposals can be dissolved and treated by making use of steel production procedure. For example, waste steel of different sources can be deal with. Waste steel is an important regenerated resource, and has important meanings in economizing natural resources, reducing the energy and raw materials consumption, reducing the environmental pollution, lowering costs and increasing employment. Waste steel be utilized to produce 1t steel economizing the 1.3t iron precise powder, reducing energy consumption of 350kg standard coal and 1.4t CO<sub>2</sub>. A large number of waste plastics are deal with, including injection through the draught of the blast furnace and hot pressing and pack into coke oven treatedly. Waste plastics with 1t are injected in blast furnace can produce the same amount of heat as 1t oil (whether Germany blast furnace can deal with 90,000 tons of waste plastics every year). Waste tire can be deal with, namely the tire can be carried on the deep cold treatment, crush, separate treatment or burning by electric stove utilizing liquid nitrogen from the oxygen machine in the steel factory. Social rubbish can be deal with, namely a large number of social rubbish in the special-purpose incinerator deal with by what the usable steel factory familiar with burning principle and the steel plant gas. Community wastewater and sewage can be deal with near community sewage in usable huge water process system of steel plant.

Green manufacturing of iron and steel enterprise should be from such respects as environment in the factory, supply of material and energy, optimization of production procedure and discharge of production process, reduce resource consumption and environmental pollution from the whole life cycle course of the steel

products production process. Its main content system is shown as Figure 4.

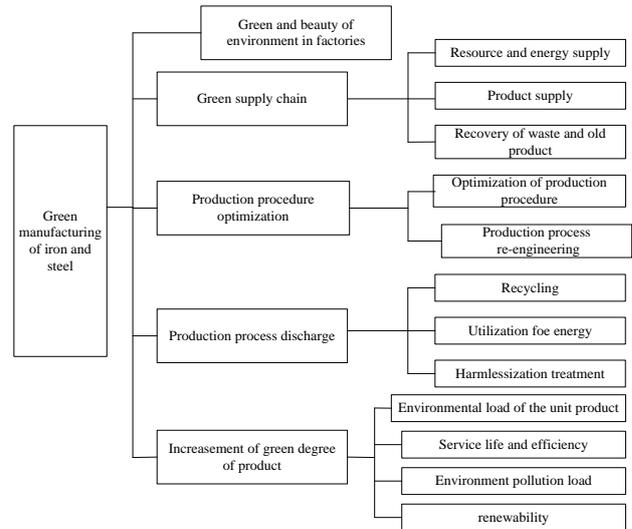


FIGURE 4 Main content system of green manufacturing of iron and steel enterprise

#### 4 Production procedure optimization of iron and steel enterprise

##### 4.1 WHOLE LIFE CYCLE MANAGEMENT OF IRON AND STEEL PRODUCT

The whole life cycle management of the iron and steel products carries on controlling technology research on the whole life cycle of the products information flow of real-time steel with the thought of the green manufacturing. Its data management mode solves steel products production process, order management; quality design, stock control, and contract retrieve of waste steel systematically. Through the whole life cycle data management of steel products effectively, strong support function is offered to improve the green degree of the iron and steel products.

It is the important embodiment in the society of iron and steel industry with friendly environment to pay attention to the performance of the steel and products, design method and retrieve to recycle. Green degree friendly to the environment that compared with other materials the steel and products demonstrated in many respects. At the same time, some indexes of green degree can be used in the friendly degree of relative environment among more different steel too.

##### 4.2 GREEN COUNTERMEASURE OF STEEL AND IRON INDUSTRY

Whether Chinese iron and steel industry could really become the green process industry with ecological nature, the key lies in whether there is green countermeasure suiting our country's national conditions. According to the actual current situation and characteristic of the steel and iron industry and his environment of our country, the

main countermeasure that the China iron and steel industry develop along the green direction are shown as following:

- 1) Optimization of the iron and steel production procedure;
- 2) Improvement resources and energy service efficiency, reduce water consumption per ton molten steel, improvement the circulation service efficiency of water;
- 3) Control the discharge of the iron and steel manufacturing process;
- 4) Recycling, utilization for energy and harmless treatment;
- 5) Improve the green degree of the steel products, such as small environmental load of the unit product, having long service life and high efficiency of the products, little environmental pollution load and good renew ability of steel products;
- 6) Formation the ecological chain of industry and give play to the social friendly function with relevant trade;
- 7) Performance and the perfection green principles and policies.

#### 4.3 GREEN KEY TECHNOLOGIES OF STEEL AND IRON INDUSTRY

The green countermeasure of iron and steel industry is relied on green technology heavily, only adopt green technology, could guarantee the implementation of the green countermeasure, and then the green strategic objective can be realized finally. There are implementing green key technology from 3 levels. (1) Popularization and promotion a batch of ripe energy-conservation and environment-protective technology. For example, coke dry quenching technology, TRT, electricity generated by the blast furnace gas, retrieve of the converter coal gas, regenerative type clean combustion, high-efficiently casting and nearly end shape casting in succession, coal injection of blast furnace, longevity of the blast furnace and renewability of LD slag. (2) Investment and development a batch of effective green technology, such as dealing with the waste plastics in blast furnace and coke oven, flue gas desulfurization, and tailings disposal technology for metal mines. (3) Exploration a batch of future green technology, such as melting iron-smelting technology and new energy development technology, the new-type coke oven technology, society's friendly offal treatment technology dealing with the old and useless tire, rubbish incinerator. Green manufacturing procedure in iron and steel enterprise is integrated further on above basis.

#### 4.4 ENERGY-CONSERVATION AND CLEANER PRODUCTION OF THE STEEL AND IRON INDUSTRY

Optimization of the steel manufacturing procedure improves the market competitiveness of the steel,

contribute to the environment again friendly and sustainable development. The global steel factory is regarding the most effective technology as motive force actively, then the steel manufacturing procedure develops with the direction from the intermittence constantly - stop - long procedure to the compactness - melt accurately in succession - shorten procedure. In fact, it is the course of using a series of power-saving technologies and cleaner production technology, so it is the embodiment of source administration even more. Proceed from height of procedure global optimization, simpler summation that but not proceed from unit process, individual equipment transformation briefly, comprehensive result can be got with maximum material obtaining rate, energy efficiency optimization and minimum manufacturing process time. Certainly, cleaner production is got and the result of environmental load reduced at the same time.

The iron and steel enterprise produces a large amount of waste gas, wastewater, and waste residue with usable heat from raw materials, coking, fritting in the production process of iron-smelting, steel-making, casting in succession and rolled steel. There are intermediate products with usable energy and semi-manufactured products among every process at the same time. It is one of the signs of enterprise's green degree to fully retrieve and utilize the energy. In the energy consumption of various kinds of industry stove, the remaining energy of the waste gas accounts for 15% - 35%, waste gas after cleaning is a good energy with easy transportation and use after combustion and no environmental pollution.

Green manufacturing is a new comprehensive strategy, which preventing the pollutant from turning into in the course of resource and production. The iron and steel enterprise green manufacturing is studied from course science, procedure technology and project, choose the resource, energy and product design rationalization, develops production technology, technical equipment, procedure of manufacturing with friendly development environment rationally, the contradiction between environment pollution and sustainable development is settled well. Its intension is mainly to deal with energy circulation at first in discharging the materials, the energy, recycling treatment and harmless treatment as small as possible finally.

1) Energy circulation is dealt with again. Concentration on chemistry is metallurgical mainly, especially chemical energy and physics heat in the iron-smelting system of the blast furnace in exhaust gas as a good energy treatment craft in converter, electric stove and steel rolling heating furnace.

2) Recycle treatment and utilization again. Treatment and utilization of materials with iron-content and available resource in converter and electric store mainly.

3) Harmless treatment. For example, concentrate on water treatment and water-saving technology mainly, the treatment technology of exhaust gas (especially flue gas desulfurization) preventing producing and treatment technology of other harmful gas.

At present, from the view of energy-conservation, cleaner production and environmental protection, the water-conservation, closed circulation and sewage treatment technology, hot water community of water of the steel factory supply technology are worth causing the special attention.

**6 Conclusion**

The development of iron and steel enterprise of 21 century faces severe challenge of lightening the environmental load of the earth. So, the iron and steel industry must transform from the manufacturing industry offering steel products simply to green manufacturing industry, from offering the function of steel products simply to the energy conversion function of playing a role in the production procedure, from rich discharge and influential family to minimum discharge and deals with some social disposals. The green of the Chinese iron and steel

industry must have excellent structural adjustment procedure carrying on green key technology from 3 levels, integrated as manufacturing procedure green for the iron and steel enterprise, so as to ensure and realize the conversion of above-mentioned functions, actively promote the cleaner production and green process of the steel and iron industry of our country. It is the only way that the steel and iron industry faces new century to take the green road. Only in this way, the steel and iron industry of our country can have sustainable development, could march towards the steel powerful country.

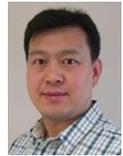
**Acknowledgments**

This research reported in the paper is supported by National Natural Science Foundation of China (71271160) and China Scholarship Council (CSC). This support is greatly acknowledged.

**References**

[1] Liu A Z, Li Y R 2004 Green design and manufacture of metallurgic machine, *Chinese Journal of machinery design and manufacture* 11 70-1

[2] Yin R Y 2002 Green manufacturing VS steel industry *Chinese Journal of iron and steel* 35 61-5  
 [3] Yin R Y 2007 Some science problems about steel manufacturing process *Chinese Journal of acta metallurgica sinica* 43 1121-8  
 [4] Yin R Y 2002 Energy-saving, clean production, green manufacturing and sustainable development of steel industry *Chinese Journal of iron and steel* 37 1-8

Authors	
	<p><b>Gongfa Li, born in 1979, Hubei China</b></p> <p><b>Current position, grades:</b> Associate professor, college of Machinery and Automation, Wuhan University of Science and Technology.  <b>University studies:</b> Ph.D. degree in mechanical design and theory from Wuhan University of Science and Technology in China.  <b>Scientific interest:</b> modelling and optimal control of complex industrial process.  <b>Publications:</b> nearly twenty papers in related journals.</p>
	<p><b>Yikun Zhang, born in 1990, China</b></p> <p><b>Current position, grades:</b> student M.S. degree in mechanical design and theory at Wuhan University of Science and Technology  <b>University studies:</b> B.S. degree in mechanical engineering and automation from Hu Bei University of Arts and Science, Xiangyang, China, 2013.  <b>Scientific interest:</b> mechanical CAD/CAE, signal analysis and processing.  <b>Publications:</b> 5.</p>
	<p><b>Guozhang Jiang, born on December 15, 1965, Tianmen, China</b></p> <p><b>Current position, grades:</b> Professor of Industrial Engineering, and the Assistant Dean of the college of machinery and automation, Wuhan University of Science and Technology.  <b>University studies:</b> the Ph.D. degree in mechanical design and theory from Wuhan University of Science and Technology, China, 2007.  <b>Scientific interest:</b> computer aided engineering, mechanical CAD/CAE and industrial engineering and management system.  <b>Publications:</b> 120.</p>
	<p><b>Honghai Liu, born in 1973, China</b></p> <p><b>Current position, grades:</b> Professor in Intelligent Systems, Head of Intelligent Systems and Biomedical Robotics, University of Portsmouth.  <b>University studies:</b> PhD in Intelligent Robotics in 2003 from Kings College, University of London, UK.  <b>Scientific interest:</b> approximate computation, pattern recognition, multi-sensor based information fusion and analytics, human machine systems, advanced control, intelligent robotics and their practical applications.  <b>Publications:</b> 300.  <b>Experience:</b> research appointments at King's College London, University of Aberdeen, and project leader appointments in large-scale industrial control and system integration industry.</p>
	<p><b>Jia Liu, born in 1990, Hubei China</b></p> <p><b>Current position, grades:</b> student M.S. degree in mechanical design and theory at Wuhan University of Science and Technology.  <b>University studies:</b> B.S. degree in mechanical engineering and automation from Wuchang institute of Technology, Wuhan, China, 2012.  <b>Scientific interest:</b> mechanical CAD/CAE, signal analysis and processing.</p>

# Spatial effect of knowledge spillover on regional economic development: an empirical study from China

**Houxing Tang\***

*School of Business Administration, Nanchang Institute of Technology, Tianxiang Road 289, Nanchang, China*

*Received 22 May 2014, www.tsi.lv*

---

## Abstract

In order to measure the spatial effect of knowledge spillover on regional economic development, a spatial Durbin model, which contains the dependent variable GDP and independent variables Capital, Labour & Knowledge stock, was constructed based on C-D production function. And then an empirical study with 31 provinces of mainland China from year 2000 to 2011 was conducted. The results show that, firstly, the Capital, Labour and Knowledge-stock all have significant positive correlation with GDP. In other words, these three factors have an important impact on the local regional economic development, but the effect of Capital is the greatest and the Labour & Knowledge-stock follow suit. Secondly, the Capital and Labour have a negative spillover effect, but the Knowledge-stock has much more positive effect. Consequently, the governments of developing regions should make full use of the spatial effect of knowledge spillover from developed regions to promote the economic restructuring and great-leap-forward development, especially when they are lack of sufficient funds to support local R&D activities.

*Keywords:* spatial effect, knowledge spillover, knowledge sharing, regional economic development, spatial durbin model

---

## 1 Introduction

The knowledge spillover is a kind of externality originating from imperfect appropriation of R&D performances. This implies that the knowledge created by anybody could be transmitted to other related people in different ways, such as reverse engineering, patents, reading scientific papers, informal communications and so on. Because of the knowledge spillover, it not only promotes the growth of a region, but also the development of other contiguous regions. Many researchers such as Jaffe (1989), Anselin (1999) had pointed out that the R&D activities had a spillover effect [1, 2]. With the help of spatial econometrics and panel data, Seyit & Ronald (2002) conducted an empirical study with 57 regions of France, Italy and Spain in order to analyse the effect of knowledge spillover on the regional economic growth in Europe. The results showed that R&D intensity and R&D spillover have significant positive effects on regional economic growth [3]. And then, the flows of knowledge among regions can increase the production efficiency and promote the economic development. Especially, with the development of new economic geography, the spatial externality is becoming the core element and the hot topic. Therefore, many researchers have put focus on the interplay of different regions. Similarly, in Regional Economics more and more attentions are put on the effect of economic spillover of the local region on neighbouring regions.

Surely, the researchers not only hope to prove the

existence of this knowledge spillover effect, but also want to measure the spatial effect and explore its characteristics. For examples, based on Griliches–Jaffe knowledge production function, Greunz (2003) proposed an improved model to investigate inter-regional knowledge spillover across European sub-national regions. The result showed that if knowledge spillover occurred within a given country, the national border turned out to seriously hamper interregional spillover on the European scale [4]. Also, an empirical study from Scherngell et al. (2007) indicated that the knowledge spillover would be local concentration when analysing the effect of knowledge spillover in Europe on the total factor productivity and the degree of knowledge spillover between industries [5]. Andrea & Chiara (2011) thought that the statistical evidence suggested that the relevance of knowledge spillover had increased over time. A region's absorptive capacity, measured by local R&D expenditure and social capital, implied a reduction of outward knowledge spillover [6]. Bernard & Lesage (2011) examined the spatial spillovers associated with public and private research expenditures in own- and other-industry sectors of 94 French regions. The empirical results showed that the largest direct and indirect effects are associated with private R&D activity which spilled across industry boundaries [7].

In China, an empirical study of the spatial pattern of China's R&D spillover at provincial level was conducted by Su (2006). The results showed that the spatial dependencies of provincial R&D knowledge production

---

\* *Corresponding author* e-mail: tanghouxing@aliyun.com

existed, and R&D spillover were also locally bounded. Under the given conditions, spatial lag model showed that an average increase of 0.22 percent in a regional patent production was caused by one percent increase in neighbouring regional patent production. Moreover, a decay process of provincial innovation was found to be existed by considering the research effort made by neighbouring regions [8]. Based on an empirical study of BRICs from year 1980 to 2008, Li & Han (2010) had proved the weak tacit knowledge spillover was positive correlation with an output of region [9]. There were knowledge spillover among the provinces and the knowledge stock had a positive effect on regional economical growth. Especially, Xu et al. (2010) pointed that the efficiency of growth was affected by regional capital and labour and absorbing capacity [10]. Zhou & Lan (2012) thought that the knowledge spillover effect weakened the profit of originated region from new knowledge, but it was good to reduce the gap among regions and promote the coordinated development of regional economy [11]. Addition to, in common theory, the effect of knowledge spillover will gradually decrease with the increasing of distance. However, Liu & Tang (2010) proved that the effect of knowledge spillover did not always descent strictly with the increasing order of contiguity matrix, and there maybe existed an optimal distance for spillover [12].

Based on these literatures review, we find that on one side, the spatial econometric model provides a useful way to measure and evaluate the effect of knowledge spillover on agglomeration, innovation and regional economic growth. However, Lesage & Pace (2009) had pointed that the ordinary regressive method would get inaccurate coefficients and we need to change a new method [13]. On the other side, there are lack of empirical studies on the effect and characteristics of knowledge spillover on the regional economic development. In other words, we not only hope to prove the existence of spatial effect of knowledge spillover, but also explore the nature of such effect. Therefore, we can deeply understand the regular pattern of knowledge sharing among regions. This will help us to make effective regional policies to promote the regional development soundly and rapidly.

Therefore, based on [14] and [12], in the section II, a spatial Durbin model is transformed from a production function whose dependent variable is GDP and independent variables are Labour, Capital and Knowledge-stock. In the section III, the data are collected and reprocessed firstly, and then, the Moran Index is calculated to roughly test the existence of spatial correlation. We use the toolbox of Matlab R2010b to calculate the model and give a detailed analysis. In the section IV, a stability analysis is conducted in order to test whether the change of parameter affect the results. In section V, the conclusions and police recommendations are given.

## 2 Constructing the spatial Durbin model

### 2.1 ANALYSIS AND SELECTION OF VARIABLES

Now, the GDP is a very important index to evaluate the regional economic development. Therefore, we use the GDP ( $G$ ) as depended variable. From the perspective of knowledge production function, we select *the total investment in fixed assets in the whole country* as the capital input variable ( $K$ ), and select the *number of employed persons* as the labour input variable ( $L$ ). How to measure the knowledge stock ( $S$ )? In most literatures, the patent is a good choice. But it has many defects. Kesidou (2004) showed that the patent was often treated as a proxy of innovation, but it only contains a formal output and neglects other complex activities useful for knowledge accumulation [15]. In other words, the patent only stands for the measurable knowledge. So, we also divide the knowledge variable  $S$  into  $S1$  and  $S2$ , which stand for the measurable knowledge and un-measurable knowledge respectively. Surely, the patent is a proxy of the  $S1$ .

As the other goods, the value of knowledge is also depreciation. Therefore, we will measure the present value of knowledge stock through *perpetual inventory method*. The formula is as follows

$$S1_t = N_t + S1_{t-1}(1-d), \quad (1)$$

where  $t$  stands for time (year),  $S_t$  stands for the knowledge stock at  $t$ ,  $N_t$  stands for the knowledge added-value at  $t$ , and  $d$  stands for depreciation rate. In other literatures, the  $d$  is set to 12% [14] and set to 15% [16]. Considering the data is from China, we also set the  $d$  to 15%. Surely, it may be quite arbitrary. Therefore, a stability analysis will be provided in section IV.

In Equation (1), we need to get the initiative knowledge stock  $S1_0$ , which is calculated as follow

$$S1_0 = N_0(1+g)/(g+d), \quad (2)$$

where  $S1_0$  stands for the knowledge stock of base year,  $N_0$  stands for number of patent application and authorized of base year, and  $g$  stands for average growth rate per year of patent application and authorized.

### 2.2 MODELLING

According to the selected variables, we suppose that they meet the function as follow

$$G = f(K, L, S1, S2), \quad (3)$$

where  $G$  stands for GDP;  $K$  stands for capital input that is the total investment in fixed assets in the whole country;  $L$  stands for labour input that is the number of employed

persons; S1 stands for measurable knowledge that is number of patent application accept and granted; S2 stands for un-measurable knowledge. According to the C—D production function, we get

$$G = AK^\alpha L^\beta S1^\gamma S2^\lambda, \tag{4}$$

where  $\alpha, \beta, \gamma, \lambda$  need to be estimated; and A is constant and using logarithm in the formula (4), we get  $LnG = \alpha LnK + \beta LnL + \gamma LnS1 + \lambda LnS2 + LnA$ , then let

$$G' = \alpha K' + \beta L' + \gamma S_1' + \lambda S_2' + c. \tag{5}$$

Many literatures have confirmed that the measurable knowledge and un-measurable knowledge both had spatial dependence [17, 14, 13].

Therefore, there is a simple hypothesis as follows:

$$S_1' = \varphi_1 WS_1' + u_1 \quad u_1 \sim N(0, \sigma_{u_1}^2 I_n), \tag{6}$$

$$S_2' = \varphi_2 WS_2' + u_2 \quad u_2 \sim N(0, \sigma_{u_2}^2 I_n), \tag{7}$$

where  $W$  stands for spatial contiguity matrix;  $\varphi_1, \varphi_2$  stands for the spatial dependence of measurable knowledge stock and un-measurable knowledge stock respectively, which reflect the spatial dependence of sample data and measure the average effect of neighbouring regions on local region;  $u_1, u_2$  are stochastic error respectively.

Obviously, the measurable knowledge and un-measurable knowledge is different but there is a correlation between them. This relationship can be reflected from stochastic error, so we suppose that the  $u_1, u_2$  meet a simple linear equation as follows:

$$u_1 = \omega u_2 + \varepsilon \quad \varepsilon \sim N(0, \sigma_\varepsilon^2 I_n). \tag{8}$$

According to the Equations (5–8), we get

$$G' = \theta_1 WG' + \theta_2 K' + \theta_3 L' + \theta_4 S_1' + \theta_5 I + \tilde{\theta}_2 WK' + \tilde{\theta}_3 WL' + \tilde{\theta}_4 WS_1' + \tilde{\theta}_5 WI + \tilde{\varepsilon}, \tag{9}$$

where  $\theta_1 = \varphi_2$ ,  $\theta_2 = \alpha, \tilde{\theta}_2 = -\alpha\varphi_2$ ,  $\theta_3 = \beta, \tilde{\theta}_3 = -\beta\varphi_2$ ,  $\theta_4 = r + \lambda\omega^{-1}$ ,  $\tilde{\theta}_4 = -(\gamma\varphi_2 + \lambda\omega^{-1}\varphi_1)$ ,  $\theta_5 = c, \tilde{\theta}_5 = -\varphi_2 c$ ,  $\tilde{\varepsilon} = -\omega^{-1}\lambda\varepsilon$ .

Now, we have found that the Equation (9) is just a spatial Durbin model.

### 3 Empirical results and analysis

#### 3.1 DATA COLLECTION AND PRE-PROCESS

In this paper, the panel data set contains 31 provinces in mainland of China with 4 variables over the period from 2000 to 2011. In the State Intellectual Property Office of the People’s Republic of China, the statistical reports start from 1985, but the data in most regions are zero. Addition to, Hainan province was founded in 1988 and had data from 1989. Chongqing became a municipality directly under the Central Government in 1997, but it has independent statistic since 1985. Generally speaking, there are complete and effective statistical data from 1989. Therefore, we set 1989 as the base year to calculate the knowledge stock. Especially, when we calculate the average growth rate per year of the patent (i.e. variable  $g$ ), the base year of Xizang is 1990, because the corresponding data is zero in 1989.

In addition to, considering the availability of data, the GDP is collected from “Gross Regional Product and Indices” in statistical yearbook from 2001 to 2012. The Labour is collected from “Number of Engaged Persons in Private Enterprises and Self-employed Individuals” from 2001 to 2012. The Capital is collected from “Total Investment in Fixed Assets in the Whole Country by Status of Registration and Region” from 2001 to 2012. The patents are collected from “Patents Application Accepted and Granted by Region” from 1985 to 2011. (Data source: Statistics Annual Report 1985-2011 from State Intellectual Property Office of the People’s Republic of China, China Statistical Yearbook 2001-2012 from National Bureau of Statistics of China).

#### 3.2 CALCULATING AND ANALYZING THE SPATIAL DURBIN MODEL

Spatial regression models exploit the complicated dependence structure between observation units, which represent countries, regions and so on. Because of this, the parameter estimates contain a wealth of information on relationships among the regions [12]. A change on any given explanatory variable of a region will impact the region itself (e.g. a direct impact) and potentially impact all other regions indirectly (e.g. an indirect impact). That is to say, the indirect impact is just the spillover effect. It implies that the ordinary regression method is not accurate. And then we need other ways to estimate the actual spillover effect rather than use the ordinary regression coefficient  $\tilde{\theta}_4$  of Equation (9). The new way is called *partial derivative summation method* in this paper.

With the help of spatial econometric toolbox and Matlab R2010b, we get the results as follows in Table 1 and Table 2.

Firstly, from the Table 1, we find that the value of t-tests of all variables surpass 2. This indicates that all the variables have significant linear correlation with GDP.

An interesting result is that the Capital and Labour have positive correlation with GDP but the W-Capital and W-Labour have negative correlation with GDP. However, the Knowledge-stock and W-Knowledge-stock have both positive correlations with GDP. As we all know, the W stands for the spatial effect. Therefore, this result shows that the capital and labour of a region make negative spillovers effect for neighbouring regions.

Secondly, because the regression coefficients are not to reflect the actual spillover effect, we will evaluate the spatial effect by Table 2. The first, in direct effect, the values of Capital and Labour reach to 0.745436 and 0.252985 respectively, and exceed greatly the value of Knowledge-stock, which equal to 0.076793. This implies that the capital (e.g. fixed asset investment) and labour rather than technology innovation are still the main driving force of regional economic development. In Table 3, the contribution share of gross capital formation to the increase of the GDP is given from year 1978 to 2011, which approves the above conclusion.

The second, in the indirect effect, there is also an interesting result. The spatial spillover effect of Capital and Labour are negative while the effect of Knowledge-stock is positive. How to understand this phenomenon? We think that the capital is treated as a kind of homogeneous and excludable resource. Once it is invested into a region, it can not be used in other regions. And similarly, the labour is a kind of heterogeneous and excludable resource. Because of its quick flowing in present China, the growth of labour (especially the advanced workers) in a region will inevitably reduce the supply of labour in other regions in a fixed period. Consequently, this indicates that it will lower the GDP in

local region because of the increasing of capital and labour in neighbouring regions. Otherwise, the knowledge is a kind of non-rivalry and partially excludable resource. And it is reused by others. This implies that the growth of knowledge stock in one region will indirectly increase the growth of knowledge stock in neighbouring regions. Consequently, this will increase the GDP in neighbouring regions. Therefore, it is acceptable that the spillover effect of Capital & Labour is negative and that of Knowledge-stock is positive. Moreover, the significant consequence of the characteristic is that how to make full use of the spatial spillover effect of knowledge.

The third, for the total effect, we find that the effect of Knowledge-stock exceed the effect of Labour. This implies that the knowledge-stock (i.e. innovative knowledge or high-tech) rather than more pure labour is playing a stronger role in regional economic development. The knowledge intensive industry will gradually substitute the labour intensive industry with the increasing input on education.

Thirdly, all these analysis indicate that knowledge spillover will play a more and more important role in regional economic development. One side, each region should put much more attention on the knowledge production. For example, increasing the science and technology innovation input. One the other side, many regions maybe are lack of necessary capital or labour, so they can efficiently utilize the knowledge spillover effect from other regions in order to promote own region powerful. For instance, they can engage in extensive technological exchanges and cooperation with other regions.

TABLE 1 Model estimation based on linear regression

Variable	Coefficient	Asymptot t-stat	z-probability
Constant	0.500267	3.664145	0.000248
Capital	0.763673	28.422736	0.000000
Labour	0.258725	10.255035	0.000000
Knowledge-stock	0.060986	5.569024	0.000000
W-Capital	-0.467287	-10.256821	0.000000
W-Labour	-0.144594	-4.362333	0.000013
W-Knowledge-stock	0.159103	7.904632	0.000000
rho	0.328976	5.726039	0.000000

TABLE 2 The estimation of spatial effect based on partial derivative summation

Direct	Coefficient	t-stat	t-prob	Lower 01	Upper 99
Capital	0.745436	29.528766	0.000000	0.674685	0.807168
Labour	0.252985	10.742312	0.000000	0.192955	0.314691
Knowledge-Stock	0.076793	7.342959	0.000000	0.047858	0.105268
Indirect	Coefficient	t-stat	t-prob	Lower 01	Upper 99
Capital	-0.304283	-8.638289	0.000000	-0.400282	-0.216711
Labour	-0.083380	-2.420611	0.015973	-0.169920	-0.000575
Knowledge-Stock	0.251915	9.882188	0.000000	0.180004	0.325630
Total	Coefficient	t-stat	t-prob	Lower 01	Upper 99
Capital	0.441153	12.948609	0.000000	0.352705	0.530985
Labour	0.169605	6.991557	0.000000	0.106852	0.227901
Knowledge-stock	0.328708	11.209642	0.000000	0.255210	0.404317

TABLE 3 Contribution share of Gross Capital Formation to the increase of the GDP

Year	Contribution Share (%)	Year	Contribution Share (%)
1978	66.0	2004	54.5
1980	26.4	2005	38.5
1985	80.9	2006	43.6
1990	1.8	2007	42.5
1995	55.0	2008	46.9
2000	22.4	2009	87.6
2001	49.9	2010	52.9
2002	48.5	2011	48.8
2003	63.2	--	--

Data source: China Statistical Yearbook 2012 from National Bureau of Statistics of China.

#### 4 Stability of Model

In our model, the parameter  $d$  is set to 15%, which seems arbitrary. So, in this section, a stability analysis is taken for different  $d$ . We set  $d$  to 10%, 12%, 18% and 20% respectively to test whether the change of  $d$  will impact the stability of the model. The detailed results are as follows in Table 4 and Table 5. From Table 4, we find that the changes of  $d$  do not change the linear correlation between GDP and other independent variables. It only changes the regression coefficients a little. From the

Table 5, we find that there is a little change on indirect effect of Labour including its value and direction. However, the absolute values of all t-test are far less than 2, which show that there is not a linear correlation between GDP and W-Labour. In Table 3, although it passes to the t-test, the indirect effect of Labour only reaches to -0.08, which is much lower compared with Capital and Knowledge-stock. All these indicate that the Labour only has much weak negative spillover effect on GDP, and the changes of the parameter  $d$  do not change this situation. So, the model in our paper is stability.

TABLE 4 Model estimation with the changes of parameter  $d$ 

Variable	d=10%		d=12%		d=18%		d=20%	
	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
Constant	0.19269	1.36682	0.20914	1.48247	0.24148	1.70667	0.24584	1.73765
Capital	0.84978	33.84629	0.84799	33.60750	0.84406	33.09362	0.84337	33.01704
Labour	0.17954	7.05091	0.18067	7.05531	0.18462	7.11745	0.18631	7.16648
Knowledge-stock	0.03635	3.27409	0.03636	3.27657	0.03575	3.23700	0.03520	3.19848
W-Capital	-0.59899	-13.30724	-0.59093	-13.02385	-0.57818	-12.60055	-0.58011	-12.68296
W-Labour	-0.09875	-3.12122	-0.09533	-2.97931	-0.08982	-2.73716	-0.09085	-2.75603
W-Knowledge-stock	0.10631	5.06790	0.10601	5.06654	0.10324	4.98099	0.10141	4.91478
rho	0.51400	10.37155	0.50396	10.04316	0.49099	9.62029	0.49596	9.76483

TABLE 5 The spatial effect with changes of parameter  $d$ 

Direct	d=10%		d=12%	
	Coefficient	t-stat	Coefficient	t-stat
Capital	0.824719	34.608679	0.823463	34.412799
Labor	0.177917	7.602651	0.179406	7.616899
Knowledge-Stock	0.056386	4.931707	0.055722	4.899498
Indirect	Coefficient	t-stat	Coefficient	t-stat
Capital	-0.309845	-6.675854	-0.306349	-6.700182
Labor	-0.012401	-0.325573	-0.008038	-0.211120
Knowledge-Stock	0.238665	5.943980	0.232735	5.944784
Total	Coefficient	t-stat	Coefficient	t-stat
Capital	0.514874	10.513328	0.517114	10.753491
Labor	0.165516	5.202213	0.171368	5.408059
Knowledge-stock	0.295051	6.457783	0.288456	6.472558

TABLE 5 The spatial effect with changes of parameter  $d$  (continued)

Direct	d=18%		d=20%	
	Coefficient	t-stat	Coefficient	t-stat
Capital	0.820461	33.584064	0.819707	33.477667
Labor	0.184482	7.656494	0.186647	7.715160
Knowledge-Stock	0.053744	4.644222	0.053164	4.598221
Indirect	Coefficient	t-stat	Coefficient	t-stat
Capital	-0.297518	-6.984009	-0.296769	-6.913834
Labor	0.002505	0.064312	0.003888	0.098670
Knowledge-Stock	0.220095	5.684304	0.218575	5.614861
Total	Coefficient	t-stat	Coefficient	t-stat
Capital	0.522944	11.549711	0.522938	11.439479
Labor	0.187347	5.691958	0.190534	5.681662
Knowledge-stock	0.273840	6.298379	0.271740	6.212776

## 5 Conclusions and recommendations

With the development of knowledge-based economy, the knowledge will substitute other production factors such as capital and land to become the main driving force of regional economic development. The spatial spillover of knowledge is becoming more and more widespread because different regions become more and more closed with each other. The knowledge and such spatial effect will play more and more important role in regional economic development. The question is that how to measure and evaluate the spatial spillover effect of knowledge and understand its characteristics.

In this paper, knowledge is divided into two parts: measurable knowledge and un-measurable knowledge. With the help of knowledge production function, a spatial Durbin model is formed. The empirical study containing 31 provinces over the period from 2000 to 2011 in China shows that there is spatial dependence among regions. The knowledge spillovers go beyond the administrative boundary and have an impact on neighbouring regions. The interesting phenomenon is that capital and labour have a negative spatial effect, but knowledge has positive spatial effect. This result indicates that spatial knowledge spillover will play a more and more important role in regional economic development especially in developing regions. In other words, although the fixed asset investment is the main driving force of regional economic development, but these regions have insufficient money. Also, from the perspective of sustainable development, these regions cannot depend on labour-intensive industry for long term. The knowledge has become the strategic resource to achieve and keep competitive advantages, but the knowledge resources are mainly concentrated in developed regions rather than developing regions because of location, history, finance supporting and so on. Fortunately, the knowledge has the characteristic of public good. Therefore, a meaningful question is that how to make full use of the spatial spillover effect of knowledge.

In this paper, many policy recommendations are given as follows. Firstly, it is essential to increase expenditure on the knowledge production such as R&D and improve the contribution share of technology innovation to GDP. Since knowledge is the most important resource to achieve competitive advantages. Therefore, it is essential to guarantee a certain amount of knowledge stock. At present, the average ratio of R&D expenditure to GDP has reached to 1.84% in 2011 and ranked third in the world. However, according to statistical bulletin in 2011 of national science & technology funds, the R&D expenditure of eight provinces (e.g. Jiangsu, Guangdong, Beijing, Shandong, Zhejiang, Shanghai, Liaoning and

## References

- [1] Jaffe AB 1989 Real effects of academic research *American Economic Review* 79(5) 957-70

Hubei) accounts for 66.5% of national expenditure. In other words, most other provinces are lack of sufficient input on knowledge production.

Secondly, it may be a good way to construct a virtual alliance of knowledge production and promote the development of collaborative innovation. The knowledge has the characteristics of stickiness and the locality. Obviously, it is a slow and difficult process for developing regions to achieve the knowledge through spillover. Therefore, it is an effective way to construct knowledge alliance and promote the collaborative innovation. It is useful to break the administrative boundary and promote spatial knowledge spillovers quickly. According to the reports, the ministry of education and ministry of finance have implemented collaborative innovation plan in 2012 in China.

Thirdly, it is essential to promote the flow of knowledge through an effective mechanism of talent flow in order to balance the knowledge gap between regions. In fact, innovative knowledge comes from innovative talent and the knowledge spillover is achieved mainly by the flow of talent. Unfortunately, the developed regions attract too many talents because of the *black hole effect*. Although the governments of developing regions provide the talents many preferential policies in order to get them stay in the locality, it still fails. Therefore, in the present situation, the most practical way is to make effective mechanism of talent flow. The developing regions can get innovative knowledge from such flow rather than being entangled in a local production of knowledge.

Fourthly, it is very important to improve the knowledge absorption capacity of developing regions. If the developing regions do not have such absorption capacity, then facing the knowledge spillover will only lead to frustration and little innovation will take place. We can take an example to explain it. Why do so many multi-national corporations invest the mainland China rather than other developing countries? Except the cost of labour, one of the critical factors is that the China has enough knowledge employees to meet the need of production. Undeniably, this brings China an opportunity for rapid development compared with other developing countries.

## Acknowledgement

The author acknowledge the financial support by National Natural Science Foundation of China (grand: 71261019), the Youth Fund Project of College Humanities and Social Sciences of Ministry of Education (grand: 11YJC630193) and the Youth Fund Project of College Humanities and Social Sciences of Jiangxi Province (grand: GL1244).

- [2] Anselin L 1999 *Spatial Econometrics Working Paper School of Social Science University of Texas at Dallas Richardson* 26 April  
 [3] Seyit K, Ronald M 2002 Knowledge Spillovers and Regional Growth in Europe European Regional Science Association ERSA

- available at <http://www.sre.wu-wien.ac.at/ersa/ersaconfs/ersa02/cd-rom/papers/373.pdf> (accessed 15 June 2009)
- [4] Greunz L 2003 Geographically and Technologically Mediated Knowledge Spillovers between European Regions *Annals of Regional Science* 37(4) 657-80
- [5] Scherngell T, Manfred M F, Reismann M 2007 Total Factor Productivity Effects of Interregional Knowledge Spillovers in Manufacturing Industries across Europe *Romanian Journal of Regional Science* 1(1) 1-16
- [6] Andrea C, Chiara D B 2011 Determinants of spatial knowledge spillovers in Italian provinces, *Socioeconomic Planning Sciences* 45(1) 28-37
- [7] Corinne A B, Lesage J P 2011 Quantifying Knowledge Spillovers using Spatial Econometrics Models *Journal of Regional Science* 51(3) 471-96
- [8] Su F L 2006 An analysis on the spatial pattern of China's provincial R&D spillovers *Studies in Science of Science* 5 696-700
- [9] Li Y, Han B T 2010 The impact of weak tacit knowledge spillover on economic growth based on paper cooperation *Studies in Science of Science* 28(10) 1547-54
- [10] Xu Y Z, Zhu Y X, Sun J 2010 Knowledge spillover and regional economy growth: A spatial empirical analysis *Science Research Management* 31(6) 105-12
- [11] Zhou G F, Lan Y N 2012 Spatial Cluster Knowledge Spillovers and Regional Economy Increasing Discrepancy Based on a Spatial Econometric Analysis in the Perspective of Cities *Journal of Industrial Technological Economics* 3 31-9
- [12] Liu M F, Tang H X 2010 Empirical Study on Effect of Knowledge Spillovers Based on Spatial Durbin Model *Science & Technology Progress and Policy* 27(18) 28-33
- [13] Lesage J P, Pace R K 2009 *Introduction to Spatial Econometrics* CRC Press New York.
- [14] Lesage J P, Fischer M M 2008 Using Dependence to Measure the Impact of Tacit Knowledge Stocks on Regional Total Factor Productivity available at: [http://www.stanford.edu/group/SITE/archive/SITE\\_2008/segment\\_7/papers/lesage.pdf](http://www.stanford.edu/group/SITE/archive/SITE_2008/segment_7/papers/lesage.pdf) (accessed 15 June 2009)
- [15] Kesidou E 2004 Knowledge Spillovers in High-tech Clusters in Developing Countries" available at: [http://www.globelicsacademy.net/pdf/EffieKesidou\\_paper.pdf](http://www.globelicsacademy.net/pdf/EffieKesidou_paper.pdf) (accessed 15 June 2009)
- [16] Lv Z W 2009 Research on the R&D Spatial Spillover and Regional Knowledge Production *Statistical Research* 26(4) 44-51
- [17] Corinne A B 2001 Science and Knowledge Flows: Evidence from the French Case *Research Policy* 30(7) 1069-78

## Authors



**Houxing Tang, July, born in 1982, Anhui Province in China**

**Current position, grades:** associate professor

**University studies:** PH.D for Management Science and Engineering from School of Information Management, Jiangxi University of Finance and Economics in 2010.

**Scientific interest:** knowledge management and regional innovation.

**Publications:** 15.

**Experience:** School of Business Administration, Nanchang Institute of Technology.

# A fuzzy set approach for a multi-period optimal portfolio selection model

Xing Yu\*

Department of Mathematics & Applied Mathematics, Hunan university of humanities, science and technology, Loudi, 417000, P.R. China

Received 1 March 2014, www.tsi.lv

## Abstract

Due to portfolio decision deals with future events and opportunities, the market information is uncertain. This paper aims to propose a fuzzy multi-period portfolio selection model to hedge against the uncertainty. A new transformation method based on qualitative possibility theory is developed to transfer the model to a crisp programming, which can be solved by an optimization technique. An example is used to illustrate our approach.

*Keywords:* Multi-period portfolio selection, Fuzzy sets, Optimization, Qualitative possibility theory

## 1 Introduction

The purpose of an investor is to predict the future return and decide how to allocate asset optimally for Maximilian the total return under some constraints, such as a budget constraints. However, the confront with two problems. One is culti-period invest strategy. Another is uncertainty. As every knows, the time horizon of an investment for an investor usually is multi-period, because he will adjust invest strategy according to market and his budget, from this point, Markovitz's single mean-variance analysis model [1] does not conform to the actual condition. A numerous cholers extended the single period portfolio to the dynamic case. Hakansson [2] gave the multi period mean variance model based on the general portfolio selection theory. Li and Ng [3] generalize Markowitz's mean-variance model to the multi-period model under discrete case, and deduced the analysis expression of efficient frontier by establishing auxiliary function.

It is clear that history data not mean the future data because the market condition is changing time and time. Conventional portfolio optimization models have an assumption that the future condition of stock market can be accurately predicted by historical data. However, no matter how accurate the past data is, this premise will not exist in the financial market due to the high volatility of market environment. To deal with imprecise information in making portfolio selection decisions. Östermark [4] used the fuzzy decision theory to study dynamic portfolio problems with a risk-free asset and risky asset and proposed a fuzzy control model. Sadjadi etc. [5], who researched the fuzzy portfolio selection problem with different borrowing and lending rate. Huang and Qiao [6] tried to study the multi period portfolio problem under fuzzy environment, and proposed a risk index of an

uncertain multi period portfolio problem. Liu [7] proposed a pair of two-level mathematical programs to calculate the upper bound and lower bound of return and transfer it into a pair of ordinary one-level linear programs.

This paper is organized as follows. In the next section, we present the fuzzy number and its four arithmetic operates. In section 3, we introduce the formulation of transformation for a fuzzy portfolio model to a crisp programming. Section 4 presents a numerical example with real data from the Chinese stock market. The paper ends with some concluding remarks.

## 2 Notation of fuzzy numbers

From [8], membership function of a trapezoidal fuzzy number  $\tilde{A} = (r_1, r_2, r_3, r_4)$  is defined as

$$\mu_{\tilde{A}}(x) = \begin{cases} 0, & x < r_1 \\ \frac{x - r_1}{r_2 - r_1}, & r_1 \leq x < r_2 \\ 1, & r_2 \leq x < r_3 \\ \frac{r_4 - x}{r_4 - r_3}, & r_3 \leq x < r_4 \\ 0, & x \geq r_4 \end{cases}, \text{ where } r_2, r_3 \text{ are the left and right modal values and } r_1, r_4 \text{ are the left and right spreads.}$$

Triangular fuzzy number is a special case of trapezoidal fuzzy number, for any  $M = (r_1, r_2, r_3, r_4)$ , which is a trapezoidal fuzzy number, when  $r_2 = r_3$ , trapezoidal fuzzy number will degenerate to the triangular fuzzy number  $M = (r_1, r_2, r_3)$ .

\* Corresponding author e-mail: hnyuxing@163.com

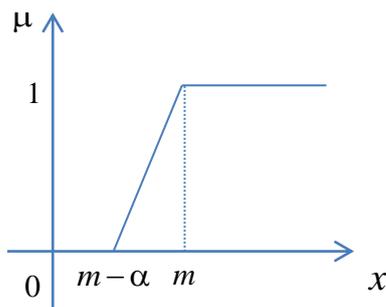


FIGURE 1

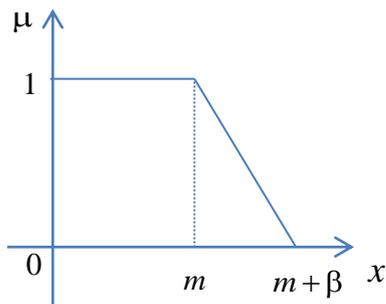


FIGURE 2

There exists certain flexibility in making the portfolio decision .Fig 1 denotes the degree of satisfaction of the expected value of the portfolio. If the expected value is larger than  $m - \alpha$  , the degree of satisfaction increase, when the expected value is larger than  $m$  the investor is completely satisfy. Fig 2 shows the degree of satisfaction of the budget spending. If the budget spending is less than  $m$  , the investor is completely satisfy. However, if the budget spending is greater than  $m$  , then the degree of satisfaction decrease. When the budget spending is greater than  $m + \beta$  , the degree of satisfaction becomes zero.

The four arithmetic operates on a trapezoidal fuzzy number as follows:

Suppose  $M_1 = (r_1, r_2, r_3, r_4)$  ,  $M_2 = (s_1, s_2, s_3, s_4)$  are two trapezoidal fuzzy numbers, then  $M_1 + M_2 = (r_1 + s_1, r_2 + s_2, r_3 + s_3, r_4 + s_4)$  for any

$$\lambda \in R \quad , \quad \lambda M_1 = \begin{cases} (\lambda r_1, \lambda r_2, \lambda r_3, \lambda r_4), & \lambda \geq 0 \\ (\lambda r_4, \lambda r_3, \lambda r_2, \lambda r_1), & \lambda < 0 \end{cases} \quad ,$$

$$M_1 \times M_2 = (r_1 s_1, r_2 s_2, r_3 s_3, r_4 s_4), r_1, s_1 > 0.$$

### 3 Problem statement and modelling

#### 3.1 INTRODUCTION TO QUALITATIVE POSSIBILITY THEORY

This section introduces qualitative possibility theory, see to Dubois and Prade [9], which is used to deal with the fuzzy constraint involving both uncertain and flexible parameters. The basic concept is introduced as follows.

Let  $U$  be a set of states and  $X$  be a set of possible consequences, a possibility distribution  $\pi$  represent the incomplete knowledge on the state on  $U$  and  $X$  be the possibility distribution representing the preference of decision maker on  $X$  . The utility of a decision  $d$  whose consequence in state  $u$  is  $x = d(u)$  for  $u \in U$  can be evaluated by combining the plausibilities  $\pi(u)$  and the utilities  $u(x)$  in a suitable way. Two quantitative criteria were proposed by Dubois and Prade to evaluate the worth of decision  $d$  regarding uncertain information:

Pessimistic criterion:

$$U_*(d) = \inf_{u \in U} \max(1 - \pi(u), \mu(d(u))). \quad (1)$$

Optimistic criterion:

$$U^*(d) = \sup_{u \in U} \min(\pi(u), \mu(d(u))). \quad (2)$$

In this paper, the pessimistic criterion is used to determine the satisfaction degree of the fuzzy constraint that contains uncertain parameters on its left-hand side and flexible parameters on its right-hand side. For example, the left-hand side is the future fuzzy cost, and the right-hand side is the flexible budget of a investor. The satisfaction degree of the decision  $d$  can be defined

$$C_d(\tilde{R}, \tilde{B}) = \inf \max(1 - \mu_{\tilde{R}}(x), \mu_{\tilde{B}}(x)), \quad (3)$$

where  $\tilde{R}$  is the possible consequences of decision  $d$  and  $\tilde{B}$  is the preference of decision maker about the consequence.

#### 3.2 MODEL FORMULATION OF THE PORTFOLIO SELECTION

The multi-period portfolio selection problem is to select a set of strategy to maximize the expected benefits during the planning horizon under some budget constraints. Since imprecision and flexibility are encountered in making portfolio decisions, a fuzzy programming model is proposed here to optimize portfolio decisions in an uncertain environment.

Notation

$n$  the total number of candidate assets

$\tilde{B}_t$  the flexible budget available for stage  $t$

$r_{it}$  fuzzy future value of candidate asset  $i$  at stage  $t$

$x_{it}$  share of asset  $i$  at stage  $t$

$b_{it}$  share of buying asset  $i$  at stage  $t$ ,  $c_{it}(b_{it})$  is buying cost

$s_{it}$  share of selling asset  $i$  at stage  $t$ ,  $c_{it}(s_{it})$  is selling cost  
 $I_{pt}$  the return of portfolio at stage  $t$   
 $W_0$  the initial wealth  
 $I_t$  the given return of an investor  
 $I'_t$  degree of tolerance for deviation, equivalents to risk

The optimal model:  $\max \prod_{t=1}^T (1+I_{pt})W_0$ , where

$$I_{pt} = \sum_{i=1}^n [r_{it}(x_{it} + b_{it} - s_{it}) - c_{it}(b_{it}) - c_{it}(s_{it})]$$

$$s.t. \begin{cases} |I_{pt} - I_t| \leq I'_t \\ \sum_{i=1}^n [c_{it}(b_{it}) + c_{it}(s_{it})] \leq \tilde{B}_t \\ x_{it}, b_{it}, s_{it} \geq 0 \end{cases}$$

In fact,  $x_{it} = x_{i,t-1} + b_{it} - s_{it}$ . And for general, For simple, given the original respective share are  $\frac{1}{n}$ , the cost of buying and selling asset is proportional, i.e., for example, transaction costs function is  $C(b_{it}) = 0.008b_{it}$  and  $C(s_{it}) = 0.008s_{it}$ ,  $I_t = 0.1, I'_t = 0.02$ . The purpose is to maximize wealth, and the first constraint means risk, the second constraint means budget.

3.3 QUALITATIVE POSSIBILITY THEORY TO TRANSFORMATION

In this paper, the approach of Inuiguchi and Ramik [10] has been extended with qualitative possibility theory to handle the fuzzy constraint containing both uncertain and flexible parameters. Then the fuzzy portfolio selection model is transformed into a linear programming model which can be solved by an optimization technique.

Consider an inequality constraint (first constraint) of the above problem, it can be divided to two inequalities:  $I_{pt} \leq 0.12, I_{pt} \geq 0.08$ . For  $I_{pt} \leq 0.12$ , it equals to

$$\sum_{i=1}^n [r_{it}(x_{it} + b_{it} - s_{it}) - 0.008b_{it} - 0.008s_{it}] \leq 0.12.$$

Set  $x'_{it} = x_{it} + b_{it} - s_{it}$ ,

$$b'_i = \sum_{i=1}^n (0.008b_{it} + 0.008s_{it}) + 0.12$$

, then the above constraint is transferred to  $\sum_{i=1}^n r_{it}x'_{it} \leq b'_i$ .

If the decision maker feels that the satisfaction degree of the constraint needs to be greater than or equal to  $\lambda_i$ , the constraint can be reformulated based on pessimistic criterion according to Eq. (3):

$$C(r_{1t}x'_{1t} + r_{2t}x'_{2t} + \dots + r_{nt}x'_{nt}, b'_i) \geq \lambda_i$$

Then the constraint can be transformed into  $\sum_{i=1}^n r_{it}^{c2}x'_{it} + \lambda_i \sum_{i=1}^n r_{it}^r x'_{it} \leq b'_i$ , where  $r_{it}^{c2}$  is the right modal values of  $r_{it}$ , and  $r_{it}^r$  is their right spreads, respectively. Similarly, we can transform another constraints  $I_{pt} \geq 0.08$  and  $\sum_{i=1}^n [c_{it}(b_{it}) + c_{it}(s_{it})] \leq \tilde{B}_t$  into a crisp objective function. So far, all the constraints are transformed to linear programming.

Next for the fuzzy objective function. Suppose that it is satisfied for an investor when the satisfaction degree of the objective function should be greater than or equal to  $\gamma$ .

$$\max \nu, \quad s.t. C(W_T, \nu) \geq \gamma, \quad \text{where}$$

$$W_T = \prod_{t=1}^T (1+I_{pt})W_0.$$

In order to avoid multiplication, we logarithmic the target function

$$\ln(W)_T = \sum_{t=1}^T \ln \left( 1 + \sum_{i=1}^n (r_{it}x'_{it} - 0.008(b_{it} + s_{it})) \right) + \ln(W_0).$$

So target of maxing the return equals to  $\max \sum_{t=1}^T \ln \left( 1 + \sum_{i=1}^n (r_{it}x'_{it} - 0.008(b_{it} + s_{it})) \right)$ .

Let  $C^{c1} = \sum_{t=1}^T \ln \left( 1 + \sum_{i=1}^n (r_{it}^{c1}x'_{it} - 0.008(b_{it} + s_{it})) \right)$  and

$$C^l = \sum_{t=1}^T \ln \left( 1 + \sum_{i=1}^n (r_{it}^l x'_{it} - 0.008(b_{it} + s_{it})) \right).$$

The target function is equivalent to  $\max \nu$ ,  $s.t. C^{c1} - \gamma C^l \geq \nu$ .

Since the problem is the maximization problem, the left edge of  $C(W_T, \nu)$  is used based on pessimistic criterion to determine the satisfaction degree of the portfolio value that is greater than  $\gamma$ .

4 Illustrated example

This section presents an example of portfolio selection problem to illustrate the approach developed. The investment has three stages. The preferred budgets for stages 1, 2, and 3 are described in fuzzy numbers (in millions): (0, 271.2, 0, 40), (0, 984.9, 0, 200), and (0, 1975.8, 250), respectively. That is  $B_1 = (0, 271.2, 0, 40)$ ,  $B_2 = (0, 984.9, 0, 200)$ ,  $B_3 = (0, 1975.8, 0, 250)$ .

Let the target satisfaction degree of objective function ( $\gamma$ ) be set to 0.95 and risk, budget ( $\lambda$ ) are set to 0.9 for all  $t$  and  $i$ , respectively. Table 1 lists the uncertain return for three stages in fuzzy numbers.

TABLE 1 Fuzzy return for 3 assets

Asset no	Fuzzy return		
	Stage 1	Stage 2	Stage 3
1	(0.3,0.3,0.045,0.045)	(0.5,0.5,0.075,0.075)	(0.45,0.45,0.067,0.067)
2	(0.1,0.1,0.015,0.015)	(0.35,0.35,0.011,0.011)	(0.2,0.2,0.015,0.015)
3	(0.1,0.1,0.015,0.015)	(0.75,0.75,0.112,0.112)	(0.5,0.5,0.015,0.015)
4	(0.55,0.55,0.075,0.075)	(0.65,0.65,0.0975,0.097)	(0.17,0.17,0.025,0.025)
5	(0.2,0.2,0.03,0.03)	(0.85,0.85,0.012,0.012)	(0.2,0.2,0.03,0.03)

From solving the model, the dynamic portfolio is:

At stage 1,  $b_{11} = 0.7203$ ,  $a_{21} = a_{31} = 0$ ,  $a_{41} = 0.314$ ,  $a_{51} = 0$   
and  $s_{11} = s_{21} = s_{31} = 0$ ,  $s_{41} = s_{51} = 0$ .

At stage 2,  $a_{12} = 0$ ,  $a_{22} = 0.577$ ,  $a_{32} = 0$ ,  $a_{42} = 0.027$ ,  $a_{52} = 0$   
and  $s_{21} = s_{22} = s_{32} = 0$ ,  $s_{42} = 0$ ,  $s_{52} = 0.516$ .

At stage 3,  $a_{51} = 0.183$ ,  $a_{52} = 0.204$ ,  $a_{53} = a_{54} = a_{55} = 0$   
and  $s_{51} = 0$ ,  $s_{52} = 0$ ,  $s_{53} = 0.1265$ ,  $s_{54} = s_{55} = 0$ .

## 5 Conclusion

This paper developed a fuzzy multi-portfolio selection model to determine invest strategy that maximizes the

target portfolio value while there is lack of reliable information. The fuzzy portfolio selection model developed was able to handle both uncertain and flexible parameters and the proposed possibilistic transformation method can convert the model into a crisp mathematical model, which can be solved by linear programming.

## Acknowledgment

This research is supported by Department of Education Topics of Hunan Province (NO 12C0749) and scientific research fund of Hunan provincial education department (NO 2A077).

## References

- [1] Markowitz H 1952 Portfolio selection *The Journal of Finance* **7**(1) 77-91
- [2] Hakansson N H 1971 Multi-period mean-variance analysis: Toward a general theory of portfolio choice *The Journal of Finance* **26**(4) 857-84
- [3] Li D, Ng W L 2000 Optimal dynamic portfolio selection: multi-period mean-variance formulation *Mathematical Finance* **10**(3) 387-406
- [4] Östermark R 1996 A fuzzy control model (FCM) for dynamic portfolio management *Fuzzy Sets and Systems* **78**(3) 243-54
- [5] Sadjadi S J, Seyedhosseini S M, Hassanlou Kh 2011 Fuzzy multi period portfolio selection with different rates for borrowing and lending *Applied Soft Computing* **11**(4) 3821-6
- [6] Huang X X, Qiao L 2012 A risk index model for multi-period uncertain portfolio selection *Information Sciences* **217**(25) 108-16
- [7] Liu S T 2011 A fuzzy modeling for fuzzy portfolio optimization *Expert systems with application* **38** 13803-9
- [8] Watada J 1997 Fuzzy portfolio selection and its applications to decision making *Tatra Mountains Mathematical Publications* **13** 219-48
- [9] Dubois D, Prade H 1999 Qualitative possibility theory and its applications to constraint satisfaction and decision under uncertainty *International Journal of Intelligent Systems* **14**(1) 45-61
- [10] Inuiguchi M, Ramik J 2000 Possibilistic linear programming: a brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem *Fuzzy Sets and Systems* **111**(1) 3-28

## Authors



Xing Yu, born on February 15, 1981, in Xianning City of Hubei province

**Current positions, grades:** Hunan university of humanities, Science and technology (China); Lector

**Professional interests:** Applicate mathematics; Finance model

**Research interests:** the optimal portfolio model; mathematical model

# A novel method of user interest drift detection engaging in individual background factors

Chonghuan Xu\*

College of business Administration Zhejiang Gongshang University, Hangzhou, China

Center for Studies of Modern Business, Zhejiang Gongshang University, Hangzhou, China

Received 1 March 2014, www.tsi.lv

## Abstract

Personalized service tends to be an emerging challenge in the field of interest mining on e-commerce platform, the issues of which include how to integrate the user's individual background factor, to hiddenly attain portal user interest behaviour, and to mine interest drift pattern. According to user interest drift problem of personalized service in network, this paper explains the user interest through an integration of individual background factor, user behaviour and interest. Meanwhile, it recommends the fuzzy logic thought to explain its impact factor weights comprehensively in order to reflect the level of the user interest on theme. And, it establishes the Hidden semi-Markov Model via user browsing path to detect whether the interest is drifted or not. Finally, the method is proved to be accurate through the experiment analysis.

*Keywords:* user interest, HSMM, background factors, drift detection

## 1 Introduction

The quality of personalized service depends on the accuracy of user interest mastered by the system. However, the user's interest will change ranging from time to other surrounding factors. In other words, the user interest drift occurred. The key issue of personalized service is how to identify the changes of user's interest accurately and timely in order to provide the interested service content.

Global scholars have made some research for user interest drift. Grabtree and Soltysiak [1] used the time window approach to solve the problem of user interest drift, whereas it only used one sample to train the user model as the recent visiting. Wu *et al.*[2] proposed a Semi-supervised classification algorithm for data streams of user interesting with concept drifts and Unlabeled data(SUN). Maloof and Michalski [3] used a forgetting mechanism to attenuate the sample. Koychev and Schwab[4] presented a method for dealing with drifting interests by introducing the notion of gradual forgetting. Ahmed *et al.*[5] proposed a novel framework to introduce a very useful measure, called frequency affinity, among the items in a HUP and the concept of interesting HUP with a strong frequency affinity for the fast discovery of more applicable knowledge. Nicoletti *et al.* [6] presented a novel method for topic detection of user interesting from online informal conversations.

This paper will continue to study the drift of the user's interest, not only from the perspective of way to proceed, but also from the perspective of the individual user background. And it would build Hidden Semi-Markov

Model to detect whether the user interest drift or not, in light of the user interest combination of background, user behaviour and content.

## 2 User interest description and mapping

### 2.1 BASIC DEFINITIONS

Definition 1. The user interest content set  $UIC$  is the collection of interest content after classification of visited resource about all users in the website:  
 $UIC = \{P_1, \dots, P_l\} \cup \{L_1, \dots, L_m\}$   
 $\cup \{T_1, \dots, T_n\} = \{UIC_1, UIC_2, \dots, UIC_M\}$ .

Where  $P$  is a web site component channels;  $L$  is a hyperlink content;  $T$  is a tag page;  $UIC$  is the classification of interest content used by the concept layered approach [7], and it has a corresponding interest concept set:  $\Sigma = \{\sigma_x \mid 1 \leq x \leq Z\}$ ,  $\exists UIC \mapsto \sigma_x$ ,  $\sigma_x$  is a characteristic concept of interest content,  $\mapsto$  means the mapping relationship from interest content to characteristic concept.

Definition 2. The user background factors set  $UBE$  is a collection of various background factors existed in individual user  $u$ , mainly containing Region, Gender, Age, Marriage, Education and Income. It is defined as a user background set:  $UBE = \{Region, Gender, Age, Marriage, Education, Income\}$ .

Definition 3. The User interest behaviour set  $UIB$  is a collection of all possible behaviour operations when  $u$  visits the  $UIC$  on the page of website. In this paper, the behaviour data is divided into several types as follows

\* *Corresponding author* e-mail: talentxch@gmail.com

when users are browsing the web: marking behaviour, such as increasing the bookmark (Book), saving the page (Save), etc.; operational behaviour, such as dragging the scroll bar (Scroll), visiting the page of time (Times), etc.; link behaviour, that is, whether click a hyperlink when you are browsing a page (Click). A set of interest behaviours is defined as follows:  $UIB = \{Book, Save, Scroll, Times, Click\}$ .

Definition 4. Let the accessing process of user  $u$  among the session of time fragment  $T$  as a accessing sequential transaction  $tr$ , defined as a tuple:  $\{tr.u, (tr.content_1, tr.time_1, tr.background_1, tr.behaviour_1), \dots, (tr.content_p, tr.time_p, tr.background_p, tr.behaviour_p)\}$ . In which,  $tr.u \in U$  denoted accessing user; Four tuples  $(tr.content, tr.time, tr.background, tr.behaviour_i)$  express as the per accessing operation of user,  $tr.content \in UIC$  denotes the detail of interest in content objection,  $tr.time(tr.time_p - tr.time_1 \leq T)$  denotes accessing timestamp;  $tr.background \in UBE$  expresses as the specific background factors of the user;  $tr.behaviour \in UIB$  expresses as the interests of specific behaviours of users. Therefore, consisting all accessing transaction  $tr$  to accessing transaction set of user in visiting the website by sequential session times:  $TR_u = \{tr_i | 1 \leq i \leq |TR_u|\}$ ,  $|TR_u|$  as total number of sessions of the user.

2.2 BACKGROUND

There are various differences among different backgrounds users, and different levels of interest in commodity. In light of the difference of their ages, occupations, backgrounds, interests, they focus on different emphasis on the information systems, and often merely focus on a subset of resources in specific areas. Internet users' interest properties are mainly determined by external factors and internal factors. External determinants include: cultural factors, social factors and family factors, while internal determinants include: life-cycle stages, occupational factors, income, lifestyle, personality factors, self-concept and psychological factors, etc., both these various factors of which will be integrated and have an influence on network behaviour of the users. In this paper, using the geographic, gender, age, marital status, educational background and income which are key impacts on the users' interest as indexes, combining with user behaviour and characteristics of its interest to obfuscate the content of user and to get their degree of interest value.

This paper introduces the idea of fuzzy logic to describe the joint mapping based on the factor weight in backgrounds and behaviour of interest.

Let the user's individual background factors be expressed as  $B_u = (u, background)$ .  $FB_B = Relation(B_u, UIC \cup UIB)$  represents the fuzzy relationship between  $B_u$  and  $UIC \cup UIB$  on the domain of  $B_u \times (UIC \cup UIB)$ , where  $u$  describes the process of interaction the user to access the background of the individual behaviour of interest.

The definition of  $W_B(content_k) \in [0,1]$  is a normalization that reflects the individual background  $FB_B$  weight.

Let  $FB_{TR} = Relation(TR_u, UIC \cup UBE \cup UIB)$  said  $TR_u \times (UIC \cup UBE \cup UIB)$  domain  $u_{TR}$  and  $UIC \cup UBE \cup UIB$  relationship between the fuzzy, interactive access to the process  $u$  described the behaviour of interest characteristics and evaluate the impact of the definition of  $W_{TR}(content_k) \in [0,1]$  for the normalized reflected in the behaviour that  $FB_{TR}$  interest in weight.

Let user  $u$  interested in browsing the contents of the purchase process of change expressed as navigation path sequences:  $S_u = \{tr_i.seq\} (tr_i \in TR_u, |S_u| = |u_{TR}|)$ ,  $tr.seq$  each order record requests, the interest in the content of  $tr_i, content_k$  where the ranks of the position. That  $FB_L = Relation(S_u, UIC \cup UBE \cup UIB)$ , said  $S_u \times (UIC \cup UBE \cup UIB)$  domain  $S_u$  and  $UIC \cup UBE \cup UIB$  fuzzy relationship between describing the process of  $u$  interested in interactive access to content and degree of concern, the definition of  $W_L(content_k) \in [0,1]$  is the normalization of interest that reflect the content of  $FB_L$  weight.

Thus, each combination of background, interests behaviour and interest Description in weight of user  $u$  can be expressed as  $W(content_k)$ :

$$W(content_k) = \theta_1 W_B(content_k) + \theta_2 W_{TR}(content_k) + \theta_3 W_L(content_k) \text{ Where } \theta_1 + \theta_2 + \theta_3 = 1 (\theta_1, \theta_2, \theta_3 \in [0,1]).$$

3 User interest drift mechanism based on HSMM

3.1 HIDDEN SEMI-MARKOV MODEL

A Hidden semi-Markov Model (HSMM) is an extension of HMM by allowing the underlying process to be a semi-Markov chain with a variable duration or sojourn time for each state. Therefore, in addition to the notation defined for the HMM, the duration  $d$  of a given state is explicitly defined for the HSMM. State duration is a random variable and assumes an integer value in the set  $D = \{1, 2, \dots, D\}$ . The important difference between HMM and HSMM is that one observation per state is assumed in HMM while in HSMM each state can emit a sequence of observations. The number of observations produced while in state  $i$  is determined by the length of time spent in state  $i$ , i.e., the duration  $d$ .

A parameter of the HSMM [8], can be expressed as a six-tuple:  $\lambda = \{N, M, \pi, A, D, B\}$  where  $N$  indicates the number of states;  $M$  is the number of observations;  $\pi = \{\pi_i\}$ ;  $A = \{a_{ij}\}$ ;  $B = \{b_j(k)\}$ ;  $P = \{p_i(d)\}$ .  $o_t$  represents observation of the  $t$  vector, which includes the first  $t$  requests objects and  $r_t$  between  $r_t$  and  $r_{t-1}$  with the time interval  $\tau_t$ , that is  $o_t = (r_t, \tau_t)$ . Representatives from the first  $a$  one to one observation vector  $b$  sequence, represents the observation vector sequence. The length  $Tst$  represents  $t$  time state.  $\epsilon_t$  represents the current state of the output will be the number of observations,  $1 \leq t \leq T$ .

### 3.2 USER INTEREST BEHAVIOUR

For the user's interest drift, this will create two hidden semi-Markov models. One is used to describe the stable interest and behaviour profile of one or a group of users, the other HSMM is used to outline the behaviours of transferred interests. To get classified in accordance with the path sequence of the user access behaviour, and make sure the observation value set corresponding to every state according to the training data (Unchanged behaviour of user interest) of unchanged user interest behaviour. Take the path sequence to access the web page of users as the basis of the classification of the drift behaviour patterns. Make sure the interest behaviour patterns have the similar path sequence into the same category.

This paper selected the following two observations to describe the user's browsing behaviour:

1) the path sequence of the user access to Web browsing;

2) the time interval between the two adjacent pages. As the user's browsing behaviour is usually from one page to another, therefore, let us assume the user's browsing actions are consistent with the characteristics of Markov chain, and can describe the chain from a state perspective. The set of all states is expressed as  $S = \{S_1, S_2, \dots, S_N\}$ , the corresponding set of observations expressed as  $V = \{v_1, v_2, \dots, v_M\}$ , discrete integer seconds time interval, set as  $I = \{1, 2, \dots\}$ . For the typical user's browsing behaviour of a class, the number of its navigation path link is another random variable, which can be considered state of the output in a given number of observations, the set of which is represented as  $\{1, \dots, D\}$ . The sequence of user browsing path is expressed as user browsing Web content objects and user time interval between  $rt-1$  and  $rt$  from one page to another page.  $O$  is a model of two-dimensional sequence of observations.  $B = \{b_i(v, q)\}$  is the output probability matrix model, where as a given state  $i \in S$ ,  $b_i(v, q)$  matches the state of the user in a page  $r_t = v \in V$  and a page with the previous time interval  $T_0$   $\tau_t = q \in I$  probability.  $P = \{p_i(d)\}$  represents the output under a given state  $i$  the number of observations for the  $d \in \{1, \dots, D\}$  of probability, and meets  $\sum d p_i(d) = 1$ , that is,  $P$  is the HSMM model the status of the residence time probability matrix.  $\pi = \{\pi_i\}$  represents the initial state probability vector, where  $\pi_i$  represents the initial state  $i \in S$  the probability.  $A = \{a_{ij}\}$ , represents the state transition probability matrix, where  $a_{ij}$  is transferred from state  $i \in S$  to  $j \in S$  the probability. The user's behaviour is an important record of interest which is defined as:  $U_{interest} = \{user, timestamp, content, background, behaviour\}$ .

### 3.3 USER INTEREST DRIFT DETECTION

First, the user browsing behaviour data from collection system are used as a sequence of observations, after pretreatment to form a training sequence to train the

model. After the model parameters are determined, the model can be used for drift detection. After a pre-measured data is required by observations, by calling HSMM algorithm module, it can calculate the average of the number of contingent probabilities. Then, user interest module in the same sentence will be to get the same interest in the value of user behaviour. If the value of the user's interest in the normal range, the user data will be added to the training data set used in the background update HSMM model parameters, and enter the service queue; Otherwise, the user will be considered to be interested in drift, and to other modules (interested in change processing module) for processing.

Drift detection implementation is shown in Figure 1.

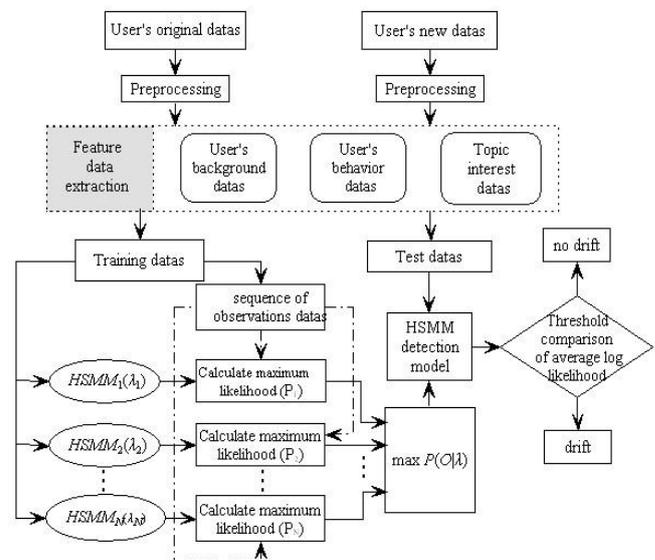


FIGURE 1 Interest drift detection

In this paper, the training data set of all sequences, the average number of contingent probabilities of the mean  $lkh$  normal behaviour as a reference point. An observation sequence can be pre-defined length threshold  $T_0$ , when the user's browsing path  $l$  to  $T_0$ , the user can calculate the average number of contingent on the probability of  $lkh^{(l)}$ . By comparing  $lkh^{(l)}$  and  $lkh$ , you can get the user relative to the model deviation. The smaller the absolute value of difference is, the lower the deviation is, and the smaller the degree of interest drifts is.

### 4 Performance analysis

The testing data is extracted from a shopping site, and this paper obtains the background factors and the most interesting of the three themes according to the user's registration information and historical data. Background factors mainly contain region, age, marital status, income, and education; while theme is divided into clothing, ornament, beauty, digital, home, motherhood, food, sports and entertainment. The user's background factors and the initial interest theme form are shown in Table 1, where the theme is arranged by the interest weight. In other words, interest topic 1 > interest topic 2 > interest

topic 3, and the empty means the interest does not exist. This paper allows user access to 150 pages, and then regards user interest category as the training set. Through user awareness, human factors analysis, the user browsing the page follows the regular pattern, in a total of nine categories, the main user interest categories of which are "Digital", "food".

We adopt the fuzzy logical thinking to describe user's interest weight jointly by providing a fuzzy processing of the user's background factors, browsing behaviour and the subject content of interest. These specific methods are defined as definition 6 to 8. The interest weight reflects

the level of the user interest in some subject as well as tests on the user interest drift if the weight changes. For the online consumers, the paper analyses the nine interest subjects to judge the interest drift. We use the 1st to 15th user visit, and each visit contains 100 samples. Table 2 takes the user1 as an example, the first line of each concept is the observation sequence of change process of user interest, where the number in the table presents the probability of this characteristic and empty presents no appearance. Among the 15 visits of user1, the original interest subjects (clothing, accessories and entertainment) change into (maternal, infant and clothing).

TABLE 1 User's background factors and the initial theme interest form (part)

Number	Gender	Region	Marriage	Income	Age	Education	Interest topic 1	Interest topic 2	Interest topic 3
1	female	eastern	No	8000	27	undergraduate	clothing	digital	amusement
2	male	northeast	Yes	6000	26	master	digital	sports	
3	female	eastern	Yes	7000	27	undergraduate	ornament	hairdressing	food
4	female	eastern	No	5500	33	undergraduate	maternal-infant	household	clothing
5	male	south	No	14000	31	doctor	hairdressing	amusement	
6	female	south	Yes	4500	38	undergraduate	digital	household	sports
7	female	eastern	No	5000	40	undergraduate	household	maternal-infant	sports
8	female	northeast	Yes	14000	46	undergraduate	clothing	hairdressing	food
9	male	southwest	No	14000	31	doctor	clothing	sports	amusement
10	male	eastern	Yes	8500	42	senior	digital	sports	amusement
11	male	south	Yes	15000	35	doctor	clothing	household	
12	male	eastern	No	7500	39	master	clothing	digital	household
13	male	eastern	No	23000	30	doctor	digital	food	sports
14	female	northeast	Yes	9000	28	undergraduate	ornament	clothing	hairdressing
15	female	north	No	15000	25	master	digital	ornament	sports
16	female	eastern	Yes	6500	46	senior	clothing	hairdressing	
17	male	north	No	6000	42	senior	clothing	hairdressing	amusement
18	male	eastern	Yes	4500	29	senior	digital	amusement	food
19	female	northeast	No	6500	27	undergraduate	clothing	ornament	sports
20	female	southwest	No	15000	26	doctor	amusement	clothing	food
21	male	eastern	Yes	4500	38	senior	household	digital	
22	female	south	No	8000	29	master	clothing	amusement	food
23	female	north	No	4000	27	undergraduate	amusement	ornament	amusement
24	female	eastern	Yes	3000	30	senior	sports	clothing	food
25	male	eastern	Yes	12000	34	master	amusement	digital	household
26	female	eastern	No	8000	34	undergraduate	clothing	household	food
27	male	southwest	Yes	15000	35	doctor	digital	maternal-infant	sports
28	female	southwest	No	6000	30	undergraduate	clothing	amusement	
29	male	north	No	13000	45	master	digital	household	sports
30	female	eastern	Yes	5000	36	senior	hairdressing	clothing	maternal-infant
31	male	northeast	Yes	8000	30	undergraduate	clothing	digital	ornament

TABLE 2 User interest sequence and chances of its weight

Interest	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
clothing	3.44	3.43	2.65	2.21	2.32	3.11	2.76	1.94	2.02	1.88	2.36	2.58	2.61	2.29	2.39
ornament	2.31	2.62	1.98	0.96	0.83	0.72	0.2								
hairdressing				0.86	1.32	1.56	1.63	0.94	0.32	0.18					
digital															
household									0.2	0.3	0.4	0.4	0.3	0.4	0.4
maternal-infant							0.03	0.04	0.05	0.98	1.32	2.63	2.91	3.67	5.41
food	0.36	0.36	0.52		1.33			1.89	1.75		1.98		0.43		0.96
sports					0.53	0.92	1.31	1.62	1.18	1.53	1.21	1.67	2.28	3.02	3.52
amusement	1.89	2.64	2.53	1.62	0.98	0.67									

Define the accuracy of shift algorithm: Detection accuracy = correct identification of a particular interest in the number of users / number of a particular user interest. The accuracy of detection is used to measure the fit degree of the adjustment of the user's interest and the actual change in the interest by the algorithm. The higher accuracy indicates that the more algorithms meet the change of user's actual interest. There are three respective algorithm accuracy comparison (HSMM model, sliding windows, and progressive forgetting), and thus the results are shown in Figure 2.

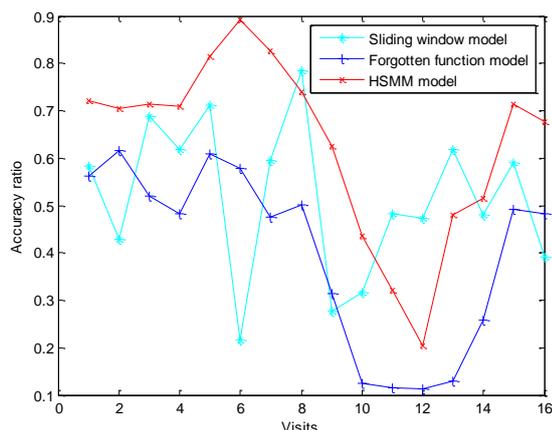


FIGURE 2 Comparison of drift detection methods

## 5 Conclusions

User interest extraction and user drift detection have great significance on interest mining applications. This paper explains the joint mapping based on the factor weight in interest background and interest behaviour, taking into account of the user's explicit and implicit interests as well as adopts HSMM model to detect whether or not the user's interest is drifted. These methods can express the user's personal interest comprehensively, and improve the accuracy and predictability on the basis of personal interest mining. There requires, however, a further discussion on how to improve the situation of interest drift after the detection.

## Acknowledgments

This work was supported in part by Ministry of Education, Humanities and Social Sciences project (Grant No. 13YJ CZH216), Natural Science Foundation of Zhejiang Province (Grant No. LQ12G01007), Specialized Research Fund for the Doctoral Program of Higher Education of China (No. 20093326120004), Zhejiang Science and Technology Plan Project (No. 2010C33016, 2012R10041-09), and the Key Technology Innovation Team Building Program of Zhejiang Province (No. 2010R50041).

## References

- [1] Crabtree Barry, Soltysiak Stuart J 1998 Identifying and tracking changing interests *International Journal of Digital Libraries* Springer Verlag 2 38-53
- [2] Xindong Wu, Peipei Li, Xuegang Hu 2012 Learning from concept drifting data streams with unlabeled data *Neurocomputing* 92(1) 145-55
- [3] Marcus A Maloof, Ryszard S Michalski 2000 Selecting Examples for Partial Memory Learning *Machine Learning*, 41(1) 27-52
- [4] Ivan Koychev, Ingo Schwab 2000 Adaptation to drifting user's interests *Proceedings of ECML2000 Workshop: Machine Learning in New information Age 2000*
- [5] Chowdhury Farhan Ahmed, Syed Khairuzzaman Tanbeer, Byeong-Soo Jeong, Ho-Jin Choi 2011 A framework for mining interesting high utility patterns with a strong frequency affinity *Inform Sciences* 181(21) 4878-94
- [6] Matias Nicoletti, Silvia Schiaffino, Daniela Godoy 2013 Mining interests for user profiling in electronic conversations *Expert Syst Appl* 40(2) 638-45
- [7] Kim Hyung-rae 2005 Learning implicit user interest hierarchy for Web personalization: Florida Institute of Technology
- [8] Yu Shun-Zheng 2010 Hidden semi-Markov models *Artif Intell* 174(2) 215-43

## Authors



**Chonghuan Xu, born on November 14, 1983, Hangzhou, China**

**Current position, grades:** lecturer, Zhejiang Gongshang University.

**University studies:** B.S. and M.S. degrees in Computer and Information Engineering from Zhejiang Gongshang University, Hangzhou.

**Scientific interest:** Operations research, electronic commerce, data mining.

**Publications:** 15.

# Empirical study of C2C logistics customer satisfaction based on AHP and FCE

Jing An\*, Bo Xu

School of Economics & Management, Changzhou Institute of Technology, Tongjiang South Road, No. 299, Changzhou, China

Received 1 March 2014, www.tsi.lv

## Abstract

Evaluation system of C2C logistics customer satisfaction was mainly studied here. Firstly, three primary indexes and nine secondary indexes were constructed as the components of the evaluation system. Then the integrated use of AHP and FCE was carried on taobao.com as the empirical research. The result indicates that the overall logistics customer satisfaction of taobao.com is good. But there is still space for improvement. Through this study, logistics customer satisfaction of taobao.com is to improved. In addition, references are provided for other C2C enterprises and logistics companies. The common development of them is therefore promoted.

Keywords: C2C logistics customer satisfaction, AHP, FCE

## 1 Introduction

As more and more people enjoy online shopping, in recent years China e-commerce has gained explosive growth. Online shopping boom is arriving. China has also rapidly changed from the distributed shopping to scale shopping. The running condition of the whole country's postal service announced by China Post Office, says that the express industry has gained a growth more than 27% for five years in a row, over half of which is from e-commerce, indicating that e-commerce logistics is very important in the logistics industry. Meanwhile as the only face-to-face process in the whole e-commerce marketing, e-commerce logistics becomes one important factor in the successful development of e-commerce [1, 2].

How to provide effective e-commerce logistics service for customers, is not only the problem e-commerce merchant cares, but also the logistics company concerns [3]. Whether customer is satisfied with the logistics services, not only affects the cooperation between merchants and customers, but also affect the business cooperation between merchants and logistics companies [4]. In current social situation, the construction of the express logistics system in our country is still in its infancy, which has not yet formed its mature system. At the same time, employee diathesis of the express logistics is generally not high, which becomes the factor that makes customer dissatisfied [5]. Scholars and researchers have made many relative studies. But current research is mainly qualitative research, there is rarely quantitative research. In addition, relative research of logistics customer satisfaction of C2C is a little.

By quantitative analysis of customer satisfaction logistics of C2C companies through AHP and FCE, they can find explicitly, which aspects they did not well. So in

the future these aspects are improved to serve customers better, thereby to increase their marketing activities. It can also standardize and improve those marketing activities, so to give impetus to the development of C2C enterprises and logistics industry.

## 2 Construction of evaluation index system

Through analysis of related literature and combining with the practical research experience [6-11], following the principle of science, integrity, operability, comparability, and analysis of qualitative and quantitative, the evaluation index system of C2C logistics customer satisfaction this study built is as follows.

### 2.1 TARGET LAYER

It is the highest level, as C2C logistics customer satisfaction (A).

### 2.2 PRIMARY INDEXES

The target layer is decomposed into three primary indicators, respectively are delivery speed and price ( $A_1$ ), corporate image ( $A_2$ ), and quality of service ( $A_3$ ), as shown in Figure 1.

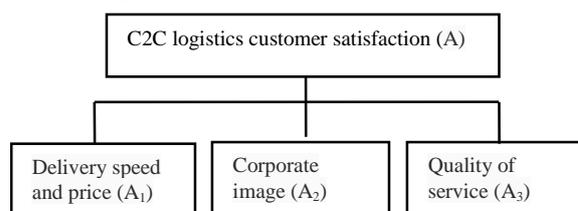


FIGURE 1 Primary evaluation indexes of C2C logistics customer satisfaction

\* Corresponding author e-mail: anj@czu.cn

## 2.3 SECONDARY INDEXES

Three primary indexes are decomposed into nine secondary indicators, as shown in Table 1.

TABLE 1 Evaluation index system of c2c logistics customer satisfaction

Evaluation target	Primary indicators	Secondary indicators
C2C logistics customer satisfaction (A)	Delivery speed and price (A <sub>1</sub> )	Speed of delivery (B <sub>1</sub> )
		Cargo tracking and feedback (B <sub>2</sub> )
		Logistics fee (B <sub>3</sub> )
	Corporate image (A <sub>2</sub> )	Enterprise brand (B <sub>4</sub> )
		Credibility of enterprises (B <sub>5</sub> )
		Enterprise staff quality (B <sub>6</sub> )
	Quality of service (A <sub>3</sub> )	Packing integrity (B <sub>7</sub> )
		Treatment of emergency (B <sub>8</sub> )
		Integrity of goods (B <sub>9</sub> )

Delivery speed and price (A<sub>1</sub>) consists of three secondary indicators: 1) the speed of delivery (B<sub>1</sub>) refers to the time from purchase to receive goods for the customers; 2) the cargo tracking and feedback (B<sub>2</sub>) refers to the network information timeliness and validity during the distribution process of goods; 3) logistics fee (B<sub>3</sub>) is to pay the price of the logistics by customers.

Corporate image (A<sub>2</sub>) consists of three secondary indicators:

1) enterprise brand (B<sub>4</sub>) refers to the business philosophy, corporate culture, corporate values and attitude to consumers, etc.;

2) credibility of enterprises (B<sub>5</sub>) mainly refers to how the enterprise credit is;

3) enterprise staff quality (B<sub>6</sub>) refers to the comprehensive quality of the main employees, which consists of the basic quality, professional quality and political quality.

Quality of service (A<sub>3</sub>) consists of three secondary indicators:

1) packing integrity (B<sub>7</sub>) refers to whether outer packaging of goods is complete;

2) treatment of emergency (B<sub>8</sub>) refers to whether logistics companies actively cooperate with the customers to deal with unexpected problems;

3) integrity of the goods (B<sub>9</sub>) refers to the quantity and quality of the goods customers received has not been changed.

## 3 Evaluation method

C2C logistics customer satisfaction is influenced by many factors [12-15], so to build index system need to consider from many aspects, in order to objectively and comprehensively reflect the customer satisfaction level of C2C logistics. This paper uses the analytic hierarchy process (AHP) to determine the weights of each layer of the index system, and makes comprehensive evaluation combining with Fuzzy Comprehensive Evaluation (FCE).

## 3.1 ANALYTIC HIERARCHY PROCESS

AHP (Analytic Hierarchy Process, AHP for short), is a multi-objective decision-making method used to handle limited solutions. Its basic method is to establish a hierarchy structure model. After the hierarchy model is set up, each two elements of every layer index are compared, so as to construct the judgment matrix. To fill the results of comparison in the judgment matrix, eigenvalue and eigenvector of the judgment matrix are solved, and then the weighting coefficient of the importance of each target is determined.

## 3.2 FUZZY COMPREHENSIVE EVALUATION

Fuzzy Comprehensive Evaluation (referred to as FCE) is a kind of effective multi-factor decision method, making comprehensive evaluation on objects affected by various factors. The evaluation result is based on a fuzzy set with quantitative description.

While FCE is used, expert investigation method is chosen. This method is to make expert evaluation questionnaire, each index is concluded according to the experience of experts. Through statistics of a few expert questionnaires, corresponding frequency of each factor is got. Using normalization, the corresponding grade degree of membership of each factor is got, so the single evaluation matrix is obtained, namely the degree of membership = (the number thought factor i belongs to comment j)/(the total number of evaluation experts).

## 4 Empirical study

### 4.1 CASE STUDIES

Here China's largest C2C online retail platform taobao.com is taken as the study case.

Everyday taobao.com has large quantities of customers, it also brings the demand of goods logistics. Sellers on taobao.com come into contact with many logistics companies, and many sellers even cooperate with several logistics companies at the same time. When a customer receives the goods sent from different logistics companies, the feeling of service is also different. The logistics situation of taobao.com is reflected in the following aspects.

The first is about the cargo information tracking problem. After people shopping on the Internet, they will be eager to know when their goods ship and where goods arrive after shipment. However, currently logistics companies have not done well in the aspect of networks. After purchase, a lot of customers check the cargo tracking information on the Internet, it often appears "temporarily unable to query". Therefore, that customers can't have effective understanding and tracking about their goods. Customer's demands are not well met; naturally, customer satisfaction would be affected a lot.

Second is the service of express and speed, and so on, which can be divided into two phenomena. One is that online sellers use one more express companies, when a customer receives the goods from different logistics companies, his feel of service is also not the same. When the difference of the service attitude is large, it will cause customer's dissatisfaction. The other is due to the logistics company itself. Logistics time delay leads to customer's waiting too long and causes their dissatisfaction, which directly affects the logistics customer satisfaction.

The third is the product packaging and damage rate. When customers find their commodity packaging is not complete, it will cause customer dissatisfaction, which even lead to the customer's request to return. Most of the time the outer packaging of goods is intact, but after opening, the inside products are found damaged. Obviously, customers will have strong displeasure, so that the satisfaction will drop sharply.

4.2 EVALUATION PROCESS

Firstly using AHP to respectively determine the weights of all levels of the evaluation system, then to use expert evaluation method, to give comprehensive grade of each index, to get fuzzy matrix of two level indexes, thus can get comprehensive evaluation vector at all levels.

Firstly to establish the judgment matrix of the secondary sub-indexes elements to A<sub>1</sub> (here ten experts are invited to participate in the questionnaire, the raw data in this article is the average value of the experts' scoring), as shown in Table 2. To adopt AHP to calculate the weights at various levels, then are checked through the consistency test.

TABLE 2 Judgment matrix of A<sub>1</sub> and the corresponding weights

A <sub>1</sub>	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	weights
B <sub>1</sub>	1	1	577/300	0.394
B <sub>2</sub>	1	1	64/30	0.408
B <sub>3</sub>	300/577	30/64	1	0.198

The consistency check are: λ<sub>max</sub>=3.001, CI=0.0005, CR=0.00086<0.1.

In the same way, the judgment matrix of the other secondary sub-indexes elements to A<sub>2</sub> and A<sub>3</sub> and the corresponding weights are obtained as shown in Table 3 and Table 4.

TABLE 3 Judgment matrix of A<sub>2</sub> and the corresponding weights

A <sub>2</sub>	B <sub>4</sub>	B <sub>5</sub>	B <sub>6</sub>	weights
B <sub>4</sub>	1	17/10	36/10	0.525
B <sub>5</sub>	10/17	1	57/20	0.342
B <sub>6</sub>	10/36	20/57	1	0.133

The consistency check are: λ<sub>max</sub>=3.009, CI=0.005, CR=0.0086<0.1.

TABLE 4 Judgment matrix of A<sub>3</sub> and the corresponding weights

A <sub>3</sub>	B <sub>7</sub>	B <sub>8</sub>	B <sub>9</sub>	weights
B <sub>7</sub>	1	551/300	727/300	0.512
B <sub>8</sub>	300/551	1	311/300	0.258
B <sub>9</sub>	300/727	300/311	1	0.230

The consistency check are: λ<sub>max</sub>=3.006, CI=0.003, CR=0.006<0.1.

Fuzzy comprehensive evaluation of those indexes by data are made, for which the raw data comes from the results summary of 135 questionnaires and 10 experts questionnaires before, to get the fuzzy matrix which is as follows:

$$R_{A1(3 \times 5)} = \begin{pmatrix} 0.062 & 0.566 & 0.324 & 0.048 & 0.000 \\ 0.076 & 0.434 & 0.379 & 0.097 & 0.014 \\ 0.076 & 0.524 & 0.331 & 0.069 & 0.000 \end{pmatrix}$$

$$R_{A2(3 \times 5)} = \begin{pmatrix} 0.028 & 0.490 & 0.448 & 0.028 & 0.000 \\ 0.056 & 0.480 & 0.410 & 0.040 & 0.014 \\ 0.065 & 0.400 & 0.434 & 0.090 & 0.000 \end{pmatrix}$$

$$R_{A3(3 \times 5)} = \begin{pmatrix} 0.076 & 0.593 & 0.270 & 0.055 & 0.006 \\ 0.041 & 0.379 & 0.476 & 0.090 & 0.014 \\ 0.041 & 0.338 & 0.428 & 0.124 & 0.069 \end{pmatrix}$$

Based on the common matrix multiplication operator ".", fuzzy operation is made. "." means generalized synthesis operator. Matrix synthesis is (Λ, V), that is the elements of B is:

$$b_j = \sum_{i=1}^m + a_i \cdot r_{ij} = (a_1 \cdot r_{1j}) \cdot (a_2 \cdot r_{2j}) \cdot \dots \cdot (a_m \cdot r_{mj}), \quad (1 \leq j \leq n)$$

The evaluation fuzzy sets of the secondary indexes are obtained:

$$B = A \cdot R = (0.394, 0.408, 0.198) \cdot \begin{pmatrix} 0.062 & 0.566 & 0.324 & 0.048 & 0.000 \\ 0.076 & 0.434 & 0.379 & 0.097 & 0.014 \\ 0.076 & 0.524 & 0.331 & 0.069 & 0.000 \end{pmatrix} = (0.076, 0.408, 0.379, 0.097, 0.014)$$

Normalized to get:

$$B_{A1} = (0.078, 0.419, 0.389, 0.100, 0.014)$$

By the same token, the results of the operations of fuzzy evaluation to other factors are:

$$B_{A2} = (0.056, 0.444, 0.405, 0.082, 0.013)$$

$$B_{A3} = (0.060, 0.407, 0.379, 0.099, 0.055)$$

In the same way, the judgment matrix of the primary indexes with application of AHP is got, and the weights are calculated, as shown in Table 5.

TABLE 5 Judgment matrix of A and the corresponding weights

A	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	weights
A <sub>1</sub>	1	107/60	88/30	0.523
A <sub>2</sub>	60/107	1	221/120	0.305
A <sub>3</sub>	30/88	120/221	1	0.172

Details are shown as in Table 6.

TABLE 6 Evaluation results of c2c logistics customer satisfaction

Primary index	weights	Secondary index	weights	fuzzy comprehensive evaluation				
				secondary evaluation vector				
Delivery speed and price (A <sub>1</sub> )	0.523	Speed of delivery (B <sub>1</sub> )	0.394	0.078	0.419	0.389	0.1	0.014
		Cargo tracking and feedback (B <sub>2</sub> )	0.408					
		Logistics fee (B <sub>3</sub> )	0.198					
Corporate image (A <sub>2</sub> )	0.305	Enterprise brand (B <sub>4</sub> )	0.525	0.056	0.444	0.405	0.082	0.013
		Credibility of enterprises (B <sub>5</sub> )	0.342					
		Enterprise staff quality (B <sub>6</sub> )	0.133					
		Packing integrity (B <sub>7</sub> )	0.512					
Quality of service (A <sub>3</sub> )	0.172	Treatment of emergency (B <sub>8</sub> )	0.258	0.060	0.407	0.379	0.099	0.055
		Integrity of goods (B <sub>9</sub> )	0.230					
				primary evaluation vector				
				0.072	0.410	0.374	0.093	0.051

### 4.3 EVALUATION RESULTS

This paper makes fuzzy comprehensive evaluation of logistics customer satisfaction of taobao.com from delivery speed and price, corporate image and quality of service. The results show that according to the principle of maximum membership degree, the value of logistics customer satisfaction of taobao.com is 0.410, the result is better, which means logistics customer satisfaction of taobao.com is ok, but there is still room for improvement.

Combining with Table 6, this paper makes detailed analysis is as follows:

In the primary indexes, the weight of delivery speed and price is the biggest, the value of which is 0.523. According to the principle of maximum membership degree, the value of maximum membership degree of FCE is 0.419, the result of evaluation is well. Under its three secondary indexes, the weight of cargo tracking and feedback is the largest, whose value is 0.408. It is not difficult to find whether can effectively provide cargo tracking and feedback will have huge impact on logistics customer satisfaction.

Under the index of enterprise image, the weight of enterprise brand is the biggest, the value of which is 0.525. From this, it can easily find that customers pay much attention to it. Therefore, how to effectively improve the brand of logistics enterprises will affect the logistics customer satisfaction of taobao.com.

Among the primary indexes, quality of service weights the smallest, whose value is only 0.172. The weight of its subordinate indicators - packaging integrity, treatment of emergency, integrity of goods---respectively is 0.512, 0.258, 0.230. Can say that within service, the influence packaging integrity making on customer's satisfaction is the biggest. According to the principle of maximum membership degree, the value of FCE is 0.407, showing that the result is good. However, taobao.com can do better in service, such as improving the treatment of emergency, integrity of goods, etc.

The consistency check are:  $\lambda_{max}=3.001$ ,  $CI=0.0005$ ,  $CR=0.00086<0.1$ .

To make fuzzy comprehensive evaluation of the primary indexes, the fuzzy evaluation vector obtained is as follows:  $B_A=(0.072, 0.410, 0.374, 0.093, 0.051)$ .

From above analysis, it can conclude that for taobao.com customers value delivery speed and price the most, second is the brand of logistics companies, the last is quality of service. Therefore, for customers after purchase, they particularly value when goods can arrive and how much the price of logistics is. Compared with the entity store shopping, the advantage of online shopping is convenient and cheap, but it also has shortcoming that customers cannot immediately get the merchandise they buy. How to shorten the time from ordering the goods to receiving the goods for customers, can not only promote better development of e-commerce logistics and taobao.com, but also improve customer satisfaction of e-commerce logistics.

### 5 Conclusions

This paper mainly studies customer satisfaction evaluation system of C2C logistics. From this research, can know what customers care much, so it can not only improve the quality of logistics service, but also let the C2C businesses choose better logistics enterprises, to improve customer satisfaction, so as to promote the rapid development of e-commerce.

This work only evaluates logistics customer satisfaction of pattern C2C, but for other e-commerce modes, such as B2B, B2C, O2O, etc., are not studied here. In addition, due to limit academic accumulation and limited time, there are many deficiencies, please comment.

### Acknowledgments

Thanks go to our research team of Changzhou Institute of Technology. Also appreciate for other people's work, such as Sunke, etc.. In one word, the authors would like to thank all those that support this research and these authors whose articles are listed in the references.

## References

- [1] Ye Zuoliang, Cai Li, Ye Zhenhua, Dai Li 2011 Third party logistics service quality and its impact on customer satisfaction of C2C *Science Research Management* 32(8) 119-26 (in Chinese)
- [2] Su Qin, Liu Ye-yi, Cao Peng 2010 Empirical Study on C2C E-commerce Service Quality *Commercial Research* (03) 213-6 (in Chinese)
- [3] Ramakrishnan Ramanathan, Joseph George, Usha Ramanathan 2014 The Role of Logistics in E-commerce Transactions: An Exploratory Study of Customer Feedback and Risk *Supply Chain Strategies, Issues and Models* 221-33
- [4] Ieva Meidutė-Kavaliauskienė, Artūras Aranskis, Michail Litvinenko 2014 Consumer Satisfaction with the Quality of Logistics Services *Procedia Social and Behavioral Sciences* 110 330-40
- [5] Adrian Micu, Kamer Aivaz, Alexandru Capatina 2013 *Economic Computation and Economic Cybernetics Studies and Research* 47(2) 1-9
- [6] Novani Santi, Kyoichi Kijima 2013 Efficiency and Effectiveness of C2C Interactions and Mutual Learning for Value Co-Creation: Agent-Based Simulation Approach *International Journal of Business & Management* 8(9) 50-62
- [7] Chen Chien-Wen, Cheng Chiang-Yu 2013 How online and offline behavior processes affect each other: customer behavior in a cyber-enhanced bookstore *Quality & Quantity* 47(5) 2539-55
- [8] Tibert Verhagen, Willemijn van Dolen 2011 The influence of online store beliefs on consumer online impulse buying: A model and empirical application *Information & Management* 48(8) 320-27
- [9] Shashank Rao, Thomas J Goldsby, Stanley E Griffis, Deepak Iyengar 2011 Electronic logistics service quality (e-LSQ): Its impact on the customer's purchase satisfaction and retention *Journal of Business Logistics* 32(2) 167-79
- [10] Ramanathan Ramakrishnan 2011 An empirical analysis on the influence of risk on relationships between handling of product returns and customer loyalty in E-commerce *International Journal of Production Economics* 130(2) 255-61
- [11] Thirumalai Sriram and Sinha Kingshuk K 2011 Customization of the online purchase process in electronic retailing and customer satisfaction: An online field study *Journal of Operations Management* 29(5) 477-87
- [12] Norizan Kassim, Nor Asiah Abdullah 2010 The effect of perceived service quality dimensions on customer satisfaction, trust, and loyalty in e-commerce settings: A cross cultural analysis *Asia Pacific Journal of Marketing and Logistics* 22(3) 351-71
- [13] Rudolf Leuschner, Douglas M Lambert, A Michael Knemeyer 2012 Logistics Performance, Customer Satisfaction, and Share of Business: A Comparison of Primary and Secondary Suppliers *Journal of Business Logistics* 33(3) 210-26
- [14] Socrates J Moschuris, George F Velis 2012 Customer Perceptions on Service Satisfaction with Third Party Logistics (3PL) Service *International Journal of Applied Logistics* 3(4) 33-47
- [15] Zaryab Sheikh, Shafaq Rana 2012 *International Journal of Academic Research in Business and Social Sciences* 2(1) 546-58

## Authors



**Jing An, born on December 22, 1981, China**

**Current position, grades:** lecturer at Changzhou Institute of Technology, China.

**University studies:** Bachelor of Information Management and Information System from Jilin University in 2004, Master of Informatics in 2007 from Jilin University, China, PhD (Informatics) degree from Jilin University in 2010.

**Scientific interest:** Information Management, Electronic Commerce and Services Science.

**Publications:** 10.

**Experience:** Teacher and active researcher at Changzhou Institute of Technology since 2010.



**Bo Xu, born on January 25, 1982, China**

**Current position, grades:** lecturer at Changzhou Institute of Technology.

**University studies:** Master Degree from Nanjing Agricultural University in 2010.

**Scientific interest:** Numerical Analysis.

**Publications:** 7 and 2 patents.

**Experience:** Teacher and active researcher at Changzhou Institute of Technology since 2010.

# Research and development of comprehensive communication experiment teaching system

**Bin Wang<sup>1, 2\*</sup>, Dashe Li<sup>1, 2</sup>, Shue Liu<sup>3</sup>**

<sup>1</sup>*School of Computer Science and Technology, Shandong Institute of Business & Technology, Yantai, China*

<sup>2</sup>*Key Laboratory of Intelligent Information Processing in Universities of Shandong, Yantai, China*

<sup>3</sup>*School of Computer Science and Technology, Binzhou Medical University, Yantai 264003,*

*Received 1 March 2014, www.tsi.lv*

---

## Abstract

This paper has developed a set of management and experiment system for communication laboratories. Its development platform is Visual Basic 6.0, using access database, winsocket programming and multicast technology. Including several sets of software and using C/S architecture, the system can jointly work with data network configuration system so as to achieve the communication experiment of data network. It can conduct VLAN isolation and IP filter using right management switch, which can make multiple servers connect to the device simultaneously and completely control the numbers of computers entering the system at some point. By off-line configuration technology and shared online database technology, the system can make the device resources be assigned automatically to solve the basic problems of many people doing experiment at the same time.

*Keywords:* comprehensive communication experiment teaching system, access, winsocket, multicast

---

## 1 Introduction

Currently, there are many problems in the electronic information experimental teaching process of many universities. The first problem is shortage of device for more users. The communication devices the operators use are very expensive. A set of device costs a minimum of 300,000 Yuan and may cost more than 5,000,000 Yuan. In general, a university buys one set or several sets of devices. However, after nearly 10 years of university increase enrolment, there are much more college students. These have brought out the problem of device shortage with more users.

The second problem is system stability. Many students use one set of device by plugging and unplugging the lines to switch between different users, making the operation very confusing and the laboratory full of network cable. Besides, a lot of data device configures through the serial port, which does not support hot plug, and the transmission distance is limited.

The third problem is the heavy burden of teachers' management. When having classes, the teachers need to coordinate multiple students to use one set of device simultaneously. In addition, in order to increase the device utilization, they definitely increase the device use time. However, the teachers would not always be able to supervise in the laboratory for long [1-5].

To solve these problems, this paper has developed a comprehensive communication experiment system (CCS). The system uses special access device to manage

authorization, which is adoptive to the management of all communication devices.

## 2 The network structure of CCS

The system should support the shared or exclusive standing-in-line for a large number of students, including various modes of control, the port control, the IP filtering, the hybrid filter between IP and port, and a specific packet filtering. Being able to authorize for each port of every device, and allocate time automatically and support the queue management, time management and equipment management. The system has three teaching modes, namely, the study of configuration software, the controller's programming, and the adjustment of control system. So the network structure of the system is shown in Figure 1. Data network configuration system uses C/S architecture, managing the entire experimental devices at a unified platform to provide configuration and testing for the students. All computers are configured with two network cards. One is the management net with its IP locked, so it must be able to connect to the CCS2000 server and all gateways. Via RS232- Ethernet gateway, the computer can access the CONSOLE port of the data product by Ethernet to obtain the configuration capability of the lowest level. The system does not have the issues of short serial port cable, not supporting hot plug or unplug and others, so a computer can configure multiple devices at the same time.

---

\* *Corresponding author* e-mail: thor@vip.163.com

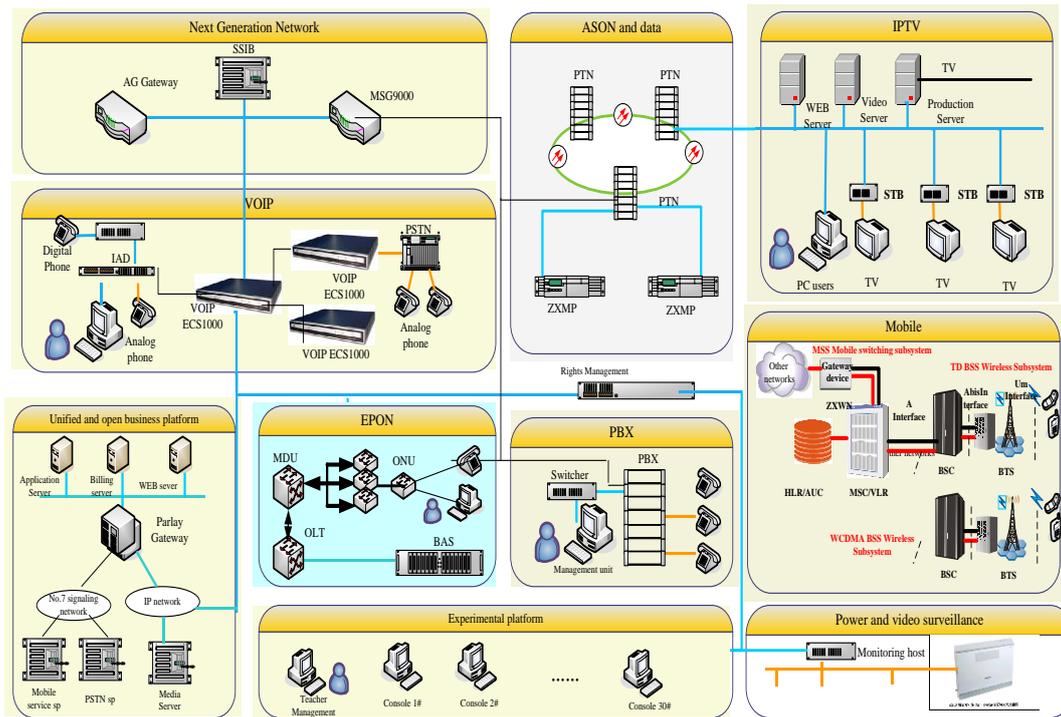


FIGURE 1 The overall structure of the system

One networking example is shown in Figure 2. All computers are configured as 129 servers with the same IP address. By CCS2000, they will not encounter IP address conflicts. Then, multiple students can make offline configuration. When the configuration is complete, they can successively connect to J10 to configure in turn. If there are multiple J10, the students can choose any one to queue [6, 7].

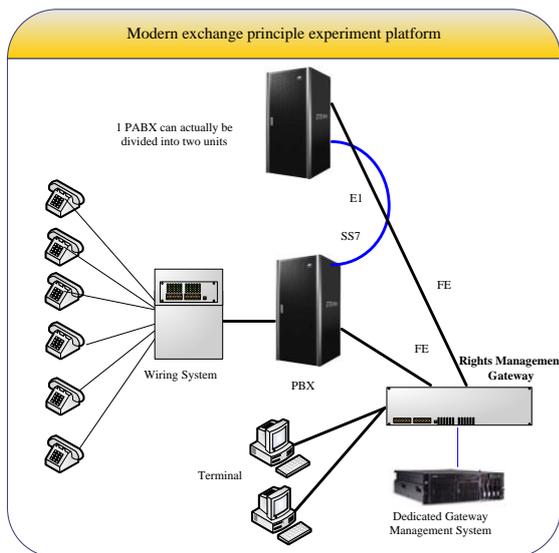


FIGURE 2 Networking example-J10

Another networking example is shown in Figure 3.

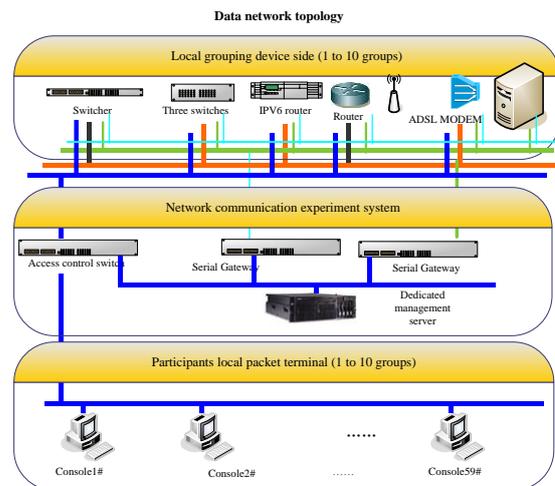


FIGURE 3 Networking example- data products

Topological graph can be divided into three layers, the upper is for experimental subjects, the middle is for the control layer, and the lower is for the student server.

In accordance with the test object, the experimental approach can be divided into two categories:

1) Those that can be classified as one category include the layer 2 switch in the middle, layer 3 switch, IPv6 router, the router, AP and ADSL MODEM. This class has a network equipment with a console port. As doing experiment, it needs to provide IP-to-serial function with serial gateway. In addition, the student port can complete the configuration of data network equipment only if it has a super terminal end.

The networking mode is to install all the student ports and the configuration of serial gateway in the same

network. The student port has access to the address and port of serial gateway by applying super terminal through TCP/IP (Winsock). CCS2000 software of serial gateway will assign a unique gateway to each student as access to the data network equipment's console port

2) Another category is the device for the server. To access such device, it requires student to have a client to the server.

The idea to access this kind of device in upper layer is: the student ports are connected with the right control switches in the middle layer, and each of them is in a different vlan for isolation. While doing the configuration experiment, student sends his/her request to a dedicated server in a queue (to access dedicated vlan communicating with the experimental object), and at this time the dedicated server will verify the student port's permission; if permission meets with the requirements, the dedicated server will allow the student port to communicate with the experimental object with the right control switch.

The dedicated configuration server controls the power gateway and the automatic hardwired gateway, which completes all kinds of networking construction and subsidiary occupation to experiments. The power gateway is an automation equipment controlled by stm32, which can complete the control over power switches of the experimental object. The automatic wired gateway is actually a switch with multi-ports (two or more layers), to which a number of data network equipment can be connected. Through the configuration of the automatic wired gateway's vlan, it can combine various topological graphs to the connected data network equipment and can accomplish various experimental networking constructions and enrich experimental content.

### 3 The software and hardware modules of the system

#### 3.1 SOFTWARE OF THE SERVER SIDE

Before the whole system can be used, the system database needs to be set according to the conditions of the networking mode, the laboratory devices and the experimental curriculum. Users can enter the editing interface to form the entire lab environment, including devices, curriculum, experiment and the students' computer information. The entire database includes the following sections: device library, gateway library, user library, curriculum and experiment library, queue library, administrator library and system data. In order to design, the system makes the following definitions:

Device group: a group of devices, which are generally not used to do experiments jointly.

Device: a network element with a separate communication port. For example, a program-controlled exchange MP, a net managed switch, a device with RS232 or RS485 port; or a device without communication, even a multi-meter.

Curriculum: the experiment courses the laboratory needs to open.

Experiment: the experiments of an experiment course.

Queue: a combination of some devices needed to complete some experiments. The queue is divided according to the device group and curriculum. A student selects one experiment and the experiment needs a queue which includes all the devices needed to complete the experiment.

Auxiliary queue: to connect the network cables of some devices. Currently, only the soft connection (VLAN connection), not the hard connection between devices can be provided. Network experiment, free of cable connection, can be achieved.

User computer: the computer used by teachers and students.

User: including user ID, user name and user password. The specified users can only use the computers belongs to them, but they can choose only one computer to log in. Users are divided to teachers and students. Teachers can manage the queue, but cannot select queue to do experiment. Students can choose experiment and queue, do experiments, but only can control whether to line up or not.

Teachers group: teachers dedicated to management. The group ID is 0 and can't be deleted.

Administrator: can get into the database to modify and increase administrator user as well as other teacher and student users.

When normal server software is in the running state, you can see the queue status, computer status, device status and gateway status.

The displayed items of the queue state can be filtered according to the device groups and courses. For example, if you select the program-controlled switch device group, only the queue of this device group is shown below.

If you select a queue in the list, the specific information of the queue is displayed on the right, including the name of the current student, the remaining time of the current student and the total remaining time of the entire queuing. If the queue is shared, the current student is defined as "all" and the total time of the queue is the remaining time of the student who has the longest remaining time.

The computer status includes the login name and the connection status. When not connected, the user name is shown as "not connected". If connected but not logged in, it shows "not registered". And after login, it will show the user name.

The connection status includes: no connection, connected but no registration, already logged in, waiting in line and being testing.

The device state is shown as "device free" or "device occupied".

Related gateway information includes "online", "connected" or "no connection". The gateway states include "right control" or "control error".

The change of the state may take up to 1 minute for the refresh utilizes cycle approach.

### 3.2 SOFTWARE OF THE USER END

First, enter a start interface to check the network settings for all users. The communication IP is required to be set correctly to decide whether the user is a teacher or a student. The operation content of a teacher is more than that of a student and a teacher can do experiments as a student. Many attributes of the user have been identified when the server database was created.

The queue status interface of the teachers end is shown in Figure 2. Except can't operate queue and enter the pages of teaching property, the students end is the same as that of the teachers. So the additional introduction is not done here.

Select a user, a teacher can do following operations:

Click < Force to Quit >, then the user is forced to quit and this status will be displayed within one minute in the classroom end.

Click < Force to Shit Forward >, the user is forced to move to the front of the front person. If the chosen user is after the third person, it will show an error. For a shared type queue, this operation has no meaning.

Click < Force to the Front >, the user is forced to move to the second position. If the chosen user is after the third person, it will show an error. For a shared type queue, this operation has no meaning. If you what the selected student can do experiment immediately, you can force the first user to exit.

By increasing time, you can increase a certain period of time by minute to the chosen student.

User actions include selecting queue and exit initiatively.

Select one course and select one experiment, then you will see the queue in correspondence of the course and meeting the experiment. Then click <Join the Queue>, if the devices of the queue have been occupied by other queue, you cannot line up and the server system will give you a feedback of error message. Otherwise, the information of the selected queue will be displayed within one minute. Currently, a student can select only one queue to line up at one time.

Select a queue and then click <Exit the Queue>, waiting for response of the server and then you can exit.

If you select a serial port to Ethernet gateway, then it will show the devices of the selected queue. If the device belongs to serial port, you can control the serial by console command, which is especially useful for data products. For the internet access device, if it supports Telnet, console can also be used.

Using the serial port gateway, you can configure the serial port of a product through network not connecting a dedicated serial port, which can get the ability of the bottom control device and provide possibility for remote experiment. You can do operation free from cable connection with the additional help of auxiliary queue.

By command, you can make the product back to the factory state.

The special operation for teachers can select optional courses, experiment and queue. The optional and not optional content can be selected to the list below. The entire program can also be saved to an external file or restored and sent to the server, so all the courses, experiments and queues optional for the students are within the control of the teachers.

### 3.3 HARDWARE OF THE SYSTEM

The system uses dedicated experimental servers, supporting for the unified clearance function of the configurations of "a specified device", "a few specified devices" and "all experiment devices" of the experimental devices. GW2112/16 supports power-on-clearance, having good ability of device identifying, able to identify the type, quantity, location, port status and other information of the second and third level switches and routers of ZET. It supports a user to connect up to five devices simultaneously by network, using simple and intuitive graphical management interfaces. Besides, it supports up to five users simultaneously to connect the configuration ports of a device over the network. Managing the devices using queuing, you can see the queue status of the device, including the information of the queuing computers and queuing time. Also, the teachers can manage the queue.

The system directly uses ZXR10-2826S and ZXR10-2852S, supporting the simultaneous management and control of 46 network elements as well as up to 10 stacks. It not only supports the exclusive authority, but also supports sharing authority. The maximum full control period is 4 seconds and the average control period is 2 seconds. The management interface can provide the number and device model information as well as the occupying information of the connected experiment devices.

## 4. Program design process

### 4.1 ADMINISTRATOR PROGRAM DESIGN PROCESS

The system can support the gateway transformation on the basis of the serial device, so as to make a remote access to the controller and also support the group management and control of the equipment's and users. Combine the resources according to the teaching program, and it can provide the exclusive resources and the sharing mode for allocating experiment resources by the way of line. Finally, the system can support the teachers' experiment development, with the experiment time, the mode of equipment network in need and the like. Through the comprehensive control management of access, consisting of the time and queue control, the authorized port switch, routing switch and gateway

control, it can realize the orderly conduct of the whole experiment, providing teachers with the convenient management and the organization, modification, preservation and effectiveness of teaching plan.

The system starts the server and server software. Before the entire system can be used, the database needs to be set according to the conditions of the network mode, experiment devices as well as the experiment courses. Establish device and gateway information needs to create a device group first and then create devices and gateways. The administrator logs in to the configuration mode, restores the database, creates a template and then device group, devices, all kinds of users, courses, experiment and experiment time according to the requirements of the experiments. And later the administrator needs to configure the queue, the device group of the queue, cable connection free operations and the supported experiments. Finally, the administrator tests the system. If the test is successful, the database can be saved, or it is needed to be reconfigured [8]. The design process is as Figure 4.

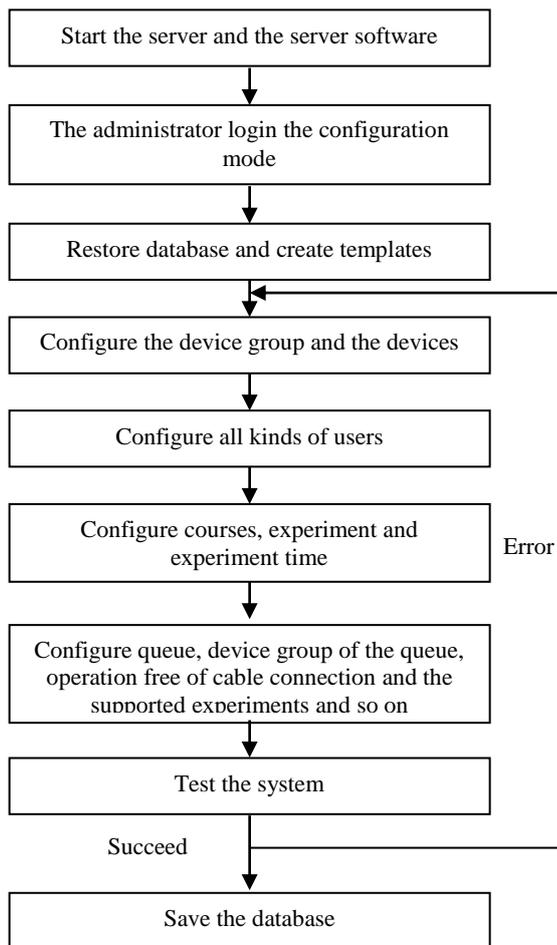


FIGURE 4 Administrator program design process

4.2 TEACHERS PROGRAM DESIGN PROCESS

Teachers can perform all the operations of students, so these operations are not included here. Device automatic identification, one key recovery, power control and other functions are for the switches, routers and others data products. The teacher computer can be any IP of the network segments the of gateway specified “server communication IP”. After starting the user end software, the teachers can login the system, set up, restore and modify the teaching programs in accordance with the requirements of the experiments and then check the status of all the devices, queues and users, manage the queues according to the situation of the experiments and do text communications with the students. After the experiment, the data products can be recovered by one key recovery [9].

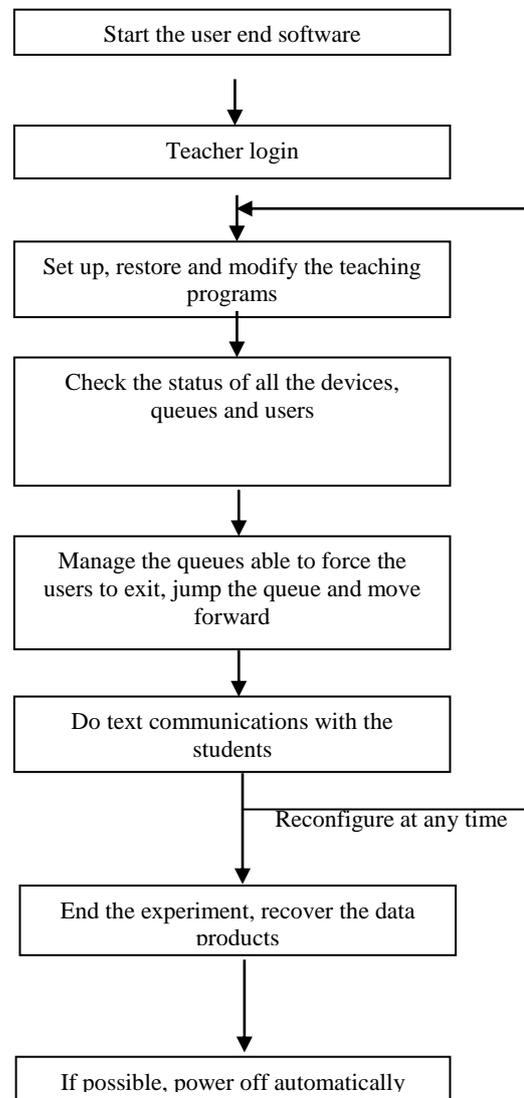


FIGURE 5 Teacher program design process

Support the setting of teaching plan, including setting up all the optional and non-optional courses, experiment and queue to be saved into the file. Therefore, the system

can support the organization, modification, preservation and effectiveness of teaching plan. The text communication can be conducted between teachers and students, so that it is convenient for teachers to help students. Teacher program design process is as Figure 5.

4.3 STUDENTS PROGRAM DESIGN PROCESS

Data network configuration system employs C/S architecture, which can manage all the experimental equipment and provide student with the configuration and testing in a unified platform. The Students program design process is as Figure 6.

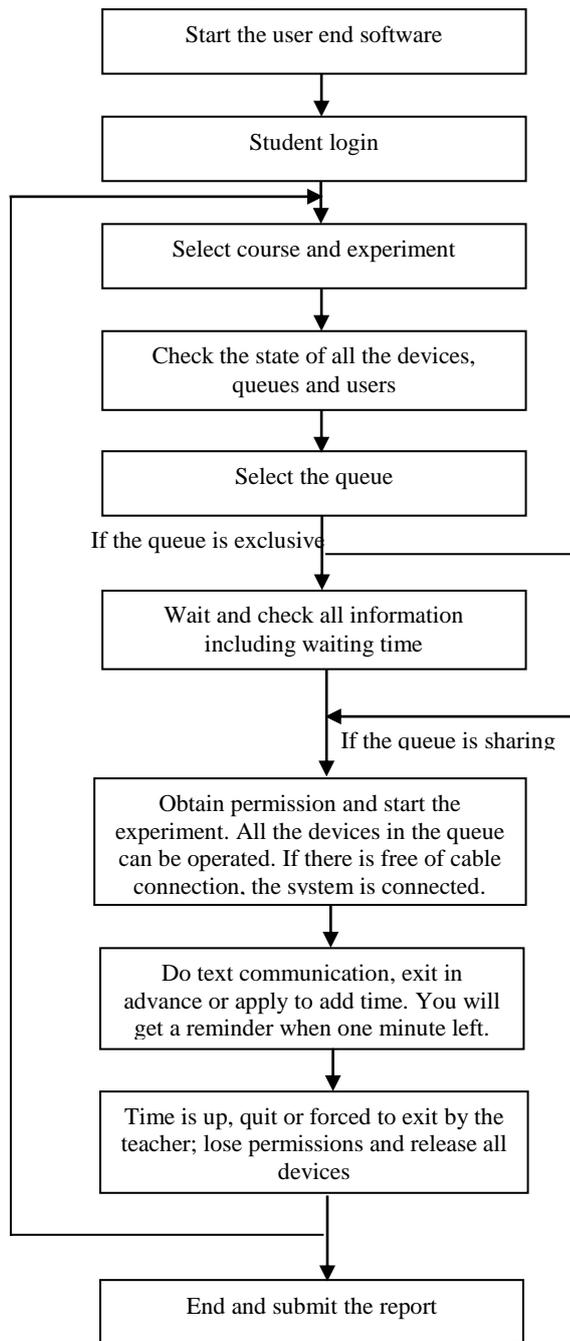


FIGURE 6 Student program design process

The system should support the queue management and operation. Students have the choice to join in the queue or drop out of it while the teachers can manage the queue, such as cutting-in-line, exit, and increase the time of individual students. The system should realize automatic connection among some devices, thus the isolation can be achieved between the device and the internship space. Furthermore, the system’s log-in is recorded to trace the equipment failure.

After CCS starts, the students firstly need to try to connect to their IP, the same with the first three bytes of the server. If it requires scanning the server, then send TCP connection to all the possible servers. Then the server will respond to provide basic information. In addition to responding to the asking, the server will do multicast to the users. After selecting a server connection, the user joins a multicast. Waiting for the multicast, if not receive any multicast after two seconds, you will receive an error. Select a test area and write the user name, then it will connect to the server. If the server makes respond and then you enter the system. Otherwise, a starting error message will be displayed [10].

After starting the user end software, the students login the system. They can select the relevant course and experiment and then check the state of all the devices, queues and users. After selecting the experiment, select the queue. If the queue is exclusive, then wait. While waiting, they can check all information including waiting time. After obtaining permission, they can start the experiment. They can do text communication, exit in advance in the experiment process. Submit the experiment report after the experiment.

5 System debugging

Data product debug configuration is generally carried out by means of the console port. The serial port line, available for the configuration and daily protection to the equipment, is a basic cable for the configuration of data product.

Data product is generally attached with a serial port configuration line by random. For ZXR10 product, one end is a DB9 serial interface (with a connection to computer’s serial port), and the other end is a RJ45 port (connected to the equipment’s console port). The schematic sketch of the serial port configuration line is shown in Figure 7.

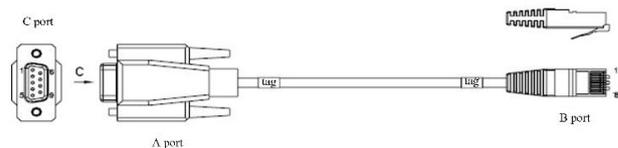


FIGURE 7 Student program design process

It uses the specialized data configuration line, and the connection device of the console port adopts the terminal mode of VT100. After the proper connection of PC and

data product, it can select to connect with TCP/IP or the serial ports such as COM1, COM2 and so on, depending on the connected serial port of the configuration wire. Meantime, it will set the port property of the selected serial port. It can be directly connected to the serial port on the server. While on the user terminal, it can use a serial connection.

Teacher's computer can be any IP of a segment in the "connecting with the server communication IP" specified by the gateway. He/she can run the super terminal by selecting 192.168.1.200 (the specified IP in the sample), 23, and accessing.

## 6 Conclusions

The most basic principle of the project is to do connection management to the terminals of the network through switches, routers, gateway devices, so as to manage the user authorization. Via Ethernet communication technology and server with C/S structure to manager all the user ends to queue, reserve and so on. These switches and routers can do packet filtering via VLAN technology and ACL technology.

Using a variety of strategies to visit access control, it analyses and filters the users' IP and port as well as the

first 80 bytes of the information in order to more accurately control the behaviour of the terminals. This provides opportunity for experiments of multiple campuses and even the entire WAN, which greatly simplifies complexity of the wiring and construction. In communication, it uses broadcast technology and multicast technology, making the whole system respond quickly and able to withstand the impact of large loads. Besides, it uses a distributed server technology, so the user on any network can obtain the information broadcasted by the server of the experimental area.

## Acknowledgments

This research was supported in part by National Natural Science Foundation (No.61070175), Shandong Province Natural Science Foundation (ZR2013FL017, ZR2013FL018), Shandong Province University Science and Technology (J12LJ03) of China, project development plan of science and technology of Yantai (2013ZH347, 2013ZH091). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

## References

- [1] Naitoh K 2202 *Artificial Life and Robotics* 6(1) 82-6
- [2] Fan Z, Yili W 2007 Explore of Bilingual Education of Computer Network *Journal of Chongqing institute of technology* 21(3) 159-161
- [3] Mantri A, Dutt S, Gupta P, Chitkara M 2012 *IEEE Trans Educ* 51(4) 432-8
- [4] Mitchell J, Canavan B, Smith J 2009 *IEEE Trans Educ* 53(4) 587-94
- [5] Jiao W-H 2010 Design and Implementation of the Experimental Teaching of Computer Network Protocol Based on TCP/IP Model. *Research and Exploration in Laboratory* 32(10) 363-7 (in Chinese)
- [6] Brennan R, Thompson K, Wilder R 1991 *IEEE NetWork Magazine* 4(2) 32-40
- [7] Courtois P J, Heymans F, Parnas D L 1985 *Communications of the ACM* 14(10) 190-9
- [8] Pechurin N K, Kondratova L P 1999 *Cybernetics and Systems Analysis* 35(5) 797-801
- [9] Gribova V 2010 *Journal of Computer and Systems Sciences International* 45(4) 613-22
- [10] Boutellier J 2006 Panoramas from Partially Blurred Video *Advances in machine vision, image processing and pattern analysis* 4153 300-7

## Authors



**Wang Bin, born in February, 1981, Yantai County, Shandong Province, P.R. China**

**Current position, grades:** lecturer of Department of School of Computer Science and Technology, Shandong Institute of Business & Technology, China.

**University studies:** M.Sc. from Shandong University.

**Scientific interest:** wireless communication, computer network.

**Publications:** more than 10 papers published in various journals.

**Experience:** teaching experience of 13 years, 8 scientific research projects.



**Li Dashe, born in February, 1978, Yantai County, Shandong Province, P.R. China**

**Current position, grades:** the Associate Professor of Department of School of Computer Science and Technology, Shandong Institute of Business & Technology, China.

**University studies:** M.Sc. from China University of Mining & Technology (Beijing) in China.

**Scientific interest:** wireless communication, computer network.

**Publications:** more than 20 papers published in various journals.

**Experience:** teaching experience of 11 years, 8 scientific research projects.



**Liu Shue, born in November, 1977, Yantai County, Shandong Province, P.R. China**

**Current position, grades:** the Lecturer of School of Computer Science and Technology, Binzhou Medical University, China.

**University studies:** M.Sc. from Yantai University in China.

**Scientific interest:** include wireless communication, computer network.

**Publications:** more than 10 papers published in various journals.

**Experience:** teaching experience of 12 years, 3 scientific research projects.

# Artificial neural network model of forecasting relative humidity in different humid and arid areas of China

Zhenfang He<sup>1, 2\*</sup>

<sup>1</sup>*Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, China, 730000*

<sup>2</sup>*University of Chinese Academy of Sciences, Beijing, China, 100049*

*Received 6 October 2013, www.tsi.lv*

## Abstract

The objective of the present study is to build different models forecasting the daily mean relative humidity (MRH) values in China with the help of the meteorological parameters. A back-propagation artificial neural network (BPANN) models was employed to identify the relationship between meteorological factors and the relative humidity in China. Weather data 1-day lag was the input layer variables, including (1) the highest atmospheric pressure, (2) the lowest atmospheric pressure, (3) the average atmospheric pressure, (4) the average temperature, (5) the highest temperature, (6) the lowest temperature, (7) precipitation, (8) the average wind speed, (9) the maximum wind speed (the average wind speed over 10 minutes), (10) the utmost wind speed, (11) hours of sunlight, (12) the relative humidity. Experimental results: in the validation period for 1-day lead, the comparison of the prediction performance efficiency of the BPANN models indicated that the BPANN models with trainbr algorithm was superior to the remaining two ones (trainlm and traingdx) in forecasting the relative humidity time series in term of correlation coefficient (R). During the training and testing periods for 1-day lead, the best performance for the given problem was arid area, followed by semi-arid area, semi-humid area, and humid area respectively. The possible cause for the results was that the impact of these factors on the relative humidity in arid area was the largest, followed by semi-arid area, semi-humid area, and humid area, respectively. From the prediction results of MRHextrema, humid area was the first; semi-arid area was the second; semi-humid area was the third; and arid area was the fourth. From the prediction results of MRHextrema, trainbr algorithm was the best in arid area, semi-humid area, and humid area; but trainlm was the best in semi-arid area. So trainbr algorithm was further employed to predict MRH for 2, 3 or 4-day lead at Urumqi City. From the training and testing effects, 1-day lead was the best, followed by 2, 3 or 4-day lead respectively. In the prediction results of MRHextrema, the best was 2-day lead; the second was 3-day lead; the third was 1-day lead; and the fourth was 4-day lead. The BPANN model results will assist researchers determining meteorological parameters to forecast MRH.

*Keywords:* Meteorological Parameters, Humid and Arid Areas, Artificial Neural Network Model, Relative Humidity, Training Algorithms

## 1 Introduction

Relative humidity as a major meteorological component of the hydrologic cycle plays an important role in climate change studies in climatic regions. The influence of relative humidity in controlled environments (e.g. industrial processes in agro-food processing, cold storage of foods such as vegetables, fruits and meat, or controls in greenhouses) is vital [1]. The black-box modelling method is one of data developing techniques in which the knowledge is abstracted in term of models. When data is not sufficient, empirical models are a good alternative method, and can provide useful results without a costly calibration time [2]. It is crucial to contain all significant variables that influence relative humidity. Too much irrelevant data would increase the network training and restraint the network from learning adequately. Unlike physical models, black-box models can be made adaptive, by transforming their parameters as a function of the actual performance that the models show [3]. Many researchers have applied ANN to predict relative humidity with the data related to weather conditions and

climate being collected for certain periods [4]. They used three-layer feed forward neural networks to predict indoor temperature and relative humidity, and temperature and relative humidity predictions got very accurate and satisfactory results [5]. They developed an artificial neural network model to predict the thermal behaviour of an open office in a modern building, and used external and internal climate data recorded over three months to build and validate models for predicting relative humidity, moreover, the results reveal that the model provides reasonably good predictions[6]. They aimed to determine simultaneously relative humidity in air, by employing an optical sensor based on a nafion-crystal violet film and ANN to perform multivariate calibration. In addition, he proposed a kinetic approach in order to improve the ANN performance [7]. They adopted artificial neural network approach to estimate the monthly mean relative humidity (MRH) values with the help of the topographical and meteorological parameters, and used latitude, longitude, altitude, precipitation and months of the year in the input layer of the ANN network, while the MRH in output layer of the network, at last, the

\* *Corresponding author* e-mail: [hzfwhy@126.com](mailto:hzfwhy@126.com)

result shows that the obtained values were in the acceptable error limits. To examine prediction suitability of artificial neural networks in terms of relative humidity in different areas, measurement data were taken from four different meteorological stations in China. In the present study, a methodology is presented to predict relative humidity simultaneously in China by using BPANN models. Thus, the methodology is of great practical importance. The applicability of the methodology is demonstrated by using three ANN training algorithms namely trainlm, trainbr, traingdx for forecasting relative humidity.

## 2 Study background

The four study stations selected for this study are known as Chongqing, Beijing, Lanzhou, Wulumuqi (Urumqi), which are located in China (Figure 1). On the basis of precipitation, China is divided into four areas, viz., humid, semi-humid, semi-arid and arid areas. Precipitation of humid area is greater than 800 mm, precipitation of semi-humid area range from 400 mm to 800mm, precipitation of semi-arid area vary from 200 mm to 400 mm, and precipitation of arid area is less than 200mm. The four study stations are located in the four areas, respectively; that is, study areas include Chongqing, Beijing, Lanzhou, Wulumuqi, which represent humid, semi-humid, semi-arid and arid area, respectively. Eastern China has a typical monsoon climate with high temperature and rainfall mainly occurring in the summer season. Northwestern China has a typical continent climate due to be far away from the ocean.

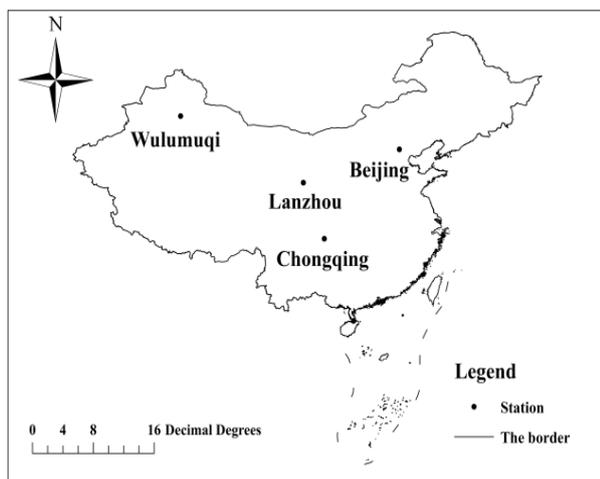


FIGURE 1. Location of the four meteorological observation stations

Table 1 shows the main meteorological parameters of the four stations. The average annual temperatures (AAT) of the four study stations are 14.8 °C, 11.8°C, 9.3°C and 6.6°C, respectively. The average annual relative humidities (AARH) of the four study stations are 81%, 52%, 53%, and 54%, respectively. The average annual rainfalls (AAR) of the four study stations are 1138mm, 576.9mm, 316.1mm, and 175.3mm, respectively. The

mean daily minimum and maximum relative humidities are 42% and 98% in Chongqing, respectively. The mean daily minimum and maximum relative humidities are 11% and 97% in Beijing, respectively. The mean daily minimum and maximum relative humidities are 17% and 91% in Lanzhou, respectively. The mean daily minimum and maximum relative humidities are 11% and 93% in Wulumuqi, respectively. There is a 2-year (2007-2008) record of the daily mean relative humidity (MRH) in the four study stations, and that is  $\{X(t), t=1, 2, \dots, n\}$ . For the ANN models, the data series were divided into a training set (January to December 2007), and a testing set (January to December 2008). The 2007 year time series are used for calibration/ training of the model, and the remaining year data are used for verification or testing purposes.

TABLE 1 The main meteorological parameters of the four study stations in China

Station	Chongqing	Beijing	Lanzhou	Wulumuqi
Latitude (oN)	28.8	39.9	36.0	43.7
Longitude (oE)	108.7	116.2	103.8	87.6
Altitude (m)	665	54	1518	918
AA T (°C)	14.8	11.8	9.3	6.6
AARH (%)	81	52	53	54
AAR (mm)	1138	576.9	316.1	175.3

## 3 Analysis Method based on ANN Model

An Artificial Neural Network (ANN) is a mathematical model or computational model that is inspired by the structure and functional aspects of biological neural networks. The true power and advantage of neural networks lies in their ability to represent both linear and non-linear relationships and to learn these relationships directly from the data being modelled. But, traditional linear models are simply inadequate when it comes to modelling data that contains non-linear characteristics. ANN is successfully applied to the areas of engineering, mathematics, medicine, meteorology, economy, neurology, psychology, electricity etc. [8, 9]. Although back propagation training has proved to be efficient in lots of applications, it has inherent limitations of gradient based techniques such as slow convergence and the local search nature. In this study, feed forward neural network architecture has been used and three improved ANN training algorithms, viz., gradient descent with momentum and adaptive learning rate back propagation (GDX) algorithm (traingdx), Levenberg-Marquardt (LM) algorithm (trainlm) and Bayesian regularization (BR) algorithm (trainbr) minimize a sum of squared error and to overcome the limitations in the standard BPANN. Three statistical criteria (or statistical indicators) were used in order to evaluate the effectiveness of back propagation artificial neural network (BPANN) models developed in this study. They are root mean square error (RMSE), mean error (ME), correlation coefficient (R), and percentage error of peak (EOP).

The global error function (E) can be calculated by using equation (1) as,

$$E = \frac{1}{2} \sum (O_i - P_i)^2, \tag{1}$$

where E is the global error function,  $O_i$  is the desired output and  $P_i$  is the output predicted by the network. The back propagation algorithm uses the gradient descent technique to adjust the weights in which the global error function, E, is minimized by modifying the weights using the following equation (2),

$$\Delta W_{ji} = -\eta \frac{\partial E}{\partial W_{ji}}, \tag{2}$$

where  $\Delta W_{ji}$  = weight increment from node  $i$  to node  $j$ ; and  $\eta$  = learning rate, by which the size of the step taken along the error surface is determined. The weights between the hidden layer and the output layer are adjusted first, followed by the weights between the hidden layer and the input layer.

The transfer function denoted by  $f(x)$ , defines the output of a neuron in terms of the induced local field  $x$ . linear function is represented by equation (3),

$$f(x) = x. \tag{3}$$

The most commonly used activation function within the neurons is the logistic sigmoid function, which takes the form shown in equation (4),

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{4}$$

The bipolarity S function, which has the advantage of a positive or negative output, can also be used, represented by equation (5). And, four statistical indicators were used to evaluate the effectiveness of the ANN models developed in this study. They are the correlation coefficient (R), mean error (ME), root mean square error (RMSE) and peak error percentage (EOP (%)), given by the following equations (6), (7) and (8).

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \tag{5}$$

$$RMSE = \sqrt{\frac{\sum (O_i - P_i)^2}{N}}, \tag{6}$$

$$ME = \frac{1}{N} \sum (O_i - P_i), \tag{7}$$

$$R = \frac{\frac{1}{N} \sum (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\frac{1}{N} \sum (P_i - \bar{P})^2} \sqrt{\frac{1}{N} \sum (O_i - \bar{O})^2}}, \tag{8}$$

where  $O_i$  = observed value for  $i^{\text{th}}$  data,  $P_i$  = predicted value for  $i^{\text{th}}$  data,  $\bar{O}$  = mean of observed value,  $\bar{P}$  = mean of predicted value, and  $n$  = number of observations. The best fit between observed and predicted values under ideal conditions would yield RMSE = 0, ME = 0, R = 1.

The percentage error of peak MRH, EOP (%) is defined as follows:

$$EO_p(\%) = \frac{P_p - O_p}{O_p} \times 100\%, \tag{9}$$

where  $P_p$  denotes the peak data of the predicted MRH,  $O_p$  is the peak data of the observed MRH and  $EO_p$  is the relative error of the maximum difference in the highest peak MRH.

Levenberg–Marquardt method is a modification of the Newton algorithm for finding an optimal solution to a minimization question. It is used to approach second order training speed and accuracy without having to compute the Hessian matrix. It uses an approximate to the Hessian matrix in the following Newton-like weight update.

$$W_{i+1} = W_i - [J^T J + uI]^{-1} J^T e, \tag{10}$$

where,  $W$  is weights of the neural network.  $J$  is Jacobian matrix of the performance criteria to be minimized.  $u$  is a scalar that controls the learning process, and  $e$  is residual error vector. When the scalar  $u$  is zero, this is just Newton’s method to use the approximate Hessian matrix. When  $u$  is large, the equation becomes gradient descent with small step size. Newton’s method is faster and more accurate near an error minimum, so the objective is to shift towards Newton’s method as quickly as possible.

#### 4 Relationships between MRH and meteorological factors

It is necessary to stress here that the problem of finding the most appropriate structure for a statistical daily MRH prediction model is perhaps one of the major problem for the modellers. Moreover, the link between meteorological factors and daily MRH is non-linear; the selected variables depend on the particular target of the prediction model. Finally, it must be stressed that observed data are affected by various kinds of noise. Generally, some degree of a priori knowledge is used to specify the initial set of candidate inputs [10, 11]. Here, the autoregressive analysis results showed that the autoregressive order was 1, with correlation coefficient R under a 95% level of significance; and this meant that modelling the time

series by using regressive models it should consider 1 past sample at least. The samples of the correlation functions computed, considering daily time series were reported in Figure 2 under a 95% level of significance.

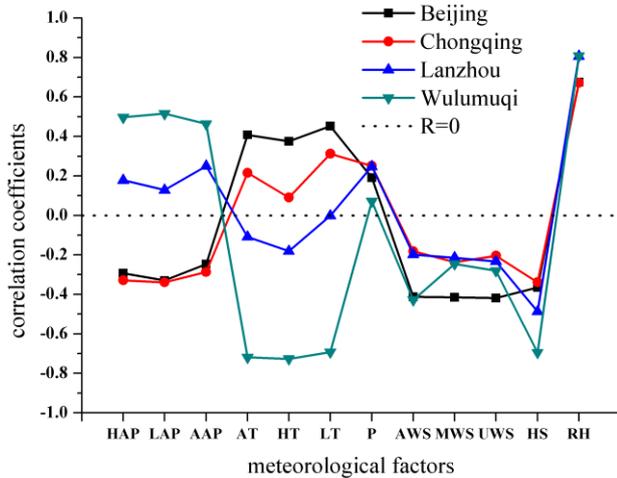


FIGURE 2 The relationship between MRH and meteorological factors

From the correlation coefficient (R) values between MRH and meteorological factors, such as AT, HS, RH 1-day lag, Wulumuqi was the best, followed by Lanzhou, Beijing, respectively, and Chongqing was the worst; that is, from the impact of these factors on the humidity, arid area was the largest, followed by semi-arid, semi-humid area, and humid area was the least. Therefore, 12 antecedent values of the meteorological factors were selected as input for modelling daily MRH. The 12 meteorological factors 1-day lag include (1) the highest atmospheric pressure (HAP), (2) the lowest atmospheric pressure (LAP), (3) the average atmospheric pressure (AAP), (4) the average temperature (AT), (5) the highest temperature (HT), (6) the lowest temperature (LT), (7) precipitation (P), (8) the average wind speed (AWS), (9) the maximum wind speed (the average wind speed over 10 minutes) (MWS), (10) the utmost wind speed (UWS), (11) hours of sunlight (HS), (12) the relative humidity (RH). In other words, the input layer consisted of 12 input nodes/variables and included a 1-day time-lag (X(t)), considering X(t) was the value of a given variable at the present time step for daily MRH. The output of the network was a prediction of daily MRH at time step t+1.

**5 Optimization of the hidden layer nodes**

In order to apply the ANN models, several network structures were tested to find the most appropriate topology. Numerous studies have shown theoretically that three-layered BP networks can precisely describe any nonlinear mapping relation [12, 13]. Therefore, three-layered BP networks were used in this paper. Sigmoid (tansig) and linear functions were used as activation functions in the neurons of the hidden layer and output neuron, respectively. The training was done for a maximum of 100 iterations. To avoid the over fitting problem, which generally appears with the application of

ANN, trains were used. The selection of the network was performed considering a minimum value of RMSE for the train data set. 12 nodes and one node were defined as the input and output layers, respectively, according to above paragraph. The number of hidden layer nodes was calculated by the trial-and-error method. 1 node was initially chosen; 6 nodes were finally selected after debugging. Figure 3 shows RMSE index evolution as a function of the number of hidden nodes for the different variances assessed in Wulumuqi station. In this study, multiple lead-time predictions of 1, 2, 3, 4 time steps were also performed for each input structure. In the case of the ANN, trainlm, traingdx and trainbr were used as learning rules. In all calculations, the best performing ANNs were feed forward back propagation with trainbr algorithm.

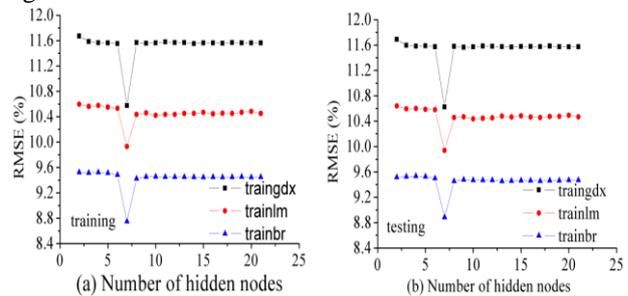


FIGURE 3 Experimental results for optimization of the hidden layer neurons (nodes) for BPANN models

**6 Prediction of the daily mean relative humidity (MRH) at 1day lead time**

The values of statistical indicators for the three training algorithms for the four stations were shown in Table 2 during training and testing periods. TABLE 2 showed the performance of all the three training algorithms was good during both training and testing periods; all the three training algorithms were able to forecast relative humidity 1 day ahead with a reasonable accuracy in all the four stations. For the trainbr training algorithm during testing period, the correlation coefficient (R) values ranged from 0.7634 to 0.9022, mean error (ME) from -0.3119% to 0.0872%, and root mean square error (RMSE) values from 7.8161% to 34.5310%. For the traingdx training algorithm during testing period, the correlation coefficient (R) values ranged from 0.6859 to 0.8569, mean error (ME) from -0.5927% to 0.3189%, and root mean square error (RMSE) values from 9.2149% to 36.7818%, whereas these figures for the trainlm algorithm were 0.7356 to 0.8769, -0.9683% to 0.3189%, 8.7155% to 31.0500%, respectively.

Table 2 summarized the results of the testing for every network configuration. The best overall performance for the given problem was achieved by the back propagation artificial neural network trained with the trainbr algorithm and the second best by the BPANN trained with the trainlm algorithm. The rest of the network performed relatively well but tended to overestimate the observed dataset. Also all the networks

performed very well for 1 day ahead predictions. Since the given problem was aiming at predicting the daily mean relative humidity (MRH), overestimating models were not of particular interest. Thus, the most promising technique seemed to be one using the feed forward neural

network trained with the trainbr algorithm. The physical meaning of this result was that the structure of this model allowed its weights to adjust to values that depict the trends of the natural system we were simulating.

TABLE 2 Comparison of trainbr, trainlm and traingdx algorithms

Algorithm	Station	area	RMSE (%)		R		ME (%)		EOP (%)	
			training	testing	training	testing	training	testing	training	testing
trainbr	Chongqing	humid	31.0583	34.5310	0.7639	0.7634	0.2486	-0.3119	-8.4280	-2.5180
	Beijing	semi-humid	12.5529	11.7988	0.7989	0.8167	-0.2219	0.0872	-11.747	-7.4643
	Lanzhou	semi-arid	7.7302	7.8161	0.8814	0.8629	0.2260	-0.1754	-3.3024	-7.3429
	Wulumuqi	arid	8.7508	8.8835	0.8844	0.9022	0.6602	-0.2056	3.4398	-8.7776
trainlm	Chongqing	humid	31.0500	38.0384	0.7638	0.7356	0.3628	-0.3609	-2.9994	-3.2772
	Beijing	semi-humid	13.6846	12.4098	0.7546	0.7947	0.1486	-0.1470	-4.5094	-9.4740
	Lanzhou	semi-arid	8.7155	8.7164	0.8460	0.8254	-0.1174	0.0996	-7.2760	-5.5173
	Wulumuqi	arid	9.9305	9.9392	0.8489	0.8769	0.9733	-0.9683	-10.701	-10.538
Traingdx	Chongqing	humid	36.7818	43.9031	0.7117	0.6859	0.3931	-0.4416	-9.8046	-6.0918
	Beijing	semi-humid	13.6846	13.1006	0.7024	0.7680	-0.1283	0.3189	-20.776	-19.359
	Lanzhou	semi-arid	9.2149	9.7276	0.8263	0.7774	0.0426	-0.1046	-14.699	-15.136
	Wulumuqi	arid	10.5758	10.6243	0.8265	0.8569	1.0433	-0.5927	-11.996	-13.381

The model calibration and validation results for the model design were analysed. In the training stage, at Wulumuqi station, the RMSE values for trainbr, trainlm and traingdx were 8.7508%, 9.9305% and 10.5758%, respectively, the ME values for trainbr, trainlm and traingdx were 0.6602%, 0.9733% and 1.0433%, and the R values for trainbr, trainlm and traingdx were 0.8844, 0.8489 and 0.8265. In the testing stage, RMSE were 8.8835%, 9.9392%, 10.6243%, ME are -0.2056%, -0.9683%, and -0.5927%, and R were 0.9022, 0.8769 and 0.8569. The RMSE values of the trainbr were smaller than trainlm and traingdx in both the training and testing stages, which implied that the calibration and validation capability of the ANN model with trainbr was better than that one with trainlm or traingdx for the given data. However, in testing phase, the magnitude of the ME values at Wulumuqi station was higher than that at Lanzhou station, implying a higher bias of the prediction results at Wulumuqi station. RMSE values also showed that the prediction results for Lanzhou station were better than those for Wulumuqi station. The R values were not much different for the two stations. For Lanzhou station, the performance of the three training algorithms was similar. However, for Wulumuqi station, the performance of the model with trainbr was better than that with trainlm and traingdx, especially for testing phase. However, the correlation coefficient (R) between input variables and daily MRH for Lanzhou station was lower than that for Wulumuqi station. This implied that the time-series data of Lanzhou station could have a higher nonlinearity than that of Lanzhou station and included more noisy data that hinder the model training process. The result of model performances for four stations indicated that the trainbr was more likely to catch the nonlinear relationship between meteorological factors and the relative humidity and to filter out the noise than trainlm or traingdx. FIGURE 4 showed the correlation coefficient (R) values between observed and predicted MRH 1 day ahead by three training algorithms with the observed daily MRH at

four stations. From the training effects of all the three training algorithms, Wulumuqi was the best, followed by Lanzhou, Beijing, respectively, and Chongqing was the worst; similarly, from the prediction results, Wulumuqi was the best, followed by Lanzhou, Beijing, respectively, and Chongqing was the worst.

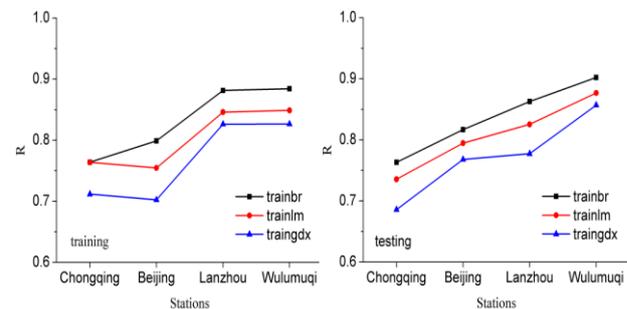


FIGURE 4 Experimental results for the correlation coefficient (R) values between observed and predicted MRH

MRH extrema are affected by random factors, which makes them difficult to predict. The EOP values in TABLE 2 reflected the models' performance in simulating the extremum at the four stations. During the training period, trainbr algorithm was the best at Wulumuqi and Lanzhou station, but trainlm algorithm was the best at Beijing and Chongqing station. During the testing period, trainbr algorithm was the best at Wulumuqi, Beijing, Chongqing station, but trainlm was the best at Lanzhou. From the prediction results of MRH extrema, Chongqing was the best, followed by Lanzhou, Beijing, respectively, and Wulumuqi was the worst. FIGURE 5 showed the comparison of the predicted daily mean relative humidity (MRH) 1 day ahead by three training algorithms with the observed daily MRH at four stations. These figures indicated that there was a very good matching between observed and simulated daily MRH at all the stations. Based on the statistical indicators and the graphical comparison, it could be deduced that all the three algorithms produced more or less same results.

However, the performance of the trainbr algorithm could be considered superior based on the performance criteria used in this study. On the other hand, the traingdx algorithm could effectively be used for large networks with little less accuracy than the trainlm algorithm and

the trainbr algorithm respectively. As a matter of fact, however, any of these three algorithms could be used for the daily mean relative humidity (MRH) prediction in the stations.

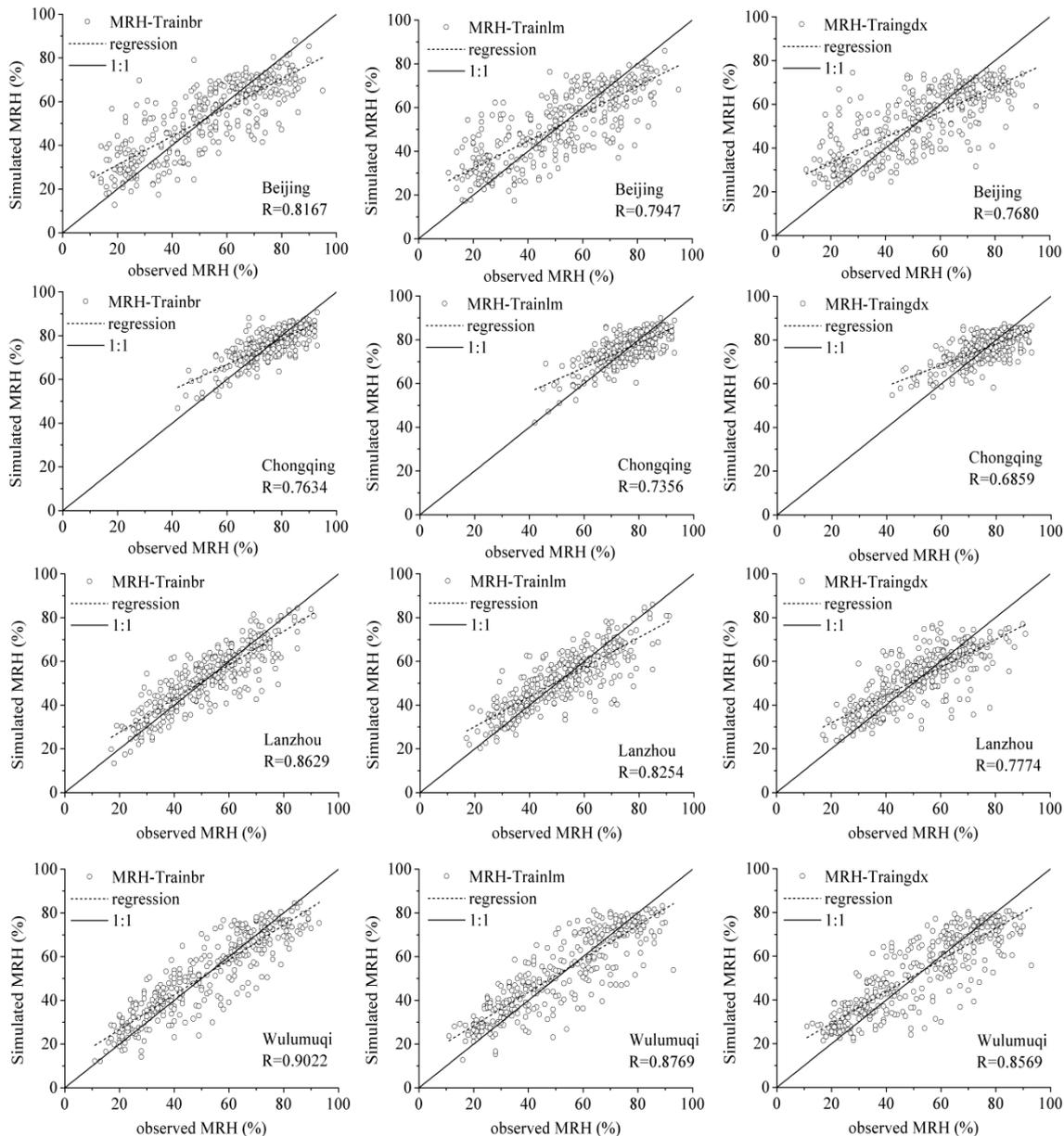


FIGURE 5 Comparison of observed and predicted MRH 1 day ahead at the four stations

It is very difficult to know which training algorithm will be the fastest or accurate for a given problem. Traingdx has adaptive learning rate. It is a faster training than traingd (the gradient descent algorithm), but can only be used in batch mode training. Trainbr (Bayesian regularization) is modification of the Levenberg–Marquardt (LM) training algorithm to produce networks which generalize well. It reduces the difficulty of determining the optimum network architecture. Trainlm (Levenberg–Marquardt) is normally used for training

purpose if enough memory is available. However, the LM algorithm appeared to be the fastest training algorithm, because the LM method must solve a linear system of equations in order to obtain the search direction, the computation becomes expensive when the number of input elements and the volume of the training data increase. Therefore, when the volume of the data is large, the standard gradient descent algorithm is used for training.

**7 Prediction of the daily mean relative humidity (MRH) at higher lead times**

As the three ANN training algorithms were determined in this study, the trainbr algorithm performed slightly better than the other two algorithms, and it was further employed to predict the daily mean relative humidity (MRH) at 2, 3 and 4-day lead at the four stations. It was worth referring that the ANN inputs used for this analysis were the same as that used for predicting daily MRH 1 day ahead. The performance of the model for Wulumuqi station in terms of correlation coefficient (R), mean error (ME), and root mean square error (RMSE) statistics along the prediction time horizon during the testing period was shown in TABLE 3. The values of the performance criteria have been acquired for the station. Taking Wulumuqi station as an example, it was obvious from this table that the R value varied from 0.9022 for 1-day lead time prediction to 0.8410 for 4-day lead time prediction, the value of ME varied from -0.2056% for 1-day lead time to -0.6214% for 4-day lead time, and the value of RMSE varied from 8.8835% for 1-day lead time to 11.2188% for 4-day lead time. ME values for 1, 4-day

lead time indicated an overestimation for the model, and an overestimation for 2, 3-day lead time. However, the magnitude of the ME values for 4-day lead-time was higher than that for 1-day lead time, implying a higher bias of the prediction results for 4-day lead-time. At Wulumuqi station, from the training effects of trainbr algorithm, 1-day lead time was the best, followed by 2, 4-day lead time respectively, and 3-day lead time was the worst; Similarly, from the prediction results, 1-day lead time was the best, followed by 2,4-day lead time respectively, and 3-day lead time was the worst. Thus, based on the performance criteria, it could be deduced that the predicted the daily mean relative humidity (MRH) for the higher lead times (2 to 4 days) were rational in this study, but the performance of the BPANN model normally reduced with an increase in the lead time. In the simulated results of MRH extrema, during the training period, 4-day lead time was the best, followed by 1,2-day lead time respectively, and 3-day lead time was the worst; during the testing period, 2-day lead time was the best, followed by 3,1-day lead time respectively, and 4-day lead time was the worst.

TABLE 3 Goodness-of-fit statistics using trainbr for different lead time forecasts at Wulumuqi station

lead time (day)	RMSE (%)		R		ME (%)		EOP (%)	
	training	testing	training	testing	training	testing	training	testing
1	8.7508	8.8835	0.8844	0.9022	0.6602	-0.2056	3.4398	-8.7776
2	13.5752	13.7296	0.8193	0.8496	1.2087	-0.5599	-4.3926	1.1988
3	58.8667	55.3289	0.7831	0.8204	0.8514	-0.9652	-6.6652	-2.5899
4	11.0688	11.2188	0.8038	0.8410	0.4782	-0.6214	3.0198	7.9712

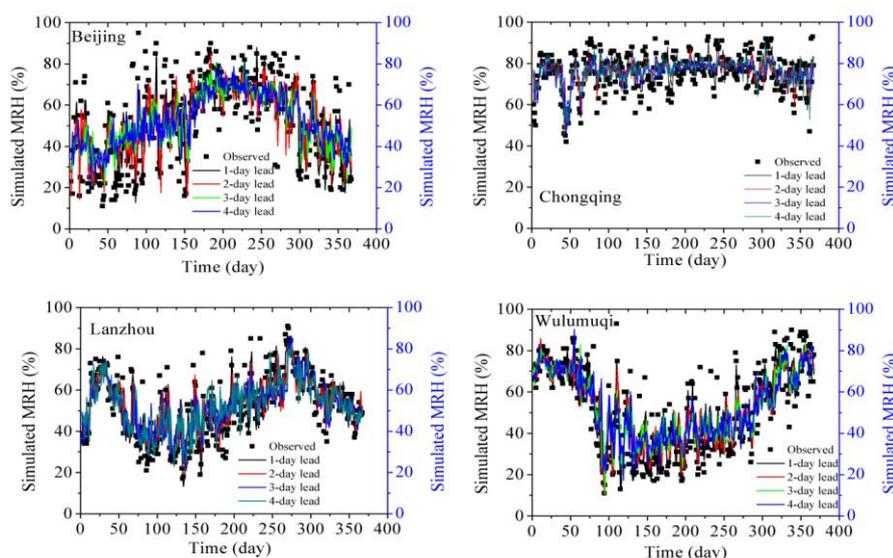


FIGURE 6 Comparison of observed and predicted MRH using trainbr from 1 to 4 days ahead at at the four stations

At the same time, FIGURE 6 also showed a comparison of observed and calculated for the daily mean relative humidity (MRH) predictions from 1 to 4 days ahead at four stations by the best performing network. The model results showed relative good prediction of the trend of the daily mean relative humidity (MRH); however the model prediction accuracy decreased slightly

with increasing horizon of prediction. In the mass, it could be deduced that the designed BPANN models could predict the daily mean relative humidity (MRH) over the area reasonably well for 1-, 2-, 3- and 4-day lead times. Thus, the BPANN technique was more appropriate where the knowledge of meteorological parameters was restricted.

## 8 Conclusion

This paper presented a multi-objective strategy for the optimal design of BPANNs when disposing nonlinear modelling relative humidity time series. The most suitable configuration for this task was proven to be a (12; 6; 1) feed forward network using trainbr (Bayesian regularization) as it showed the most accurate predictions of relative humidity 1-day in advance. From the prediction outcome 1day lead time, arid area was the best, followed by semi-arid area, semi-humid area, respectively, and humid area was the worst. The possible reason was that the impact of these factors on the relative humidity in arid area was the largest, followed by that in semi-arid, semi-humid area, and that in humid area is the least. In predicting the extremum, trainbr algorithm was the best at Wulumuqi station, Beijing station and

Chongqing station, but trainlm was the best at Lanzhou station. From the prediction results of MRH extrema, Chongqing was the best, followed by, Lanzhou, Beijing, respectively, and Wulumuqi was the worst. The BPANN model trained with trainbr algorithm was further used to predict relative humidity at 4 stations with higher lead times (2- 3- and 4-daily lead times). From the prediction results for higher lead times at Wulumuqi station, 1-day lead time was the best, followed by 2, 4-day lead time respectively, and 3-day lead time was the worst. It was found that the relative humidity prediction was reasonably good for all the three lead times, but the accuracy of prediction decreased with increasing lead times. In predicting the extremum at Wulumuqi station, 2-day lead time was the best, followed by 3, 1-day lead time respectively, and 4-day lead time was the worst.

## References

- [1] Martinez-Romero A, Ortega J F, de-Juan J A, Tarjuelo J M, Moreno M A 2010 Development of relative humidity models by using optimized neural network structures *Spanish Journal of Agricultural Research* **8**(S2) 162-71
- [2] Krishna B, Satyajji Rao Y R, Vijaya T 2008 Modelling groundwater levels in an urban coastal aquifer using artificial neural networks *Hydrological Processes* **22**(8) 1180-8
- [3] Ruano A E, Crispim E M, Conceicao E Z E, Lucio M 2006 Prediction of building's temperature using neural networks models *Energy and Buildings* **38**(6) 682-94
- [4] Lu T, Viljanen M 2009 Prediction of indoor temperature and relative humidity using neural network models: model comparison. *Neural Computing & Applications* **18**(4) 345-57
- [5] Mustafaraj G, Lowry G, Chen J 2011 Prediction of room temperature and relative humidity by autoregressive linear and nonlinear neural network models for an open office *Energy and Buildings* **43**(6) 1452-60
- [6] Raimundo Jr I M, Narayanaswamy R 2001 Simultaneous determination of relative humidity and ammonia in air employing an optical fibre sensor and artificial neural network *Sensors and Actuators B-Chemical* **74**(1) 60-8
- [7] Yasar A, Simsek E, Bilgili M, Yucel A, Ilhan I 2012 Estimation of relative humidity based on artificial neural network approach in the Aegean Region of Turkey *Meteorology and Atmospheric Physics* **115**(1-2) 81-7
- [8] Liu S Q, Lu X F, Liu R M, Zeng W H, He M C 2012 The application of the Bp neural network in the estimation of non-point source pollution *Procedia Environmental Sciences* **13**(2012) 263-73
- [9] Fleming Z L, Monks P S, Manning A J 2012 Review: Untangling the influence of air-mass history in interpreting observed atmospheric composition *Atmospheric Research* **104**(2012) 1-39
- [10] Thirumalaiah K, Deo M C 2000 Hydrological Forecasting Using Neural Networks *Journal of Hydrologic Engineering* **5**(2) 180-9
- [11] Wang G 2013 Research on supply chain demand prediction based on BP neural network algorithm *INMATEH-Agricultural Engineering* **40**(2) 27-34
- [12] Jiang S, Ren Z, Xue K, Li C 2008 Application of BPANN for prediction of backward ball spinning of thin-walled tubular part with longitudinal inner ribs *Journal of Materials Processing Technology* **196**(1) 190-6
- [13] Wang H 2013 Research on supply chain performance evaluation of fresh agricultural products *INMATEH-Agricultural Engineering* **40**(2) 35-42

Author	
	<p><b>Zhenfang He, born in January, 1983, Lanzhou County, Gansu Province, P.R. China</b></p> <p><b>Current position, grades:</b> the PHD of University of Chinese Academy of Sciences, China.</p> <p><b>University studies:</b> received her B.Sc. in GIS from Shandong Agricultural University in China. She received his M.Sc. from Northwest University in China.</p> <p><b>Scientific interest:</b> Her research interest fields include computer application.</p> <p><b>Publications:</b> more than 20 papers published in various journals.</p> <p><b>Experience:</b> She has completed five scientific research projects.</p>

# Dynamic pricing model of monopolistic manufacture based on the after-sale service

Yulan Zhou\*

*School of Economy & Management, Southwest Petroleum University, No.8, Xin Du Road, Chengdu, Sichuan, China*

*Received 6 October 2013, www.tsi.lv*

---

## Abstract

Dynamic pricing is concerned by business and academia as a pricing method, and has also made extensive research in this field. But the dynamic pricing theory with multi period is not mature considering monopolistic environment and after-sale service of manufacture. Because consumer not only pays attention to the product itself, but more emphasis on after-sale service of product with the changes of consumption concept and increasingly fierce market competition. Therefore good after-sale service is an important reason for consumers to purchase repetitively, which has become the key to the success of manufacture. This paper puts forward to the demand function with learning character, and constructs the multi-period dynamic pricing model on account of monopolistic manufacture and after-sale service level. Then it has important theoretical and practical significance when the conclusions are applied to the monopolistic manufacture. The research findings show the product price of manufacture fluctuates with oscillation both in the short and long term. But it is gradually reduced to a constant value for the magnitude or extent of the price oscillation with certain rate convergence in the long term. Finally, the price may tend to consumer's reservation price or unit operating cost of manufacture.

*Keywords:* Dynamic Pricing Model, After-sale Service, Monopolistic Manufacture

---

## 1 Introduction

The study of pricing theory is nothing more than from macro and micro perspective according to the existing literatures. The first kind of product pricing is about economic analysis based on macro perspective, which summarized the price formation mechanism by using the basic economics principles, such as the marginal utility theory of Adam Smith (1776), the labour value theory of Marx (1845), the equilibrium price theory of Marshall (1920) and the incomplete competition theory putted forward by Chamberlain (1938) and Robinson (1933), etc.. It has derived many products pricing methods based on the above pricing theory, such as elastic analysis method, correlation product pricing method and complete sets of products pricing method, etc. [1]. The second is the specific product pricing method because the cost profit for traditional pricing mode was broke out by many economists with the development of market and economy, and a series of pricing models were developed. These pricing models are developed to the practical application and the direction of diversification, which in addition to considering the internal variety and other factors, but also considering the external market volatility and the competitive environment. Dynamic pricing is that the manufacture adjusts product prices for different customer groups in real time in order to obtain the maximum benefit. As an important pricing technology of revenue management, dynamic pricing has been integrated into the various optimization software of management, which provides an important support for manufactures to develop

a reasonable price decision, and has been widely used in the tourism industry, such as aviation, automobile, railway, etc. [2-3].

There is great development about dynamic pricing model of the monopolistic manufacture. Smith and Achabal (1998) construct the dynamic optimal control model according to the retail commodity with clearance at the end of season, which considers the change of sales volume is influenced by lower prices and the seasonal change [4]. Zhao and Zheng (2000) study the demand is non homogeneous Poisson process based on Gallego and van Ryzin (1994). Their results show the optimal price increases with the decrease of inventory quantity in a certain period, and it is monotone under certain conditions [5]. Constantinos Maglaras, Joern Meissner maximize firms' total expected revenues over a finite horizon and put forward to dynamic pricing strategies for multiproduct revenue management problems [6]. Xuanming Su (2010) studies a monopolist firm selling a fixed capacity, and his findings show that the firm's expected profits will increase when the presence of speculators increase, and the speculative behaviour may generate incentives which lead to lower capacity investments [7]. Josef Broder and Paat Rusmevi chientong (2012) build a stylized dynamic pricing model, in which the purchasing decisions for a sequence customers make a monopolistic prices [8]. Bhalla and Manaswini (2012) solve a monopolist's optimal price strategy, and his results show the prices and per period profits will increase over the period of time [9]. But many researches seldom consider the dynamic pricing from the manufacturer's own point of view. The

---

\* *Corresponding author* e-mail: wqed1976@126.com

consumer's behaviour is not only related to consumption preferences, consumption habits, income and other factors of consumers, but also more closely linked with the product quality, reputation and after-sale service etc. Therefore, the after-sale service is the key to win for manufacture, which has already become the new focus of market competition [10].

In summary, the dynamic pricing theory with multi period is not mature considering monopolistic environment and after-sale service of manufacture. The literature is still relatively small, which is need to be researched further in-depth and systematically. This paper discusses the dynamic pricing strategy based on after-sale service under the monopoly environment. Firstly, it establishes the demand model with learning nature on account of the early and current period purchase price, the after-sale service. Secondly, it also solves the optimal dynamic pricing strategy of manufactures, and the result show that the price of product tends to a constant value with oscillation according to the dynamic pricing rules of manufacture. Finally, the optimal price and after-sale service of every period are solved, and it proves that the sales price tends to customer's reservation price or unit cost under certain conditions.

## 2 Model assumptions and model construction

### 2.1 MODEL ASSUMPTIONS

#### 2.1.1 The market structure

Assume that there is only a monopolistic manufacturer selling a product in a particular industry. For example, it is inevitable to form a monopoly in a certain period when new products promote to the market because of the novelty of product, the adoption of new technology, enhance of function and the change of appearance, which is mainly reflected in the monopoly of proprietary technology.

#### 2.1.2 The price and production cost

Assume that  $p_t$  is the sales price of product in period t, and  $c$  is the unit production cost of product, which is equal in each period. Assume  $p_t \geq c$ , where  $t=1, 2, \dots$

#### 2.1.3 The objective function

The goals are not the same for different types of manufacture in different development period, such as the maximum of profits, value, and market share or customer satisfaction and so on. This paper assumes that the manufacture's goal is to maximize profit.

#### 2.1.4 The cost function of after-sale service

Assume it is independent for production cost and after-sale service provided by manufacture and the cost function of after-sale service is

$$C(a_t) = \frac{ka_t^2}{2}, \quad (1)$$

where k is the cost coefficient of after-sale service, this coefficient is closely related to the product type. It is higher when the complexity of product becomes greater. It means that the cost paid more when the manufacturer raises a unit of after-sale service. And  $a_t$  is the level of after-sale service for product in period t. The larger  $a_t$  indicates the cost of after-sale service provided by manufacture to consumers is greater, and the increase is in multiples square level.

#### 2.1.5 The homogeneity assumption

In addition to the same product price and after-sale service level of manufacture in each period, the other aspects in different periods are also homogeneous.

#### 2.1.6 The assumption of discount coefficient

The products of manufacture generally have a certain life cycle, so it is worth considering the time value of money. Assume the discount coefficient between the two cycles of the products is  $\delta$ , where  $\delta \in (0, 1)$ .

### 2.2 DEMAND MODEL

When manufacturer provides after-sale service to consumers, their demand function in period t can be represent as

$$D_t = M - \beta p_t + \lambda a_t. \quad (2)$$

Here,  $D_t$  is the demand function of manufacturer in period t. M represents the market capacity of product. And  $\beta$  is the price-sensitive parameter, which indicates the reverse relationship between the demand and price. Then  $\lambda$  is the sensitive coefficient of after-sales service to the demand, which shows the positive changes relation between the demand and after-sales service. The larger coefficient indicates the current demand of manufacturer,  $D_t$ , will be increased when manufacturer increases the after-sale service,  $a_t$ .

Eq. (2) portrays only the current relationship between the demand function of manufacturer and the prices, after-sale service level. In fact, the current price and after-sale service have not only certain effect on the current demand, but also influence the sales of product for the future periods. Therefore, the two assumptions are made as

follows. The first is learning hypothesis. Assume there has learning process for consumer to the deal, but this learning process of consumers can be traced back to the previous transaction only. The second is the diffusion effect hypothesis of price or after-sale service. The lower price in period t-1 will lead to the increase of the demand in period t, and the higher price will reduce the future demand. But there is positive relationship between the product demand and after-sale service. This assumption implies that the current price and after-sale service will be affected by the next consumer surplus, which is determined by the difference between them. So Eq. (2) can be rewritten as follows according to the two assumptions.

$$D_t = M - \beta p_t + \lambda a_t + ma_{t-1} - np_{t-1}, m, n \in [0, 1], \quad (3)$$

where m and n are the learning coefficients of after-sale service and price in period t-1. And  $ma_{t-1} - np_{t-1}$  is learning effect of consumer, which is determined by the transaction price and after-sale service. Here, assume the influence degree is much bigger than the last for the current market price and after-sales service levels to the demand, which is  $\beta > n$  and  $\lambda > m$ .

### 2.3 MODEL CONSTRUCTION

According to Eq. (1), Eq. (2) and the unit cost hypothesis, the profit function in period t can be written as follows.

$$\pi_t = (p_t - c)D_t - \frac{ka_t^2}{2} = (p_t - c)(M - \beta p_t + \lambda a_t + ma_{t-1} - np_{t-1}) - \frac{ka_t^2}{2}. \quad (4)$$

Then the total profits of product in infinite period can be converted into the following decision problems.

$$\pi = \max \sum_{t=1}^{\infty} \delta^{t-1} [(p_t - c)(M - \beta p_t + \lambda a_t + ma_{t-1} - np_{t-1}) - \frac{ka_t^2}{2}], \quad (5)$$

where the discount coefficient for the first stage is equal to 1. And  $p_0$  and  $a_0$  are respectively the sale price and after-sale service level in the beginning period.

Assume the optimal total profit of manufacturer is  $\Pi_t(p_t, a_t)$  from the period t to the end period. So Eq. (5) can be rewritten as the Bellman equation.

$$\Pi_t(p_t, a_t) = \max \{ [(p_t - c)(M - \beta p_t + \lambda a_t + ma_{t-1} - np_{t-1}) - \frac{ka_t^2}{2} + \delta \Pi_{t+1}(p_{t+1}, a_{t+1})] \}. \quad (6)$$

Eq. (6) shows that the dynamic programming problem is closely related to the sales price and after-sale service level, but there is not much relationship with the time. Therefore, the dynamic pricing problem with monopoly is converted into the maximization problem of profit.

### 3 Model analyses

It reveals the dynamic pricing rule of the monopolistic manufacture on account of after-sales service level according to the optimal solution and its property of model by partial differential analysis.

#### 3.1 PRICING RULES GIVEN AFTER-SALE SERVICE LEVEL

In order to solve the after-sale service level, the first derivative of  $\pi_t$  for  $p_t$  can be solved, and assume it is equal to zero according to Eq. (4) when the after sale service level is given. That is

$$\frac{\partial \pi_t}{\partial p_t} = p_t'(p_{t-1}) = \frac{M + \lambda a_t + ma_{t-1} - np_{t-1} + c\beta}{2\beta} = 0. \quad (7)$$

The second order derivative of Eq. (7) can be expressed as follows:

$$\frac{\partial^2 p_t'(p_{t-1})}{\partial p_{t-1}} = \frac{-n}{2\beta} < 0. \quad (8)$$

The profit function can reach the maximum value at a certain point because the second order derivative is less than 0. By using the mathematical induction, the optimal selling price of monopolistic manufacture is

$$p_t^* = \frac{\sum_{i=1}^t (2\beta)^{t-i} (-n)^{i-1} (M + \lambda a_{t-i+1} + ma_{t-i} + c\beta) + (-n)^t p_0}{(2\beta)^t}. \quad (9)$$

Relaxing the assumption of after-sales service level, assume that  $a_1 = a_2 = \dots = a_t = a$ , Eq. (9) can be simplified as follows:

$$p_t^* = \frac{[1 - (-n/2\beta)^t][M + (\lambda + m)a + c\beta]}{2\beta + n} + (-n/2\beta)^t p_0. \quad (10)$$

There is  $\beta > n$  according to the hypothesis of demand function. While  $t \rightarrow \infty$ , so the limit of Eq. (10) can be written as follows.

$$p_t^* = \lim_{t \rightarrow \infty} p_t^* = \lim_{t \rightarrow \infty} \left\{ \frac{[1 - (-n/2\beta)^t][M + (\lambda + m)a + c\beta]}{2\beta + n} + (-n/2\beta)^t p_0 \right\} = \frac{M + (\lambda + m)a + c\beta}{2\beta + n}. \quad (11)$$

The following conclusion can be drawn from the above analysis.

**Conclusion 1** The price of monopolistic manufacturer is  $p_t = \frac{\sum_{i=1}^t (2\beta)^{t-i} (-n)^{i-1} (M + \lambda a_{t-i+1} + m a_{t-i} + c\beta) + (-n)^t p_0}{(2\beta)^t}$  when the after-sale service level is fixed. And the product price is  $p_t = \lim_{t \rightarrow \infty} p_t = \frac{M + (\lambda + m)a + c\beta}{2\beta + n}$  while  $a_1 = a_2 = \dots = a_t = a$  and  $t \rightarrow \infty$ .

The sale price of product tends to a fixed value when the cycle of product sales is long enough and there is the same after-sale service according to the conclusion 1. This price has positive correlation with the market capacity, learning coefficient of after-sale service and unit operation cost, but there is reverse connection with the coefficient of price learning and price sensitivity. And this price does not reflect the trajectory before the price tends to be fixed, so it is worth discussing the change trajectory. Because of  $\beta > n$ , there is  $-\frac{n}{2\beta} \in (-\frac{1}{2}, 0)$  according to Eq. (10) while  $a_1 = a_2 = \dots = a_t = a$ . So the first half of  $p_t$  is always greater than 0 whatever the number of product cycle change, and the second part,  $(-\frac{n}{2\beta})^t p_0$ , will appear the motion rules of vibration with the increase of cycle number. When t is even cycle, then  $-\frac{n}{2\beta} > 0$ , but when t is odd cycle, there has  $-\frac{n}{2\beta} < 0$ . The changes of odd cycle and even cycle will lead to the regular changes of  $p_t$ .

Taking  $p_t = \frac{M + (\lambda + m)a + c\beta}{2\beta + n}$  of Eq. (11) into  $p_t$  of Eq. (10),  $p_t$  can be represented as follows.

$$p_t = [1 - (-\frac{n}{2\beta})^t] p_t + (-\frac{n}{2\beta})^t p_0 = p_t - (p_t - p_0) (-\frac{n}{2\beta})^t = \frac{M + (\lambda + m)a + c\beta}{2\beta + n} - \frac{M + (\lambda + m)a + c\beta - (2\beta + n)p_0}{2\beta + n} (-\frac{n}{2\beta})^t \tag{12}$$

There are two kinds of situations for the change rule of  $p_t$ , such as  $p_t > p_0$  and  $p_t < p_0$ .

(1) The changes rule for  $p_t > p_0$

Firstly, due to  $-\frac{n}{2\beta} \in (-\frac{1}{2}, 0)$  and  $-(p_t - p_0) < 0$ ,  $(-\frac{n}{2\beta})^t$  will reduce with the increase of t when t is even cycle, such as  $t = 2n, n = 1, 2, \dots$ . So  $p_t$  will increase with the increase of t. Secondly,  $(-\frac{n}{2\beta})^t$  will increase with the increase of t when t is odd cycle, such as  $t = 2n - 1, n = 1, 2, \dots$ , but  $p_t$  will decrease.

(2) The changes rule for  $p_t < p_0$

Firstly, due to  $-(p_t - p_0) > 0$ ,  $(-\frac{n}{2\beta})^t$  will reduce with the increase of t when t is even cycle, such as  $t = 2n, n = 1, 2, \dots$ , but  $p_t$  will also decrease with the increase of t. Secondly,  $(-\frac{n}{2\beta})^t$  will increase with the increase of

t when t is odd cycle, such as  $t = 2n - 1, n = 1, 2, \dots$ , but  $p_t$  will also increase with the increase of t due to  $-(p_t - p_0) > 0$ .

**Conclusion 2** The sales price of product,  $p_t$ , will increase(reduce) with the increase(reduce) of t due to  $a_1 = a_2 = \dots = a_t = a$  and  $p_t > p_0 (p_t < p_0)$  when t is even cycle, such as  $t = 2n, n = 1, 2, \dots$ , and there always is  $p_t < p_t^* (p_t > p_t^*)$ . And  $p_t$  will increase (reduce) with the increase (reduce) of t when t is odd cycle, namely,  $t = 2n - 1, n = 1, 2, \dots$ , and there is  $p_t > p_t^* (p_t < p_t^*)$ .

This conclusion shows the sale price of product has oscillation law with the increase of period number in a certain extent. The sales price of product is fluctuating in the short term, but tends to a fixed value in the long run. Because consumer is not familiar with all aspects of the product information when new product promotes to market, the price will fluctuate more frequently with the change of demand. And the price of product stabilizes when consumer is familiar with the product when the product is used by a certain cycle.

But this conclusion does not consider the after-sale service level, which is derived in harsh conditions. Then the following will discuss the pricing rules when the after-sale service level is a constant value, and there have only two cases analyzed, such as strictly increasing and strictly decreasing of the after-sale service

(1) The change of  $p_t$  for  $a_t < a_{t+1}$

Firstly, there have  $(-n)^{i-1} > 0$  and  $a_t < a_{t+1}$  when t is even cycle. So

$$p_t = \frac{\sum_{i=1}^t (2\beta)^{t-i} (-n)^{i-1} (M + \lambda a_{t-i+1} + m a_{t-i} + c\beta) + (-n)^t p_0}{(2\beta)^t} > \frac{\sum_{i=1}^t (2\beta)^{t-i} (-n)^{i-1} [M + (\lambda + m)a_0 + c\beta] + (-n)^t p_0}{(2\beta)^t} \tag{13}$$

Then solving the limit of Eq. (13) according to  $-\frac{n}{2\beta} \in (-\frac{1}{2}, 0)$ , so the limit of  $p_t$  is

$$\overline{p_t} = \lim_{t \rightarrow \infty} p_t = \lim_{t \rightarrow \infty} \frac{\sum_{i=1}^t (2\beta)^{t-i} (-n)^{i-1} (M + \lambda a_{t-i+1} + m a_{t-i} + c\beta) + (-n)^t p_0}{(2\beta)^t} > \lim_{t \rightarrow \infty} \frac{\sum_{i=1}^t (2\beta)^{t-i} (-n)^{i-1} [M + (\lambda + m)a_0 + c\beta] + (-n)^t p_0}{(2\beta)^t} = \overline{p_0} \tag{14}$$

where  $\overline{p_t}$  is the limit of sale price in the period t, and  $\overline{p_0}$  represents the limit of sale price for  $a_t = a_0$ .

Secondly, when t is even cycle, there is  $(-n)^{i-1} < 0$ , so

$$p_t^{\cdot} < \frac{\sum_{i=1}^t (2\beta)^{t-i} (-n)^{i-1} [M + (\lambda + m)a_0 + c\beta] + (-n)^t p_0}{(2\beta)^t} \quad (15)$$

Similarly, the limit of  $\overline{p}_t^{\cdot}$  is

$$\overline{p}_t^{\cdot} = \lim_{t \rightarrow \infty} p_t^{\cdot} < \frac{M + (\lambda + m)a_0 + c\beta}{2\beta + n} = \overline{p}_0 \quad (16)$$

(2) The change of  $p_t^{\cdot}$  for  $a_t > a_{t+1}$

Firstly, there have  $(-n)^{i-1} > 0$  and  $a_t < a_{t+1}$  when t is odd cycle, so

$$p_t^{\cdot} < \frac{\sum_{i=1}^t (2\beta)^{t-i} (-n)^{i-1} [M + (\lambda + m)a_0 + c\beta] + (-n)^t p_0}{(2\beta)^t} \quad (17)$$

The limit of Eq. (17) is

$$\overline{p}_t^{\cdot} = \lim_{t \rightarrow \infty} p_t^{\cdot} < \frac{M + (\lambda + m)a_0 + c\beta}{2\beta + n} = \overline{p}_0 \quad (18)$$

where  $\overline{p}_t^{\cdot}$  is the limit of the sale price in period t, and  $\overline{p}_0$  is the limit of the sale price for  $a_t = a_0$ .

Secondly, there has  $(-n)^{i-1} < 0$  when t is even cycle, so

$$p_t^{\cdot} > \frac{\sum_{i=1}^t (2\beta)^{t-i} (-n)^{i-1} [M + (\lambda + m)a_0 + c\beta] + (-n)^t p_0}{(2\beta)^t} \quad (19)$$

The limit of Eq. (19) is

$$\overline{p}_t^{\cdot} = \lim_{t \rightarrow \infty} p_t^{\cdot} > \frac{M + (\lambda + m)a_0 + c\beta}{2\beta + n} = \overline{p}_0 \quad (20)$$

**Conclusion 3** The limit of product sales price in period t is greater than the price limit for  $a_t = a_0$  when t tends to infinite with odd cycle and  $a_t < a_{t+1}$ , that is  $\overline{p}_t^{\cdot} > \overline{p}_0$ , but there is  $\overline{p}_t^{\cdot} < \overline{p}_0$  while t is odd cycle. The limit of product sales price is less than the price limit for  $a_t = a_0$  when t tends to infinite with odd cycle and  $a_t > a_{t+1}$ , that is  $\overline{p}_t^{\cdot} < \overline{p}_0$ , but there is  $\overline{p}_t^{\cdot} > \overline{p}_0$  while t is even cycle.

Conclusion 3 shows there has a lower bound for the sale price of product with the increase of cycle number and after-sale service level when t is odd cycle, in other words, the sale price is increasing. The sales price of product has an upper bound when t is an even cycles, that is to say the price is decreasing. In the same way, there have similar change rules for the decline of after-sale service level. It

implies the sales price is dynamic with oscillation from a certain perspective. Therefore, the sales price has rules to follow in certain conditions.

### 3.2 THE PRICING RULES FOR THE CHANGE OF PRICE AND AFTER-SALES SERVICE LEVELS SIMULTANEOUSLY

The part regards after-sales service level and sales price as the decision variables simultaneously. In order to obtain the maximum profits, manufacture should determine the optimal sale price and after-sales service level according to the actual situation of manufacture. Solving the first-order partial derivatives of Eq. (4) for  $p_t$  and  $a_t$ , and make them equal to zero, so there have

$$\frac{\partial \pi_t}{\partial p_t} = M - 2\beta p_t + \lambda a_t + m a_{t-1} - n p_{t-1} + c\beta = 0, \quad (21)$$

$$\frac{\partial \pi_t}{\partial a_t} = \lambda(p_t - c) - k a_t = 0. \quad (22)$$

After finishing Eq. (21) and Eq. (22),  $p_t$  and  $a_t$  are

$$p_t = \frac{M + \lambda a_t + m a_{t-1} - n p_{t-1} + c\beta}{2\beta}, \quad (23)$$

$$a_t = \frac{\lambda(p_t - c)}{k}. \quad (24)$$

Taking  $a_{t-1} = \frac{\lambda(p_{t-1} - c)}{k}$  of Eq. (24) into Eq. (23), there has

$$p_t(p_{t-1}) = \frac{kM - c\lambda^2 + c\beta k + (m\lambda - kn)p_{t-1} - m\lambda c}{2\beta k - \lambda^2}. \quad (25)$$

Similarly, the optimal price and after-sales service level can be derived through the mathematical induction, namely the conclusion 4.

**Conclusion 4** The optimal price and after-sales service level are equal to the initial sales price and after-sales service respectively when the profit of manufacture achieves maximization. That is

$$p_1^{\cdot} = \frac{A_1 + k(ma_0 - np_0)}{A_2}, a_1^{\cdot} = \frac{\lambda(p_1^{\cdot} - c)}{k}, \quad (26)$$

$$p_t^{\cdot} = \frac{A_3(1 - \eta^{t-1})}{A_2(1 - \eta)} + \eta^{t-1} p_1^{\cdot}, a_t^{\cdot} = \frac{\lambda(p_t^{\cdot} - c)}{k}, t = 2, 3, \dots$$

where  $A_1 = km - c\lambda^2 + c\beta k$ ,  $A_2 = 2\beta k - \lambda^2$ ,  $A_3 = \Gamma_1 - m\lambda c$ ,  $A_4 = m\lambda - kn$  and  $\eta = \frac{A_4}{A_2}$ .

Conclusion 4 means the manufacture's optimal sale price and after-sale service is affected by the initial sale prices of product and after-sale service level. The sale price and after-sale service will be gradually reduced along with the increase of the sales cycle number when the initial price is higher, but the reduction extent will be smaller. In another word, it is smaller for the initial price and after-sale service level influenced by the late price and after-sale service level. But there is no further explanation for the convergence of price, which is reduced to what extent in conclusion 4. Therefore, the following discusses the degree of convergence effect.

According to previous consumption experience, relevant alternative product prices and other product information, consumer as a whole has a subjective expectation value for product, which is the reserve price. This price is the most acceptable price for consumers. Assume the reserve price as a constant value is  $p^*$ , and it does not change with the change of cycle number. It is different for the reservation price and sales price. Assume the reserve price of consumer is higher than the sales price of product, once the product price is higher than the reserve price, so the demand will be reduced to 0. Therefore, there is  $p^* \geq p_t$ . It also assumes that the sale price is generally greater than the unit operation cost, namely  $p_t \geq c$ . From the above analysis, there has  $p_t \geq p_t^* = \lim_{t \rightarrow \infty} p_t^* \geq c$ .

So the Eq. (26) can be rewritten as follows.

$$p_t^* = \frac{A_3}{A_2(1-\eta)} - \frac{A_3\eta^{t-1}}{A_2(1-\eta)} + \eta^{t-1}p_t = \frac{A_3}{A_2(1-\eta)} + \eta^{t-1}[p_t - \frac{A_3}{A_2(1-\eta)}] \quad (27)$$

Assume  $A_5 = p_1 - \frac{A_3}{A_2(1-\eta)}$ ,  $p_t^*$  is

$$p_t^* = \frac{A_3}{A_2(1-\eta)} + \eta^{t-1}A_5 \quad (28)$$

The value of  $p_t^*$  and  $p_t^*$  are also different due to the different value of  $\eta$ , therefore there are five kinds of cases.

(1) If  $t$  is odd cycle, there has  $\eta^{t-1} > 0$ , so  $p_t^* = \lim_{t \rightarrow \infty} p_t^* = p^*$  when  $t$  tends to infinity for  $\eta < -1$  and  $A_5 > 0$ . But when  $t$  is even cycle, there has  $\eta^{t-1} < 0$ , then  $p_t^* = c$ . If  $t$  is even cycle, there has  $\eta^{t-1} > 0$ , so  $p_t^* = c$  when  $t$  tends to infinity for  $\eta < -1$  and  $A_5 < 0$ . But when  $t$  is even cycle, there has  $\eta^{t-1} < 0$ , then  $p_t^* = p^*$ .

(2) When  $t$  tends to infinity, there has  $p_t^* = \frac{A_3}{2A_2} + (-1)^{t-1}A_5 = \frac{A_3}{2A_2} + (-1)^{t-1}[p_1 - \frac{A_3}{2A_2}]$  for  $\eta = -1$ .

(3) When  $t$  tends to infinity, there has  $p_t^* = \lim_{t \rightarrow \infty} [\frac{A_3}{A_2(1-\eta)} + \eta^{t-1}A_5] = \frac{A_3}{A_2(1-\eta)}$  for  $-1 < \eta < 1$ .

(4) There has  $p_t^* = c$  for  $\eta = 1$  when  $t$  tends to infinity.

(5) There has  $p_t^* = p^*$  for  $\eta > 1$  and  $A_5 > 0$  when  $t$  tends to infinity. And there has also  $p_t^* = c$  for  $\eta > 1$  and  $A_5 < 0$  when  $t$  tends to infinity.

According to the analysis of the above five cases the conclusion 5 can be obtained as follows

**Conclusion 5** The sale price of product has four kinds of choices when the cycle number tends to infinity. Firstly, the sale price of product is equal to the unit operating cost, that is  $p_t^* = c$ . Secondly, the sale price of product is equal to the unit operating cost and the consumer's reservation price, which is  $p_t^* = \lim_{t \rightarrow \infty} p_t^* = p^*$ . Thirdly, the sale price is

$$p_t^* = \frac{A_3}{2A_2} + (-1)^{t-1}[p_1 - \frac{A_3}{2A_2}] \text{ for } \eta = -1. \text{ Finally, the sale price is } p_t^* = \frac{A_3}{A_2(1-\eta)} \text{ for } -1 < \eta < 1.$$

The sale price of product tends to the consumer's reservation price or the unit operating cost of manufacture with the increase of sale cycle while it does not consider the case for  $-1 \leq \eta < 1$  from the conclusion 5. This confirms the truth of much realistic economic life. The sale price of product is reduced gradually when the sale is sluggish for the product promoting to market or it is the end of product life cycle. And the product will exit the market while the price tends to the unit operating costs of manufacture. The sale price is high, which is generally lower than the reserve price of consumer when the product is sold well or is promote to the market for the first.

#### 4 Example analyses

The change trend of the optimal sale price and after-sale service level is studied by the example of an agricultural machinery manufacturer in the pursuit of profit maximization. In order to study the change trend, assume that the initial conditions of manufacture are  $k = 0.5, m = 20, c = 80, \lambda = 0.1, \beta = 0.3, a_0 = 20, n = 3, p_0 = 120$  and  $T_1 = 175$  according to Eq. (26). The optimal sale price and after-sale service level for the first ten periods can be derived, as shown in table 1.

TABLE 1 the calculation results for the first ten periods

periods(t)	The optimal price ( $p_t^*$ )	The optimal after-sales service ( $a_t^*$ )
1	142.07	12.41379
2	134.1241379	10.82483
3	126.8434979	9.3687
4	123.8728265	8.774565
5	123.2342014	8.64684
6	123.1580581	8.631612
7	123.152913	8.630583
8	123.1527139	8.630543
9	123.1527094	8.630542
10	123.1527094	8.630542

It is seen from table 1, the optimal price of manufacture tends to 123.152 in the seventh period, and the optimal after-sales service tends to 8.63 in the sixth period.

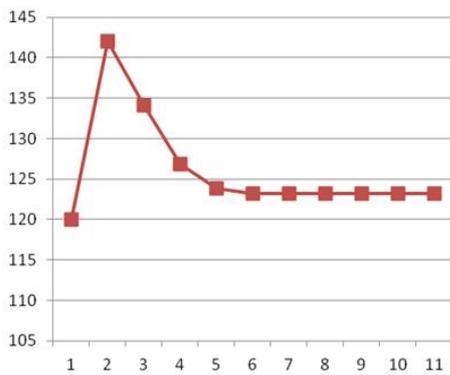


FIGURE 1 the optimal price trend

Figure 1 shows that the optimal price of manufacture is at a higher level during initial period. But it gradually reduced and tends to be a stable value, which is consumer's reservation price or unit operating costs. Because the manufacture selects the higher initial price generally in order to recover the upfront cost of as soon as possible.

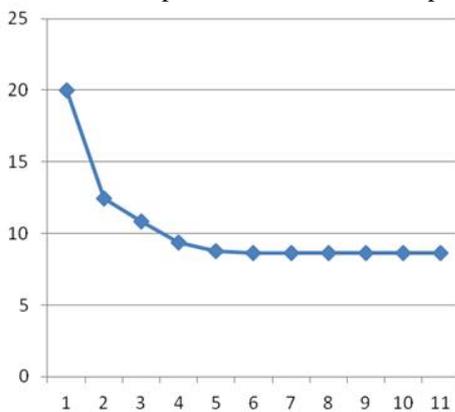


FIGURE 2 the optimal after-sales service trend

Figure 2 shows the optimal after-sale service level of will gradually reduce and tends to a constant value with the increase of cycles. Higher initial after-sale service level is also to recover the initial investment as soon as possible. Because the cost of after-sale service provide by manufacture is decreased, so the optimal after-sale service level will also reduce gradually.

### 5 Conclusions

This paper proposes the demand function of consumer with learning character, and then constructed the multi period dynamic pricing model for monopolistic manufacture according to the purchase price and after-sale service level in the prior and current period. In order to try to reveal the dynamic pricing rules of manufacture, the optimal sale price and after-sales service level is solved by dynamic programming method, then the limit value of product sale price is derived considering after-sales service. The findings show as follows. Firstly, the product pricing of manufacture fluctuates with oscillation both in the short and long term, but in the long term it tends to a constant value. Secondly, the oscillatory amplitude or degree of product price is gradually reduced with a certain convergence when the period number tends to infinity. The sale price may tend to consumer's reservation price and unit operation cost of manufacture.

There are two possible research directions to study the pricing problem. Firstly, the demand function with learning can be developed from linear to nonlinear model. Secondly, the uncertain demand function will be considered to study the effect to pricing.

### References

- [1] Crawford C M, Di Benedetto C A 2000 *New Products Management, 6th edition*, McGraw-Hill: NY
- [2] Ladany S P, Arbel A 1991 Optimal cruise-liner passenger cabin pricing policy *European Journal of Operational Research* **55**(2) 136-47
- [3] Bitran G, Caldentey R, Mondschein S 1998 Coordinating clearance markdown sales of seasonal products in retail chains *Operations Research* **46**(5) 609-24
- [4] Stephen A, Smith S, Achabal D D 1998 Clearance pricing and inventory policies for retail chains *Management Science* **44** (3) 285-300
- [5] Zhao W, Zheng Y S 2000 Optimal dynamic pricing for perishable assets with non-homogeneous demand *Management Science* **46**(3) 375-88
- [6] Maglaras C, Meissner J 2006 Dynamic pricing strategies for multiproduct revenue management problems *Journal Manufacturing & Service Operations Management* **8** (2) 136-48
- [7] Xuanming Su 2010 Optimal pricing with speculators and strategic consumers *Management science* **56** (1) 25-40
- [8] Broder J, Rusmevichientong P 2012 Dynamic pricing under a general parametric choice model *Operations Research* **60**(4) 965-80
- [9] Manaswini Bhalla 2012 Dynamic pricing under social learning with strategic consumers *IIM Bangalore Research Paper No.367* [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2114054](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2114054) /2 April 2014
- [10] Yulan Zhou 2013 Competitive dynamic pricing model based on after-sale service *International Journal of Applied Mathematics and Statistics* **50**(20) 422-30

<b>Author</b>	
	<b>Yulan Zhou, born in December, 1978, Chengdu County, Sichuan Province, P.R. China</b>
	<b>Current position, grades:</b> the Lecturer of School of Economy & Management, Southwest Petroleum University, China.
	<b>University studies:</b> received her B.B.A. in Business Management from Southwest Petroleum University of Nanchong in China. He received his M.B.A. from Southwest Petroleum University in China.
	<b>Scientific interest:</b> His research interest fields include price theory, corporate culture
	<b>Publications:</b> more than 5 papers published in various journals.
<b>Experience:</b> He has teaching experience of 9 years, has completed nine scientific research projects.	

# Study on local government public expenditure and multi-factor productivity in china based on instrument variable model

Jian Jin<sup>1</sup>, Jianxiang Wang<sup>2\*</sup>

<sup>1</sup>School of Economics, Hebei University, Baoding City, Hebei Province, China, 071002

<sup>2</sup>New Campus Administrative Committee, Hebei University, Baoding City, Hebei Province, China, 071002

Received 19 March 2014, www.tsi.lv

## Abstract

Based on the conventional C-D Production Function Model, this paper adopted Instrument Variable Model to measure the multifactor productivity growth of 223 cities at prefecture level and above in China, and probed into its relationship with local government public expenditure. It is shown that relationship between total public expenditure of local government and city multifactor productivity growth in China is significantly negative, which does not mean that local government public expenditure in China is inefficient, but because a considerable part of it is put into social security, health and medical care, and other public services. Further research by different productivity levels show that the faster productivity grows, the more deeply market-driven is the economics, the weaker is the negative correlation of local government public expenditure and productivity growth. Science & technology and educational expenditure of local government positively affect multifactor productivity growth in China cities significantly, however in varying degrees.

*Keywords:* instrument variable model, local government public expenditure, multi-factor productivity, decision reference

## 1 Introduction

Literatures show that there is a close relationship between public expenditure and productivity growth; public expenditure may enhance productivity by technical progress or by improving efficiency of single productive elements such as capital or labour. Many literatures have probed into this problem, such as Arrow, Kenneth and Kurz, Mordecai in [1], who first brought public investment into macroscopical production function model; Barro, who proposed that public service positively affect production in [2]. Adam [3] argued that quality and productivity in delivering and administering public service was of great importance. The United Nations [4] suggested that national accounts should measurement performance of the general government, and International Monetary Fund [5] suggested detailed procedure of government finance statistics. OECD (2000) inspected China's public expenditure problem, and talked about her efficiency of public expenditure in [6]. John Baldwin, Wulong Gu and Ryan Macdonald [7], Sir Tony Atkinson [8], the UK Office for National Statistics [9], Statistics New Zealand [10] and OECD [11] shared their experience and proposals of government performance measurement. Some other scholars, such as Dongping Fu in [12], Zhenye Li in [13], Jiejun Zhu in [14] and Ge Jin in [15] etc [16-21], have also investigated into problems about public expenditure and productivity.

On one hand, cities that assemble kinds of productive elements are cores of economic growth, and on the other

hand, investment to cities have been much more than that to rural areas in China, research on relationship between local public expenditure of government and city productivity growth is of more importance. Taking these factors into consideration, this paper probed into relationship between public expenditure of local government and city multifactor productivity.

Reviews of literatures show that though many literatures focus on the relationship between public expenditure and productivity growth, few is focused on local government public expenditure and productivity growth, and few is based on Instrument Variable Models. The paper adopts the Instrument Variable Model to probe into relationship between local government public expenditure and multifactor productivity of prefectural-level and above cities in China.

Based on conventional Cobb-Dauglass Production Function Model, this paper positively analyze with data of 223 cities at prefecture level and above of 30 provinces in China from 1990 to 2009. Because many data are not available, Tibet is excluded. And some cities at prefecture level have become prefecture level city not long ago, there are very few data about them, they are excluded too. According to our analysis, relationship between local government public expenditure and city multifactor productivity in China is significantly negative, which does not mean that local government public expenditure in China is inefficient, but because a considerable part of it is put into social security, health and medical care, and other public services. Further research by different

\* *Corresponding author* e-mail: wangjianxiang626@163.com

productivity levels show that the faster productivity grows, the more deeply market-driven is the economics, the weaker is the negative correlation of local government public expenditure and productivity growth. Science & technology and educational expenditure of local government positively affect multifactor productivity growth in China cities significantly, however in varying degrees.

The paper is organized as follows: in the next section, we show the source and processing of data. Section 3 gives the Instrument Variable Model of multifactor productivity we adopted. Then we test the model in the following section. Section 5 concludes.

## 2 Variables and Data

Taking availability of data into account, the sample of this paper takes from 223 cities at prefecture level and above of 30 provinces from 1990 to 2009. The reason we choose cities at prefecture level and above is that they are the main body of city function and are of stronger agglomeration effects, so are of better representativeness. Cities at prefecture level and above are more than 223. We choose only 223 ones for the following reasons: Because many data are not available, Tibet is excluded; and some cities at prefecture level have become prefecture level city not long ago, they are excluded too. Excluding cities in Tibet and those upgraded to cities of prefecture level, the number of the remained cities that meet our research criteria is 223.

For these cities at prefecture level and above, *China City Statistical Yearbook* and *Comprehensive Statistical Data and Materials on 50 Years of New China Cities* provide statistical materials for two kinds of concept for city. The first concept of city means city proper, that is to say, the downtown area and the suburb area, while countries and cities of country-level excluded. The other concept of city means the whole city, that is to say, not only the downtown area and the suburb area, but also countries and cities of country-level affiliated with the prefecture level city are included. Here we adopt the first concept; the reason is that countries and cities at country level are not main body of city function, while data for city proper are preferred.

To carry out our research, we need data for output, capital and labor input as well as local government public expenditure of the 223 cities. Data of 1990 to 1998 is taken from *Comprehensive Statistical Data and Materials on 50 Years of New China Cities*, 1999 to 2009 from *China City Statistical Yearbook*. Output data are GDP, labor input data are number of employed persons. Theoretically speaking, capital input shall use data of fixed capital stock per year. Because of availability of data, we took total investment in fixed assets as a replacement. For public expenditure of local government, we take intra-budgetary government expenditure; because it is typically relevant to size of city, we adjust them with GDP, and use the relative indicator of ratio of public

expenditure of local government to GDP. Some data for these indicators are missed, and we fill them with Moving Average Method.

This paper will also study relationship between expenditure of local government and multifactor productivity of cities in China by regions. The general processing divides China into three regions by geographic location, that is to say, Eastern Region, Central Region and Western Region. This method is very simple, while of much disadvantage in research. It is well known that productivity level and productivity growth of cities in the same geographic district may differ significantly. Some times, there may even be significant differences among productivity level and productivity growth of the cities in the same province. To avoid this problem, this paper used the cluster analysis method based on productivity growth during the period of 1990 to 2009 to divide 223 cities into groups, and then probe into the relationship of their local government public expenditure and city productivity of each group respectively.

## 3 Instrument Variable Model of Multifactor Productivity

This paper handles with panel data. According to econometric theory and practice, dynamic models with panel data typically are troubled with endogenous explanatory variable problem. To settle this problem, productivity analysts seek help from Instrument Variable Model. We have tried some other methodologies in our previous research, none of them served well than Instrument Variable Model. Therefore, this paper will also make use of this method.

The key step of Instrument Variable Method is to seek appropriate instruments. Inappropriate instrument variable usually leads to bad, even wrong conclusions. Many econometricians devote to research on choice of better instrument variables. In the fields of productivity measurement with Instrument Variable Model, Anderson & Hsiao in [22], Arrelano & Bond in [23] and Blundell & Bond in [24] and [25] have probed into this problem and suggested practicable instruments respectively. Jian Jin reviewed the Instrument Variable Model of productivity growth measurement in literatures in [26].

Anderson and Hsiao's research is based on the differenced form of the original equation,  $y_{it} = \rho y_{i,t-1} + x'_{it}\beta + \alpha_i + \varepsilon_{it}$ . Difference cancels the individual fixed effects possibly correlate with the exogenous variables which means  $E(x'_{it}\alpha_i) \neq 0$ . But the difference of the lagged endogenous variable  $y_{i,t-1} - y_{i,t-2} = \rho(y_{i,t-2} - y_{i,t-3}) + (x'_{i,t-1} - x'_{i,t-2})\beta + \varepsilon_{i,t-1} - \varepsilon_{i,t-2}$  is obviously correlated with the error term  $\varepsilon_{it} - \varepsilon_{i,t-1}$ . To avoid this problem, Anderson and Hsiao suggested to use level instruments  $y_{i,t-2}$  or the lagged difference instruments  $y_{i,t-2} - y_{i,t-3}$

as instruments for the differential regression estimators  $y_{i,t-1} - y_{i,t-2}$  and proved that level instruments superior to the latter.

In the same way, Arrelano & Bond (1991) eliminates the individual effects by differencing to get equation:

$$y_{it} - y_{i,t-1} = \rho(y_{i,t-1} - y_{i,t-2}) + (x'_{it} - x'_{i,t-1})\beta + \varepsilon_{it} - \varepsilon_{i,t-1} \tag{1}$$

And under some certain presumption, Arellano & Bond proved that  $\{x_{i,t-j}, y_{i,t-j} : j \geq 2\}$  are efficient instrument variables for this differential form equation. Now for each year, researchers on productivity measurement can find efficient instruments. For  $t=T$ , Equation (1) changes into  $y_{iT} - y_{i,T-1} = \rho(y_{i,T-1} - y_{i,T-2}) + (x'_{iT} - x'_{i,T-1})\beta + \varepsilon_{iT} - \varepsilon_{i,T-1}$ , and we have a series of instruments variables  $y_{i1}, y_{i2}, \dots, y_{i,T-2}, x'_{i1}, x'_{i2}, \dots, x'_{i,T-1}$ .

There is a serious problem. Because making use of information contained in differences only, the estimator suggested by Arellano & Bond is rather inefficient when instruments are weak. Blundell & Bond in [25] proved that both the elasticities of output to capital and to labour are very small and inaccurate.

Aimed at shortcomings of Arellano-Bond estimator, Blundell & Bond in [24] suggest making use of additional level information beside the differences, combining moment restrictions on differential and level instruments, and resulting in a so-called GMM system-estimator. According to Blundell & Bond, for  $t=T$ ,  $dy_{i1}, dy_{i2}, \dots, dy_{i,T-1}, dx'_{i1}, dx'_{i2}, \dots, dx'_{i,T-1}$  are usable instruments.

Once instruments found, we can solve equation and compute productivity growth with Instrument Variable Method.

#### 4 Measurement of productivity and clustering division of cities

The model handled with is as follow:

$$y_{it} = \beta_k k_{i,t-1} + \beta_l l_{i,t-1} + \alpha_i + f_i + \eta_t + \varepsilon_{it}, \tag{2}$$

where  $k$  and  $l$  are capital and labour input respectively,  $f_i$  denotes fixed effects of each city,  $\eta_t$  denotes time tendency faced by all the cities.

At first, we exclude time mean values of all the variables, so that the following processing does not need to deal with time specific dummy variables. Then we drops the individual fixed effects by first order difference, the equation now is:

$$y_{it} - y_{i,t-1} = \beta_k (k_{i,t-1} - k_{i,t-2}) + \beta_l (l_{i,t-1} - l_{i,t-2}) + (\varepsilon_{i,t} - \varepsilon_{i,t-1}). \tag{3}$$

We solve this model with difference instrument variables method and Eviews 5.1, compute multifactor productivity with surplus method, which is then regressed with adjusted public expenditure of local government.

It is well known that there exists imbalance in economic development of different regions and cities, and productivity in different cities also differs significantly. To find out relationship between local government public expenditure and city productivity in cities of different economic level, we clustered the 223 cities by their productivity growth during the period of 1990 to 2009.

TABLE 1 Cluster of cities by productivity growth rate during 1990 to 2009

Productivity growth rate	Cities included
<b>higher (47 cities)</b>	Beijing, Tianjin, Shijiazhang, Tangshan, Taiyuan, Shenyang, Dalian, Anshan, Changchun, Haerbin, Daqing, Shanghai, Nanjing, Wuxi, Xuzhou, Changzhou, Suzhou, Hangzhou, Ningbo, Wenzhou, Fuzhou, Taizhou, Xiamen, Nanchang, Ji'nan, Qingdao, Zibo, Dongying, Yantai, Zhengzhou, Wuhan, Changsha, Guangzhou, Shenzhen, Zhuhai, Shantou, Foshan, Zhongshan, Dongguan, Chongqing, Chengdu, Panzihua, Mianyang, Kunming, Yuxi, Xi'An, Urumchi
<b>Lower (80 cities)</b>	Xingtai, Chengde, Cangzhou, Langfang, Hengshui, Jincheng, Shuozhou, Wuhai, Fuxin, Tieling, Chaoyang, Liaoyuan, Tonghua, Baishan, Songyuan, Baicheng, Jixi, Hegang, Shuangyashan, Qitaihe, Mudanjiang, Heihe, Suqian, Quzhou, Tongling, Anqing, Huangshan, Chuzhou, Sanming, Nanping, Jingdezhen, Xinyu, Yingtan, Kaifeng, Hebi, Jiaozuo, Xuchang, Luohe, Sanmenxia, Shangqiu, Xiaogan, Huanggang, Shaoyang, Yiyang, Chenzhou, Huaihua, Zhangjiagang, Chaozhou, Meizhou, Shanwei, Heyuan, Yangjiang, Qingyuan, Yunfu, Wuzhou, Qinzhou, Guigang, Fangchenggang, Sanya, Zigong, Luzhou, Deyang, Guangyuan, Suining, Neijiang, Leshan, Yibin, Nanchong, Liupanshan, Tongchuan, Yan'An, Hanzhong, Weinan, Jiayuguan, Jinchang, Baiyin, Tianshui, Xi'Ning, Yinchuan, Shizuishan
<b>intermediate (96 cities)</b>	Qinhuangdao, Handan, Baoding, Zhangjiakou, Datong, Yangquan, Changzhi, Hohhot, Baotou, Chifeng, Fushun, Benxi, Dandong, Jinzhou, Yingkou, Liaoyang, Panjin, Huludao, Jilin, Siping, Qiqihar, Yichun, Jiamusi, Nantong, Lianyungang, Huaiyin, Yancheng, Yangzhou, Zhenjiang, Taizhou, Jiaxing, Huzhou, Shaoxing, Jinhua, Zhoushan, Hefei, Wuhu, Bengbu, Huainan, Maanshan, Huaibei, Fuyang, Putian, Quanzhou, Zhangzhou, Longyan, Pingxiang, Jiujiang, Zaozhuang, Weifang, Jining, Taian, Dezhou, Weihai, Linyi, Laiwu, Rizhao, Luoyang, Pingdingshan, Anyang, Xinxiang, Puyang, Nanyang, Huangshi, Shiyan, Jingzhou, Yichang, Xiangfan, Ezhou, Jingmen, Zhuzhou, Xiangtan, Hengyang, Yueyang, Changde, Yongzhou, Shaoguan, Jiangmen, Zhanjiang, Maoming, Huizhou, Zhaoqing, Jiayang, Nanning, Liuzhou, Guilin, Beihai, Yulin, Haikou, Guiyang, Zunyi, Qujing, Baoji, Xianyang, Lanzhou, Kramer Iraq

Cities are clustered into three groups. The outcome of the cluster shows that, cities with higher productivity growth rate are mostly with higher economic growth rate, the four municipalities and most of the provincial capital cities are classified into this group, others are most cities in coastal provinces with higher open degree; cities with lower productivity are mostly low-economic-level ones, some of whom are cities in the Western Region, and some had upgraded into prefecture level cities just a few years ago. Most cities of intermediate level productivities are of intermediate economic levels, having a long developing history, some of whom may had experienced some higher productivity growth rate periods, while have slowed down ever since about 15 years ago.

**5 Relationships between Public Expenditure of Local Government and City Multifactor Productivity**

**5.1 FOR 223 CITIES AS A WHOLE**

According to the relevant economic theory, besides government public expenditure, there are many other factors such as economic development level, economic structure, education level of the labours, etc., that influence city productivity growth. To control influence of other factors than public expenditure of local government on productivity, we take economic structure, education level of labours and economic development level into accounts. Then we have model of 223 cities as follow:

$$\ln a = 4.078 - 1.354 \ln g + 0.294 \text{sein} + 0.289 \text{tein} - 0.030 \text{stu} - 0.084 \ln \text{pgdp} \quad (4)$$

(5.55)                      (-3.49)                      (2.84)                      (2.41)  
(-2.02)                      (-3.61)

where *a* is multifactor productivity growth, *g* public expenditure of local government, *sein* ratio of the secondary industrial value-added in GDP, *tein* the tertiary industrial value-added in GDP, *stu* number of students enrolment in regular institutions of higher education, and *pgdp* per capita GDP. The secondary and the tertiary industrial value-added in GDP reflects influence of economic structure on city productivity, number of students enrolment in regular institutions of higher education reflects that of education level of the labours, and per capita GDP reflects that of economic development level.

By Equation (4), it seems that relationship between public expenditure of local government and city multifactor productivity in China is negative, public expenditure of local government does not enhance productivity growth. This conclusion is inconsistent with normal economic knowledge.

To find out the reason of the outcome of Equation (4), taking in consideration of the fact that the parts of public expenditure playing a key role in productivity are expenditures on science & technology and education, we regression with science & technology expenditure and

education expenditure as independent variables, economic structure, education level of labors and economic development level as controlled variable. It shows that influence of the controlled variables is all insignificant, so they are all dropped out of the model, and we have:

$$\lg a = 2.112 + 0.056 \ln \text{sci} + 0.178 \ln \text{edu} \quad (5)$$

(27.75)                      (5.02)                      (12.62)

where *a* is still multifactor productivity growth, and *sci* local government expenditure on science & technology fields and *edu* local government expenditure on education.

Equation (5) shows that local government expenditure on science & technology and education enhances multifactor productivity in these 223 cities significantly. Increase of 1 percent in local government expenditure on science & technology leads to multifactor productivity growth of 0.056 percent, and increase of 1 percent in local government expenditure on education leads to multifactor productivity growth of 0.178 percent.

**5.2 RELATIONSHIP OF CITIES OF DIFFERENT PRODUCTIVITY LEVEL**

We had clustered cities into three groups in accordance with their productivity. Now we shall check influence of public expenditure of local governments on multifactor productivity by group.

For the first group, we have

$$\lg a_1 = 1.149 - 0.712 \lg g_1 + 0.608 \text{sein}_1 + 0.472 \text{tein}_1 - 0.040 \lg \text{stu}_1 - 0.070 \lg \text{pgdp}_1 \quad (6)$$

(1.097)                      (-1.99)                      (4.18)                      (2.27)  
(-2.02)                      (-1.66)

where the subscript “1” means group one, that is to say, cities with higher multifactor productivity during the period of 1990 to 2009.

Equation (6) shows that for cities with higher productivity, total public expenditure of local governments is negatively correlation with productivity growth; an increase of 1 percent in local government public expenditure leads to a decrease of 0.712 percent in city multifactor productivity. That is obviously contrary to the aim of the government.

Taking into the fact that parts of the government public expenditure are put into social security, social assistance and pension, and that these expenditure has very little to do with city multifactor productivity, we set up another model with science & technology expenditure and education expenditure of the local government as independent variables, economic structure, education level of labours and economic development level as controlled variable, city multifactor productivity still the dependent variable. The regression equation shows that:

$$\lg a_1 = 0.620 - 0.0011 \lg sci_1 + 0.0461 \lg edu_1 + 0.517 \lg sei_1 + 0.268 \lg stei_1 - 0.0491 \lg stu_1, \quad (7)$$

(0.69)            (-0.05)            (1.46)            (3.92)            (1.96)            (-2.52)

where the subscript "1" means group one, cities with higher multifactor productivity during the period of 1990 to 2009.

That is to say, for cities with higher productivity, public expenditure of local governments on science & technology does not enhance city multifactor productivity growth significantly, while an increase of 1 percent in that on education enhances productivity growth by 0.046 percent.

For cities with lower productivity, we carry out the same regression. Firstly, we regress with local government public expenditure as independent variable, then with local government public expenditure on science & technology and education as independent variables. Now we have:

$$\lg a_3 = 6.025 - 2.0871 \lg g_3 - 0.0331 \lg stu_3, \quad (8)$$

(19.49)            (-5.80)            (-3.56)

$$\lg a_3 = 3.604 - 0.0271 \lg stu_3 + 0.0221 \lg sci_3 + 0.0621 \lg edu_3, \quad (9)$$

(22.3)            (-2.86)            (1.54)            (3.44)

where the subscript "3" means group three, that is to say, cities with lower multifactor productivity during the period of 1990 to 2009.

Equation (8) shows that just as situations of all the 223 cities and in higher productivity cities, total public expenditure of local government of cities with lower productivity is negatively correlation with productivity growth, the controlled variable education level of labours, denoted by the number of students enrolment in regular institutions of higher education, shows the same negative correlation with multifactor productivity. Other controlled variables do not influence the dependent variable significantly, and are dropped out of the equation.

Equation (9) shows that public expenditure of local governments on science & technology and education and the controlled variable education level of labours influenced productivity significantly, public expenditure positively and education level of labours negatively, while other controlled variables are all dropped out. Coefficients 0.022 and 0.062 mean that an increase of 1 percent in public expenditure of local governments on science & technology enhances city multifactor productivity growth by 0.022 percent, and that of education enhances it by 0.062 percent.

For cities with intermediate productivity, public expenditure of local governments influences productivity significantly, while all controlled variable does not work well so that they are taken out of the equation. Both total public expenditure of local government and expenditure on science & technology and education perform the same. The regression equations for the two situations are:

$$\lg a_2 = 5.084 - 1.1581 \lg g_2, \quad (10)$$

(18.61)            (-3.44)

$$\lg a_2 = 3.331 + 0.0071 \lg sci_2 + 0.0831 \lg edu_2, \quad (11)$$

(24.67)            (0.586)            (4.59)

where the subscript "2" means the second group, cities with intermediate multifactor productivity.

Just as situations in cities with higher and lower productivity, total public expenditure of local governments with intermediate productivity is negatively correlation with their multifactor productivity growth, while public expenditure of local governments on science & technology and education enhance productivity growth significantly, an increase of 1 percent in government public expenditure enhances productivity growth by 0.007 and 0.083 percent respectively.

### 5.3 COMPREHENSIVE ANALYSES

Putting the entire three productivity situation together, we find that in each group, just as the situation of 223 cities as a whole, total public expenditure of local government negatively influence the city multifactor productivity, while public expenditure of local government on science & technology and education positively enhance it significantly. That is not inconsistent. Nor the negative coefficients in equations for total public expenditure are inconsistent with economic theory. In fact, the reason is that local government public expenditure is divided into many parts, such as expenditure on science & technology, on education, on social security, on social support, on pension, and so on. Most of the expenditure does not enhance city multifactor productivity significantly, except for that on science & technology and on education. So it is not unusual that the coefficients for total government public expenditure are negative, while at the same time, local government public expenditure on science & technology and on education enhances city multifactor productivity significantly.

For relationship between public expenditure of local government on science & technology and multifactor productivity, the higher the productivity, the lower the regression coefficient is. That means that for cities with higher productivity, effect of public expenditure of local government on science & technology on productivity is less, while for cities with lower productivity, the effect is more obvious. The reason is that, for cities with higher productivity, because of higher economic level, the level of science & technology and degree of marketization is higher, productivity growth depends more on the market, so influence of the same amount of public expenditure on science & technology is less obvious than cities with lower productivity; for cities with lower productivity, the level of science & technology and degree of marketization is lower, and their city productivity growth depends more on government support, so the influence of public expenditure on science & technology is more obvious.

We also notice that in cities with higher and lower productivity level, just as the situation of 223 cities as a whole, educational level of labours has negative correlation with city multifactor productivity. There are different reasons. For higher-productivity-cities, social and economic development level are relatively higher, educational level of labours is also higher, according to the law of diminishing marginal benefit, educational level of labours does not show obvious enhancing effect on city multifactor productivity. As for the cities with lower multifactor productivity, the social economic development level and education development level are all lower, though the State and all levels of government attached more importance to the development of education, because of the lag of the mechanism of education to enhance productivity growth and economic growth, the expenditure of local government on education in these cities does not show obvious effect on city multifactor productivity at the present stage.

## 6 Conclusions

On the basis of Instrument Variable Model, positive analysis with panel data of 223 cities at prefecture level and above in China shows that, total public expenditure of local government and city multifactor productivity correlation negatively, while public expenditure of local governments on science & technology and on education has positive correlation with productivity. Regression coefficients of total public expenditure for higher, intermediate and lower cities are -0.712, -1.158 and -2.087 respectively. The reason is that, public expenditure of local government are most put into public services, which includes not only science & technology and education that enhance productivity growth significantly, but also social security, medical and health care and other analogous public service that do not boost the economic development remarkably.

For cities with higher productivity, because of higher economic level, demands for social security, medical and

health care and other analogous public service demands are less, the local governments can put more of its expenditure on science & technology and education; while for cities with lower productivity, because of lower economic level, demands for social security, medical and health care and analogous public service demands are more, the local governments of those cities must put more for these demands, so expenditure on science & technology and education have to be a less proportion.

Relationship between public expenditure of local government on science & technology and multifactor productivity in cities of different multifactor productivity level shows that, the higher the productivity, the lower the regression coefficient is. The reason is that, for cities with higher productivity, the level of science & technology and degree of marketization is higher, productivity growth depends more on the market, influence of public expenditure on science & technology is less obvious than that with lower productivity; for cities with lower productivity, just the opposite.

As for relationship between public expenditure of local government on education and multifactor productivity, cities with intermediate productivity have the highest coefficient, 0.083, and that of cities with higher and lower productivity are 0.046 and 0.062 respectively. It shows that, for cities with higher productivity, because of higher level of economy and education, marginal effect of education on productivity growth decreases; and for cities with lower productivity, because of weaker economic and education foundation, the effect of public expenditure of local and central government on education of these cities in recent years has not fully manifest itself due to mechanism of action of education on productivity and economic growth.

## Acknowledgments

The work was financially supported by National Planning Office of Philosophy and Social Science of China (No.12BTJ012).

## References

- [1] Kenneth J A, Mordecai K 1970 *Public Investment, the Rate of Return and Optimal Fiscal Policy* John Hopkins Press: Baltimore
- [2] Barro R J 1990 Government Spending in a Simple Model of Endogenous Growth *Journal of Political Economy* 98(5) 103-5
- [3] Jr Adam E E 1979 Quality and Productivity in Delivering and Administering Public Services *Public Productivity Review* 3(4) 26-40
- [4] United Nations Statistical Commission 2008 *System of National Accounts* 2008 <http://unstats.un.org/unsd/nationalaccount/sna2008.asp/> 26 Feb. 2014
- [5] International Monetary Fund 2001 *Government Finance Statistics Manual 2001* <http://www.imf.org/external/pubs/ft/gfs/manual/> 18 Mar 2014
- [6] OECD 2006 *Challenges for China's Public Spending: Toward Greater Effectiveness and Equity* OECD Publishing Paris
- [7] Baldwin J, Macdonald R, Gu W 2010 Integrated Productivity Accounts Contributions to the Measurement of Capital *The Canadian Productivity Review* 6 12-25
- [8] Atkinson T 2005 *Atkinson Review: Final Report, Measurement of Government Output and Productivity for the National Accounts*, Palgrave Macmillan: London
- [9] UK Office for National Statistics 2007 *The ONS Productivity Handbook*, Palgrave Macmillan: London
- [10] Statistics New Zealand 2010 *Measuring Government Sector Productivity in New Zealand: a Feasibility Study* <http://www.stats.govt.nz/> 22 Mar 2014
- [11] OECD 2001 *The OECD Productivity Manual: A Guide to the Measurement of Industry-Level and Aggregate Productivity* <http://www.oecd.org/std/productivity-stats/> 16 Apr 2014
- [12] Fu D 2005 Government Outlay and Economy Growth Taisheng from New View Angle 11 34-6 (*In Chinese*)
- [13] Li Z, Jin Y 2006 A Comparative Study of Productivity and Economic Growth Effects Between Public and Private Capital of Yangtse Delta China *Industrial Economy* 3 19-23 (*In Chinese*)
- [14] Zhu J, Hu Y 2005 Government Outlay Effect Ion Improve and Economy Growth *Contemporary Finance & Economics* 2 22-7 (*In Chinese*)

[15] Jin G, Shi J 2010 Multi-Types of Public Expenditure and Economic Growth: A Framework Economic Research Journal 7 43-56 (In Chinese)

[16] Wang Y, Zhang Y 2012 Research on Efficiency and Efficiency Dynamic Change of China Accounting Firms Based on DEA-Malmquist Index Model Advances in Information Sciences and Service Sciences 4(12) 182-92

[17] Um J H, Park S B, Choi H, Jung H 2012 Development of Bio-Cloud Service for Genomic Analysis Based on Virtual Infrastructure Advances in Information Sciences and Service Sciences 4(13) 297-307

[18] Xu L, Bao S, Mai Y 2011 Comparative Researches of Financial Productivity Journal of Convergence Information Technology 6(8) 235-43

[19] Hou O C L, Hsu H, Yang J M 2010 An Empirical Investigation of Research Productivity on Text Mining - Bibliometrics View Journal of Next Generation Information Technology 1(3) 29-35

[20] Faed A, Ashouri A, Wu C 2011 Maximizing Productivity Using CRM Within the Context of M-Commerce International Journal of Information Processing and Management 2(1) 33-42

[21] Qi W 2013 The Analysis of Network Structure Model and Innovation Network Characteristics for Industrial Cluster: using Complicated Network Perspective Journal of Digital Information Management 11(4) 277-83

[22] Anderson T W, Hsiao C 1982 Formulation and Estimation of Dynamic Models Using Panel Data Journal of Econometrics 18 47-82

[23] Arellano M, Bond S 1991 Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment equations Review of Economic Studies 58 277-97

[24] Blundell R W, Bond S R 1998 Initial Conditions and Moment Restrictions in Dynamic Panel Data Models Journal of Econometrics 87 115-43

[25] Blundell R W, Bond S R 2000 GMM Estimation with Persistent Panel Data: An Application to Production Functions Econometric Reviews 19(3) 321-40

[26] Jin J 2009 Instrument Variable Method for productivity growth measurement Journal of Management Science & Statistical Decision 6(3) 5-14 (In Chinese)

Authors	
	<p><b>Jian Jin, born in June, 1972, Baoding City, Hebei Province, China</b></p> <p><b>Current position, grades:</b> Professor of School of Economics, Hebei University, China.  <b>University studies:</b> Ph.D in Statistics at Dongbei University of Finance &amp; Economics in China.  <b>Scientific interest:</b> Statistics, Quantitative Economics.  <b>Publications:</b> More than 50 papers published in various journals.  <b>Experience:</b> teaching experience of 20 years, 15 scientific research projects.</p>
	<p><b>Jianxiang Wang, born in May, 1967, Baoding City, Hebei Province, China</b></p> <p><b>Current position, grades:</b> Lecturer of New Campus Administrative Committee, Hebei University, China.  <b>University studies:</b> Bachelor of Economics in Economics from Hebei University in China.  <b>Scientific interest:</b> Statistics, Quantitative Economics.  <b>Publications:</b> More than 30 papers published in various journals.  <b>Experience:</b> teaching experience of 24 years, 7 scientific research projects.</p>

# Study on pre-loan evolutionary stable strategy of bank-enterprise for preventing moral hazard

Qihong Zheng<sup>1, 2\*</sup>, Liangrong Song<sup>1</sup>

<sup>1</sup>Business School, University of Shanghai for Science & Technology, Shanghai, China, 200093

<sup>2</sup>College of Computer Science and Information Technology, Zhejiang Wanli University, Ningbo, China, 315100

Received 14 January 2014, www.tsi.lv

## Abstract

Based on the information asymmetry between enterprises and banks at the stage of loan application, a replicator dynamic model of bank-enterprise evolution at the pre-loan stage was established and analysed by using the evolutionary game theory (EGT) and the stability theory of nonlinear differential equation. A numerical simulation was also performed in details, which displayed intuitively how banks and enterprises achieved stable cooperation through long-term evolution. The results showed that, to effectively prevent pre-loan moral hazard, it was vitally important for commercial banks to improve their screening ability, increase disguised costs of enterprises, and formulate proper sanctions and appropriate amounts of penalty as per the local loan atmosphere also with the profit which the loan investment projects will make.

*Keywords:* information asymmetry, evolutionary game, duplicator dynamic model, decision-making

## 1 Introduction

A principal-agent relationship can be used to describe the relationship between commercial banks and enterprises that they issue loans to. In the case of information asymmetry, enterprises may adopt some practices, which go against banks' fund security measures when pursuing maximum profit, causing undesirable moral hazard. From the submission of the loan application to the loan contract signing (i.e., the pre-loan stage), enterprises aiming to obtain the loan may lie about their actual financial situations, loan repayment abilities, credibility, and the size of the risks involved in the project to be invested in. Once the loan is approved, it is easy for bad debt to be created. Banks should be highly cautious in choosing appropriate enterprises applying for loan, and make the correct decisions to prevent moral hazard, which is very important.

Stiglitz and Weiss studied the issue of bank-enterprise decisions under the hypothesis of information asymmetry in which banks would limit their loan approval ratios, causing credit rationing. Stiglitz and Weiss concluded that if banks increased interest rates, debtors might be stimulated to make high-risk investments, causing bigger moral hazard [1]. Bester further explored the role of mortgages in bank-enterprise decisions, and concluded that proper selection of loan interest rates and mortgages could screen out high-risk contracts [2]. Blumberg analyzed the loan application situations of enterprises at start-up and found that the most important factors for loan rejection were: loan commitment, loan repayment sign, and the success rate of the project invested in [3].

Xiaohong Dong discussed that in single-stage credit decisions, the harm brought by moral hazard could be overcome by a certain amount of mortgages, but in the long-term, for cooperation between banks and enterprises, one method to overcome moral hazard for banks could be offering certain preferential measures to enterprises [4]. Yanxi Li constructed a universal model of bank-enterprise decisions to prevent moral risk by setting incentives [5]. Xintian Zhuang proposed a pricing model of loan interest rates, using the maximum principle under conditions of information asymmetry [6]. Sulin Pang studied moral risk aversion methods in the cases of combined and non-combined risks, and proved that sufficient mortgages and proper ration could reduce the pre-loan moral hazard of bank-enterprise decisions by creating a credit risk decision model [7, 8].

Literatures mentioned above chose to convert credit contracts into either an optimal control problem or a classic game, in order to find the optimal interest rate, ration, or mortgage value. Such methods generally have relatively strict assumed conditions that banks should be isolated from enterprises and they also should be required to be "rational entities". However, in reality, banks and enterprises belong to different groups and the individual decisions in a group may be affected by the decisions of others. So almost no "complete rationality" exists. The evolutionary game theory can help solve this kind of problems. The EGT is based on "bounded rationality", in which individual decisions are mutually influenced and realized through the processes of imitation, study and mutation, etc. With a dynamic analysis of the decision-making behaviours, the EGT can help to derive

\* Corresponding author e-mail: qihong1010@163.com

conclusions, which reflect real-world situations. Taylor was the first to propose the duplicator dynamic model based on the EGT [9]. The model, derived from biological evolution, combines an evolutionary stable strategy with a replicator dynamic to simulate population evolution processes and stable states. Many scholars have subsequently carried out researches on the model. In recent years, the EGT has been applied to economic management applications such as inter-enterprise cooperation [10]; industrial cluster competition and cooperation behaviour [11], financial innovation [12], and credit guarantee [13].

In terms of using EGT to solve practical problems, the existing literatures lack analysis of stable points in depth, and lack numerical simulation. Based on the EGT described in the literatures [14, 15] and the stability theory of nonlinear differential equation, this paper studies how banks design loan contracts and sanction mechanisms at the pre-loan stage, to help banks and enterprises reach a stable cooperation state through long-term evolution, while effectively preventing pre-loan moral hazard.

**2 Evolutionary Game Model of Bank-Enterprise at the Pre-Loan Stage**

**2.1 MODEL CRITERIA AND PAY-OFF MATRIX OF BANK-ENTERPRISE AT THE PER-LOAN STAGE**

At the loan application stage, commercial banks and enterprises applying for loan shall meet the following hypotheses:

**Hypothesis 1:** let the fund owned by the enterprise be  $W$ , the amount needed for the investment project be  $I$ , let the enterprise require a loan of  $B$  ( $B = I - W$ ) from

banks, let the rate of return after project success be  $r'$ , and the profit be calculated by  $R$  ( $R = (B + W)(1 + r')$ ). Let the project success rate be  $p$ , let the rate of enterprise's default on repayments be  $s$ , let the loan interest rate be  $r$ , let the risk-free interest rate be  $\rho$ , let the value of the enterprise's mortgage to banks be  $C_0$  ( $C_0 < W$ ), and let the bank's examination cost be  $C_b$ .

**Hypothesis 2:** the enterprise applying for loan may make a decision of "applying based on facts" or "deceiving for loan" due to the moral hazard created by the bank-enterprise information asymmetry. Banks should make a decision of "approving the loan" or "rejecting the loan" by examining application materials.

**Hypothesis 3:** if the enterprise chooses to apply based on facts, but banks choose to reject the loan, banks and enterprises will suffer the respective opportunity losses of  $OL_b$  ( $OL_b = (r - \rho)B$ ) and  $OL_e$  ( $OL_e = (r' - r)B$ ). If the enterprise chooses to "deceive for loan", it must pay a disguise cost  $CL$  ( $CL < OL_e$ , which is deduced by the individual rationality of enterprise), and if banks examine the enterprise's attempt to "deceive for loan", then it will reject the loan and punish enterprise, in other words, make enterprises suffer a reputation loss  $RL$ , by increasing the loan interest rate and mortgage value in the next loan application.

It can be concluded from the above criteria that, if banks choose to approve the loan, then the individual rationality of bank needs to be satisfied.

$$p(1-s)(1+r)B + (1-p)s \cdot C_0 > (1+\rho)B \tag{1}$$

According the above analysis, the pay-off matrix of the bank-enterprise game is shown in Table 1:

TABLE 1 Pay-off matrix of the bank-enterprise game at the pre-loan Stage

		Enterprise	
		Applying Based on Facts	Deceiving for Loan
Bank	Approving the Loan	$p(1-s)(1+r)B + (1-p)s \cdot C_0 - (1+\rho)B - C_b$	$-(1+\rho)B + C_0 - C_b$
	Rejecting the Loan	$p(R - (1-s)(1+r)B) - s(1-p)C_0$	$p \cdot R - C_0 - CL$
		$-C_b - OL_b, -OL_e$	$-C_b, -CL - RL$

**2.2 BALANCE POINTS IN THE PROCESS OF EVOLUTIONARY GAME**

Let  $X$  be the ratio of the bank group's adoption of the strategy of "approving the loan", then  $1 - X$  is the ratio of its adoption of the strategy of "rejecting the loan"; let that  $Y$  be the ratio of the enterprise group's adoption of the strategy of "applying based on facts", then  $1 - Y$  is the ratio of its adoption of the strategy of "deceiving for loan"; and the fitness function is expressed by the expected profit.

When banks adopt the strategies of "approving the loan" or "rejecting the loan", the fitness functions and the average fitness functions  $U_{a1}, U_{a2}, \bar{U}_a$  are respectively:

$$\begin{aligned}
 U_{a1} &= Y[p(1-s)(1+r)B + (1-p)s \cdot C_0 - (1+\rho)B - C_b] + (1-Y)[-(1+\rho)B + C_0 - C_b] \\
 U_{a2} &= Y(-C_b - OL_b) + (1-Y)(-C_b) \\
 \bar{U}_a &= XU_{a1} + (1-X)U_{a2}
 \end{aligned} \tag{2}$$

Similarly, when enterprises apply based on fact or deceive for loan, the fitness functions and the average fitness functions  $U_{b1}, U_{b2}, \bar{U}_b$  are respectively:

$$\begin{aligned}
 U_{b1} &= X[p(R - (1-s)(1+r)B) - (1-p)s \cdot C_0] + \\
 & (1-X)(-OL_e) \\
 U_{b2} &= X(P \cdot R - C_0 - CL) + (1-X)(-CL - RL) \quad (3) \\
 \bar{U}_b &= YU_{b1} + (1-Y)U_{b2}
 \end{aligned}$$

According to the duplicator dynamic equation, the following two-dimensional differentiable dynamic system can be obtained:

$$\begin{cases}
 \frac{dX}{dt} = X(U_{a1} - \bar{U}_a) = X(1-X)[Y(p(1-s)(1+r)B - (1-(1-p)s)C_0 + OL_b) - (1+\rho)B + C_0] \\
 \frac{dY}{dt} = Y(U_{b1} - \bar{U}_b) = Y(1-Y)[X(-p(1-s)(1+r)B + (1-s(1-p))C_0 - RL + OL_e) + RL - OL_e + CL]
 \end{cases} \quad (4)$$

2.3 EVOLUTIONARY STABLE STRATEGY

Let  $\frac{dX}{dt} = \frac{dY}{dt} = 0$ , five balance points of the equation can be obtained, respectively, A(0,0), B(0,1), C(1,0), D(1,1) and E(X<sub>0</sub>,Y<sub>0</sub>), among which,  $X_0 = \frac{-RL - CL + OL_e}{-p(1-s)(1+r)B + (1-(1-p)C_0) - RL + OL_e}$ , and  $Y_0 = \frac{(1+\rho)B - C_0}{p(1-s)(1+r)B - (1-(1-p)s)C_0 + OL_b}$ ,  $X_0, Y_0 \in [0,1]$ .

According to the stability theory of nonlinear differential equation [16], the stability of balance points can be determined by the sign of the Jacobian matrix's characteristic root. By calculating the determinant *DetJ*, and trace *trJ*, of the Jacobian matrix *J*, of the differential dynamical systems (i.e., equation (4)), we can obtain:

$$J = \begin{bmatrix} (1-2X)[Y(p(1-s)(1+r)B + OL_b) - (1+\rho)B + C_0] & X(1-X)[p(1-s)(1+r)B - (1-(1-p)s)C_0 + OL_b] \\ Y(1-Y)[-p(1-s)(1+r)B + (1-s(1-p))C_0 - RL + OL_e] & (1-2Y)[X(-p(1-s)(1+r)B - RL + OL_e) + RL - OL_e + CL] \end{bmatrix} \quad (5)$$

$$\begin{aligned}
 DetJ &= (1-2X)[Y(p(1-s)(1+r)B - (1-(1-p)s)C_0 + OL_b) - (1+\rho)B + C_0] \\
 & \times (1-2Y)[X(-p(1-s)(1+r)B - RL + OL_e + (1-s(1-p))C_0) - OL_e + RL + CL] \\
 & - X(1-X)[p(1-s)(1+r)B + (1-p)s \cdot C_0 + OL_b] \\
 & \times Y(1-Y)[-p(1-s)(1+r)B + (1-s(1-p))C_0 - RL + OL_e] \quad (6)
 \end{aligned}$$

$$\begin{aligned}
 trJ &= (1-2X)[Y(p(1-s)(1+r)B - (1-(1-p)s)C_0 + OL_b) - (1+\rho)B + C_0] \\
 & + (1-2Y)[X(-p(1-s)(1+r)B - RL + OL_e + (1-s(1-p))C_0) - OL_e + RL + CL] \quad (7)
 \end{aligned}$$

By calculating the above-mentioned values and signs of *DetJ* and *trJ*, in each balance point, we can judge their stability situation and thus obtain the following bank-enterprise evolutionary stable strategies.

**Assumption 1:** when

$CL < p(1-s)(1+r)B - (1-s(1-p))C_0$ , i.e., when the enterprise's disguise cost of gaining a loan is smaller than the difference between the bank's expected profits and mortgage value, the bank will adopt different degrees of punishment, which may lead to the following two results:

Case 1: if  $RL > OL_e - CL$ , i.e., if the amount of penalty is greater than the difference between the enterprise's opportunity loss and the disguise cost, then, in the long run, it is impossible to achieve a stable and ideal partnership for banks and enterprises, no matter whether the local loan atmosphere is good or not.

Case 2: if  $RL < OL_e - CL$ , i.e., if the amount of penalty is smaller than the difference between the enterprise's opportunity loss and the disguise cost, then, after development and evolution, enterprises will choose to "deceive for loan" and the bank will choose to reject the loan application.

**Proof:** in Assumption 1, when the enterprise's disguise cost *CL* is less than the critical value  $p(1-s)(1+r)B - (1-s(1-p))C_0$ , makes the enterprise's profits from applying loan based on fact smaller than the profits from deceiving for loan. In the following section, the Jacobian determinant and trace are calculated, in order to obtain the stability of the balance points (see Table 2).

TABLE 2 Bank-enterprise Evolution Path (when the enterprise's disguise cost  $CL$  is smaller)

Size of Punishment $RL$	Balance Point	Sign of $DetJ$	Sign of $trJ$	Local Stability	Phase Diagram
Case 1: $RL > OL_e - CL$ Note: there are five balance points and $X_0, Y_0 \in [0,1]$	A(0,0)	-	$\pm$	Saddle point	
	B(0,1)	-	$\pm$	Saddle point	
	C(1,0)	-	$\pm$	Saddle point	
	D(1,1)	-	$\pm$	Saddle point	
	E( $X_0, Y_0$ )	+	0	Centre	
Case 2: $RL < OL_e - CL$ Note: there are four balance points and $X_0 \notin [0,1]$	z	+	-	ESS	
	B(0,1)	+	+	Unstable point	
	C(1,0)	-	$\pm$	Saddle point	
	D(1,1)	-	$\pm$	Saddle point	

FIGURE 1 Phase Diagram in case 1

FIGURE 2 Phase Diagram in case 2

From case 1 shown in Table 2 we can see that when  $RL > OL_e - CL$ , it is impossible to achieve a stable and ideal partnership for banks and enterprises, no matter what the initial ratios of strategies adopted by bank/enterprise groups are. In the process of this evolutionary game, enterprises have a small disguise cost, if banks choose to approve the loan at the very beginning, enterprises tend to choose to “deceive for loan”, then, the bank group tends to increasingly reject loan applications. Gradually, as larger sanctions are introduced, the enterprise's rationality tends towards the strategy of applying based on fact. Therefore, the model has no ESS in repeated cycles (see Figure 1 for evolution path).

From case 2 in Table 2, we can deduce that when  $RL < OL_e - CL$  (critical value), both banks and enterprises in the game tend towards the stable point A(0,0), i.e., the evolution result of bank-enterprise decision is to choose the strategies of “rejecting the loan” and “deceiving for loan” (see Figure 2 for evolution path).

**Assumption 2:** when  $CL > p(1-s)(1+r)B - (1-s(1-p))C_0$ , i.e., the enterprise's disguise cost is bigger than the difference between the bank's expected profits and mortgage value, banks will adopt different degrees of punishment, which may lead to the following two different results:

Case 1: if  $RL > OL_e - CL$ , i.e., when the amount of penalty is bigger than the difference between enterprise's opportunity loss and the disguise cost, then, after development and evolution, banks will choose to approve loan applications and enterprises will choose to apply for loans based on fact.

Case 2: if  $RL < OL_e - CL$ , i.e., when the amount of penalty is smaller than the difference between the enterprise's opportunity loss and the disguise cost, then, the evolution result depends on the initial loan atmosphere. If the ratio of cheating by enterprises and rejection by banks is large, the bank-enterprise evolution result will choose to reject the loan and “deceive for loan”; if the initial ratio of “honest” enterprise and “approving” bank is high, the evolution result will be that banks approve loan applications and enterprises apply for loans based on fact.

**Proof:** in Assumption 2, when the enterprise's disguise cost,  $CL$ , is less than the critical value  $p(1-s)(1+r)B - (1-s(1-p))C_0$ , the enterprise's expected profits from applying based on fact are bigger than which from deceiving for loan. Similarly, the aforementioned Jacobian determinant and trace can be calculated, to obtain the stability of the balance point, which is shown in Table 3.

TABLE 3 Bank-enterprise Evolution Path (when the enterprise's disguise cost  $CL$  is bigger)

Size of Punishment $RL$	Balance Point	Sign of $DetJ$	Sign of $trJ$	Local Stability	Phase Diagram
Case 1: $RL > OL_e - CL$ Note: there are five balance points and $X_0, Y_0 \in [0, 1]$	A(0,0)	-	$\pm$	Saddle point	<p>FIGURE 3 Phase Diagram in case 1</p>
	B(0,1)	-	$\pm$	Saddle point	
	C(1,0)	+	+	Unstable point	
	D(1,1)	+	-	ESS	
Case 2: $RL < OL_e - CL$ Note: there are four balance points and $X_0 \notin [0, 1]$	A(0,0)	+	-	ESS	<p>FIGURE 4 Phase Diagram in case 2</p>
	B(0,1)	+	+	Unstable point	
	C(1,0)	+	+	Unstable point	
	D(1,1)	+	-	ESS	
	E( $X_0, Y_0$ )	-	0	Saddle point	

From case 1 in Table 3 we can see that when  $RL > OL_e - CL$  (critical value) banks and enterprises in the game tend towards the stable point D(1,1) (see Figure 3 for evolution path) i.e., the evolution result of bank-enterprise decision is to choose the strategies of “approving the loan” and “applying based on facts” respectively.

From case 2 in Table 3 we can see that when  $RL < OL_e - CL$  (critical value), two players of the game have two local stable points (see Figure 4 for evolution path) and the evolution result depends on the value of the initial ratio of the strategies adopted by bank group and enterprise group. If the initial value is close to A(0,0) i.e., the initial loan atmosphere is enterprises tend to deceive and banks tend to reject the result of decision evolution is to choose strategies of “rejecting the loan” and “deceiving for loan”; if the initial value is close to D(1,1) i.e., the initial atmosphere is that enterprises tend to be honest and banks tend to approve the loan, the result of evolution is to choose the strategies of “approving the loan” and “applying based on facts” respectively.

### 3 Numerical Simulation of Bank-Enterprise Evolutionary Game Model

To display the bank-enterprise evolutionary game more intuitively, we make a numerical simulation for the previous model for further discussion.

Let  $B=1$  million the mortgage value  $C_0=0.8$  million the loan interest rate  $r=7\%$ ; the bank's safe investment interest rate  $\rho=2.5\%$  the success rate of the enterprise's project  $p=0.9$  the rate of default rate  $s=0.15$ .

1) Let parameters  $CL, RL$  meet conditions of Assumption 1.

Disguise cost loss  $CL=0.02$  million  $RL$  amount of penalty is 0.12 million (in accordance with case 1) and 0.05 million (in accordance with case 2). Figures 5 and 6 show the phase diagrams of the system in both cases. Figure 5 shows five balance points and the solutions around the central point form a closed trajectory without any stable point. Figure 6 shows only four balance points and the trend of solution tends towards point (0, 0), indicating the existence of stable point (0, 0).

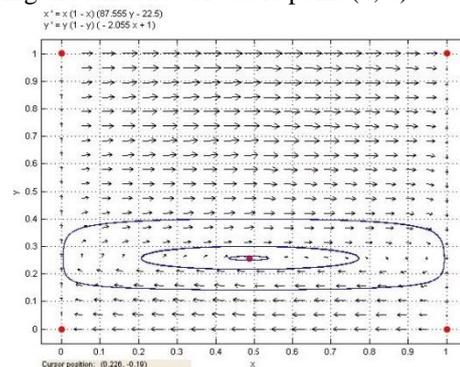


FIGURE 5 Phase Diagram in case 1

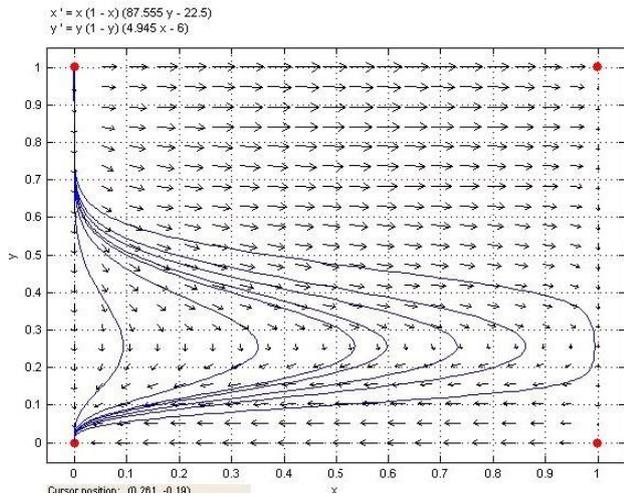


FIGURE 6 Phase Diagram in case 2

Let the initial values of  $(X, Y)$  as  $(0.8, 0.3)$  and  $(0.2, 0.1)$ ; and  $(0.1, 0.2)$  and  $(0.2, 0.7)$  respectively and make simulations for case 1 and case 2. Results show that when the disguise cost  $CL$  is relatively small and the penalty  $RL$  is relatively large, the game players have no stable point and the rate of bank-enterprise decision-making behaviours displays periodic fluctuations (see Figure 7); when both  $CL$  and  $RL$  are small the decisions tend towards the point  $(0, 0)$  (Figure 8) i.e., the evolution result is that the respective players either choose to reject the loan and deceive for loan.

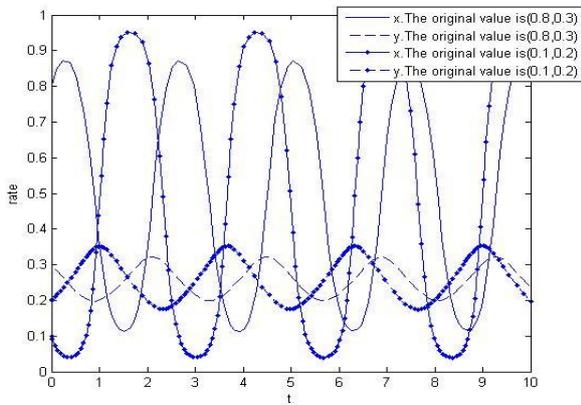


FIGURE 7 Simulation result in case 1

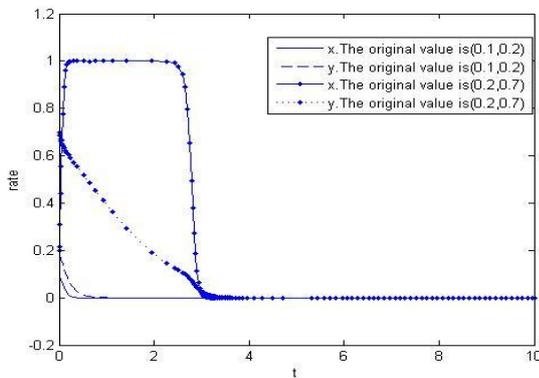


FIGURE 8 Simulation result in case 2

2) Let the parameters  $CL, RL$  meet the conditions of Assumption 2.

Let the disguise cost  $CL=0.05$  million and let  $RL$ , the bank's penalty for deceiving equal 0.15 million (meet case 1) and 0.05 million (in accordance with case 2). Figures 9 and 10 shows the phase diagrams of the system in both cases. Figure 9 shows four balance points and the direction of solutions points to point  $(0, 0)$ . Figure 10 shows that the five balance points and solutions tend towards two points, those points being  $(1, 1)$  and  $(0, 0)$ .

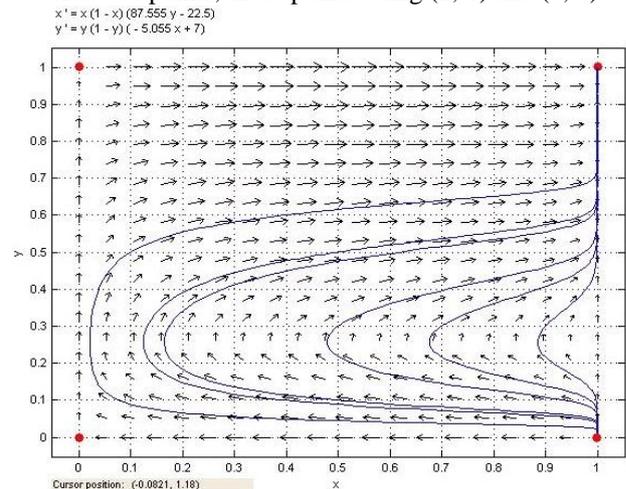


FIGURE 9 Phase diagram in case 1

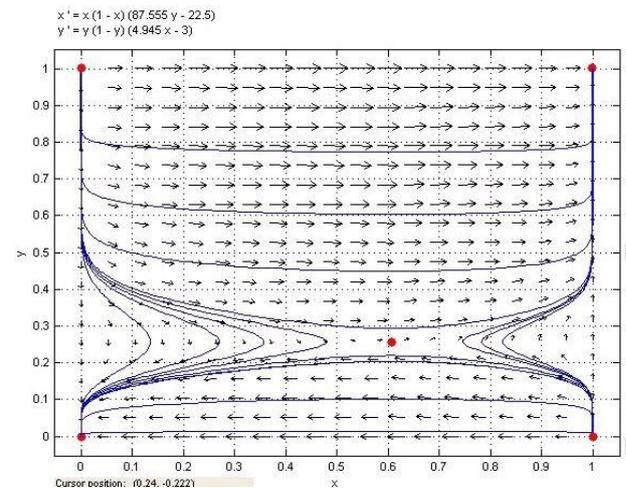


FIGURE 10 Phase diagram in case 2

Let the original values of  $(X, Y)$  for each respective case as  $(0.1, 0.2)$  and  $(0.7, 0.1)$  and  $(0.1, 0.2)$  and  $(0.8, 0.9)$  and make simulations for cases 1 and 2. The results showed that when the disguise cost  $CL$  is increased, the punishment  $RL$  increases two players of game tend towards the point  $(1,1)$  (Figure 11) and the result of the evolution is that banks and enterprise choose to approve the loan and apply for the loan based on fact; when the amount of penalty  $RL$  is small different original values lead to tremendously different evolution results. When the original values of  $X$  and  $Y$  are  $(0.1, 0.2)$  evolution result is that the business and enterprise respectively choose to reject the loan and "deceive for loan"; when the original values of  $X$  and  $Y$  are  $(0.8, 0.9)$  the evolution

result is that the business and enterprise respectively choose to approve the loan application and apply for a loan based on fact (Figure 12).

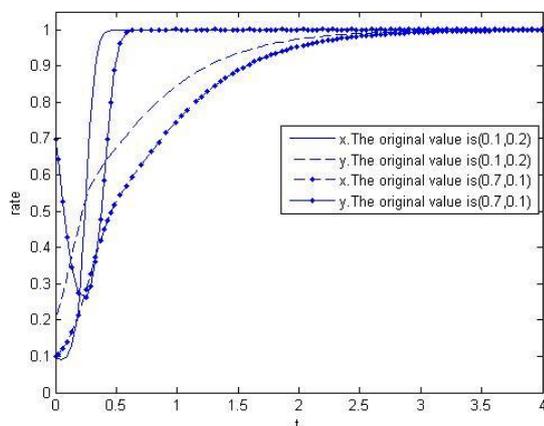


FIGURE 11 Simulation result in case 1

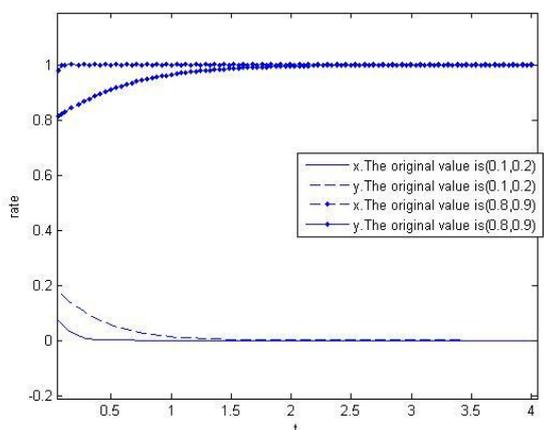


FIGURE 12 Simulation result in case 2

#### 4 Conclusions and Enlightenment

This paper established an evolutionary game model of bank-enterprise at the pre-loan stage and made a numerical simulation. According to the analysis results, the enterprise's disguise cost and the bank's punishments on cheating enterprise play an instrumental role in preventing moral hazard.

Firstly, when the enterprise's disguise cost is so low that the enterprise's profits from deceiving for loan are higher than those from applying based on fact, banks and

enterprises will not reach an ideal state in the long-term evolutionary process, even if banks impose large sanctions for enterprises. Furthermore, when the degree of punishments is very low, banks and enterprises may evolve into a state in which cooperation cannot be achieved. Therefore, banks should increase the enterprise's disguised costs using a variety of means, such as examining application materials more rigorously, enhancing the screening ability of commercial banks, improving industrial standards, and enhancing the professional ethics of banking personnel, in order to avoid the opportunities caused by banking personnel's corruptive behaviours.

Secondly, when the enterprise's disguise cost is so large that the enterprise's monetary gains from deceiving for loan are lower than those from applying based on fact, banks and enterprises can reach a stable cooperation state only if banks make appropriate sanction mechanisms according to the local loan atmosphere. If the initial loan atmosphere is favourable, i.e., the proportions of cheating enterprises and banks with stint loans are small, banks only need to focus on increasing the enterprise's disguise cost. If the initial loan atmosphere is unfavourable, i.e., the proportions of cheating enterprises and banks with stint loans are high, a proper punishment amount (at least larger than the difference between the enterprise's opportunity loss and disguise cost) needs to be set, to ensure that evolution can reach a stable state enterprises apply based on fact and banks approve the loan application.

Therefore, while increasing enterprises' disguised costs, banks should design appropriate sanction measures for enterprises according to the local loan atmosphere, enterprise asset appraisals and the profits of the loan investment project. Additionally, banks should add cheating enterprises to a blacklist, and impose restrictions on or make other sanctions such as higher interest for the enterprises' future loans, improvements and upgrades in business credit reference systems and the sharing of blacklists between various commercial banks.

#### Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant 71171135.

#### References

- [1] Stiglitz J, Weiss A 1981 Credit rationing in markets with imperfect information *American Economics Review* 71(3) 393-410
- [2] Bester H 1985 Screening vs. Rationing in credit markets with imperfect information *American Economics Review* 75(4) 850-5
- [3] Blumberg B, Letterie W 2008 Business starters and credit rationing *Small Business Economics* 30(2) 187-200
- [4] Dong Xiaohong, Wang Huanchen, Wu Jin 1998 Credit Decision-making Mechanism Evading Moral Hazards *Journal of Shanghai Jiaotong University* 32(3) 1-5 (in Chinese)
- [5] Li Yanxi 1999 Incentive Model Preventing Moral Hazards in Bank Loan Risk Management *Journal of Management Science* 2(2) 62-5 (in Chinese)
- [6] Zhuang X, Huan X 2002 An Analysis on Bank Loan Pricing Strategies Based on Information Asymmetry *System Engineering* 20(3) 20-3 (in Chinese)
- [7] Pang S, Li R 2005 An Analysis on Credit Risk Decision Contract Model Avoiding Moral hazard *System Engineering Theory and Practice* 25(6) 114-20 (in Chinese)
- [8] Pang S 2007 Credit Risk Decision-making Mechanism with the Rate of Default of Repayment *System Engineering Theory and Practice* 27(10) 31-9 (in Chinese)

[9] Taylor P D, Jonker L B 1978 Evolutionary stable strategies and game dynamics *Mathematical Biosciences* 40(1/2) 145-56

[10] Luo J 2012 Inter-enterprise Cooperation Default Punishing Mechanism Based on Evolutionary Game Theory *System Engineering* 30(1) 27-31 (in Chinese)

[11] Tian Z, Sun Q 2012 Evolutionary Game of Competition and Cooperation Behaviors in Industrial Cluster under Agglomeration Economics—Duplicator Dynamic Analysis Based on ESS Strategy *Scientific and Technological Progress and Countermeasures* 29(6) 52-6 (in Chinese)

[12] Lee C K 2007 Analysis on the evolutionary game of innovative financial system *Journal of American Academy of Business* 5(1) 338-43

[13] Guth W, Menger F, Ockenfels A 2007 An evolutionary analysis of buyer insurance and seller reputation in online markets *Theory and decision* 63 265-82

[14] Smith J M 1982 *Evolution and the theory of games* Oxford University Press: Oxford

[15] Weibull J W 1995 *Evolutionary Game theory* The MIT Press: Massachusetts

[16] Ma Z, Zhou Y 2012 *Ordinary Differential Equation Qualitative and Stability Analysis* Science Press: Beijing (in Chinese)

<b>Authors</b>	
	<p><b>Qihong Zheng, born in September, 1981, Ningbo, Zhejiang, China</b></p> <p><b>Current position, grades:</b> Doctoral candidate of Business School the University of Shanghai for Science and Technology.  <b>University studies:</b> College of Computer Science and Information Technology, Zhejiang Wanli University.  <b>Scientific interest:</b> Management of financial risk.  <b>Publications:</b> More than 12 papers published in various journals.  <b>Experience:</b> Teaching experience of 10 years; 3 scientific research projects.</p>
	<p><b>Liangrong Song, born in June, 1966, Shanghai, China</b></p> <p><b>Current position, grades:</b> Professor, doctoral supervisor, head of Department of accounting, head of Small and medium-sized banking research center the University of Shanghai for Science and Technology.  <b>University studies:</b> Graduated from University of Shanghai for Science and Technology in 2002 a doctor's degree in management.  <b>Scientific interest:</b> Management of financial risk.  <b>Publications:</b> More than 130 papers published in various journals and more than 10 teaching materials.  <b>Experience:</b> more than 25 scientific research projects.</p>

# Study on the city planning for geological disasters defence based on the model of safe city planning

Yuhui Xu<sup>1, 2</sup>, Qiuyue Luo<sup>1, 2\*</sup>

<sup>1</sup>Key Laboratory of New Technology for Construction of Cities in Mountain Area, Chongqing University, Chongqing, China, 40030

<sup>2</sup>Faculty of Architecture and Urban Planning, Chongqing University, Chongqing, China, 40030

Received 12 February 2014, www.tsi.lv

## Abstract

It is difficult to take geological disasters defence into city planning, which can greatly prevent lands for construction from geological disasters and ensure the rationality, safety and high-efficiency of land-use. Based on the model of safe city planning and discussing from the angle of safe land arrangement, this paper proposes the system of the city planning for geological disasters defence. It focuses on the defence of mutant geological disasters and takes the geological hazards assessment as the foundation. It makes the detailed regulations of the “specific control and management”, which guarantees both the defence and the control of disasters from technology aspect. Besides, it puts forward methods with highly couple of different measures to defend different disasters in city planning system which can enhance the operability to answer the reality.

*Keywords:* city planning for geological disasters defence, safe city planning, mutant geological disasters, specific control and management, city planning system

## 1 Introduction

### 1.1 THE CONCEPTION OF SAFE CITY

The conception and research emphasis of “safe city” are different in countries all around the world. In Occident, because of the social issue, “safe city” is usually connected with prevention of crime. While “safe city” in Japan attracts more emphasis on defending natural disasters like earthquake, flood, storm etc [1]. In China, most related researches of “safe city” focus on disasters prevention and mitigation in cities. As a future development pattern of cities, like the eco city and the healthy city, the safe city attaches importance to urban emergency, share of security resources and disasters prevention. Meanwhile, it also emphasizes the urban ability of sustaining the daily safety condition [2-4].

### 1.2 THE CONCEPTION AND MODEL OF SAFE CITY PLANNING

The conception of safety in city planning should be based on the foundation of urban public safety<sup>3</sup>. Therefore, the safe city planning needs the construction with an aggregative model which integrates safe decision-making, safe land arrangement, safe facilities planning and safe policies and regulations on technology level (Figure 1).

The safe decision-making means aims and strategic deployment about prevention and mitigation of disasters. It also needs to construct a wholesome and effective

urban safe system to guarantee the safety of cities by the all-around and multilevel mean [5].



FIGURE 1 The model of the safe city planning

The safe land arrangement means making the safe decision-making practicable in urban space by using technological means of city planning. It includes the safety of urban lands, evacuation ability of roads and the planning of emergency refuge space [6].

The safe facility planning covers planning of facility for disasters defence, salvation and rebuilding [7]. It is an important infrastructure which guarantees cities' regular operation in the periods of preventing, avoiding and replying disasters.

The safe policies and regulations which have rigid restrain in city planning system [8] guarantee the implement of the planning achievement to be legal provisions.

\* Corresponding author e-mail: 18875039071@139.com

**2 Progress of study on the geological disasters defence in city planning**

On one hand, the geological disaster owns features of strong destructive power, concatenate diffusivity, hard reconstruction and long recovery [9]. Thus city planning usually takes defence as the principle to avoid the hazardous area to the greatest extent when choose lands for construction. On the other hand, in China, due to the large population and limited land, the lands threatened by geological disasters can be used after the hazard control if conditions permitted. In worldwide, there exist some advances in the research of prediction, investigation and emergency measures of geological disasters. Some are researches on individual cases and some universal methods and applications have been proposed. More and more researchers have realized the intimate relationship between geological disasters defence and land utilization

in city planning and carried out some researches [10, 11]. However, researches on the connection of geological disasters defence and city planning is insufficient. In spite that some regulations like “Regulation on Geologic Disasters Defence” have already been promulgated in China, there still exist some deficiencies in the application of city, land and traffic planning [12].

**3 The frame of city planning for geological disasters defence**

City planning should pay sufficient attention to geological disasters defence. However, there exists blank in research on the combination between them. While as one kind of special public staff, city safety is lack of economic benefit and less intervened by the market organization<sup>4</sup>. The consideration coming from the system of city planning is especially important (Figure 2).

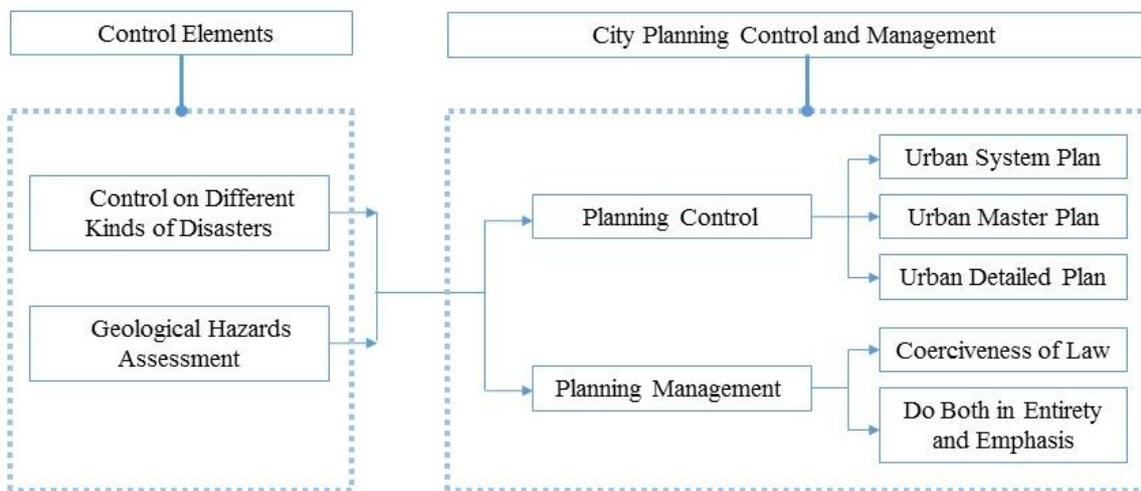


FIGURE 2 Frame of city planning for geological disasters defence

**4 The mathematical model and content of city planning for geological disasters defence**

**4.1 CONGRUITY AND CLASSIFY OF DISASTERS**

The geological disaster in this paper refers to the geological-related disasters caused by natural factors or human activities that endanger people's lives and property. It includes rock fall, landslide, debris flow,

ground collapse, ground fissure and land subsidence [13]. They are divided into two categories: mutant and gradual geological disasters. As mutant geological disasters has kept a quite high proportion in quantity statistics of geological disasters from 2006 to 2013, China (usually up to 97%) (Figure 3), they are emphasized in this paper (Table 1).

TABLE 1 Category of mutant and gradual geological disasters

Category of Disasters	Definition	Disasters Including	Measures
<b>Mutant</b>	disasters that happen in sudden and finish in short time	rock fall landslide debris flow	emphasis; especially in data of avoiding
<b>Gradual</b>	disasters that happen and develop slowly, and develop over time	ground collapse ground fissure land subsidence	not emphasis; proposing related code and guide

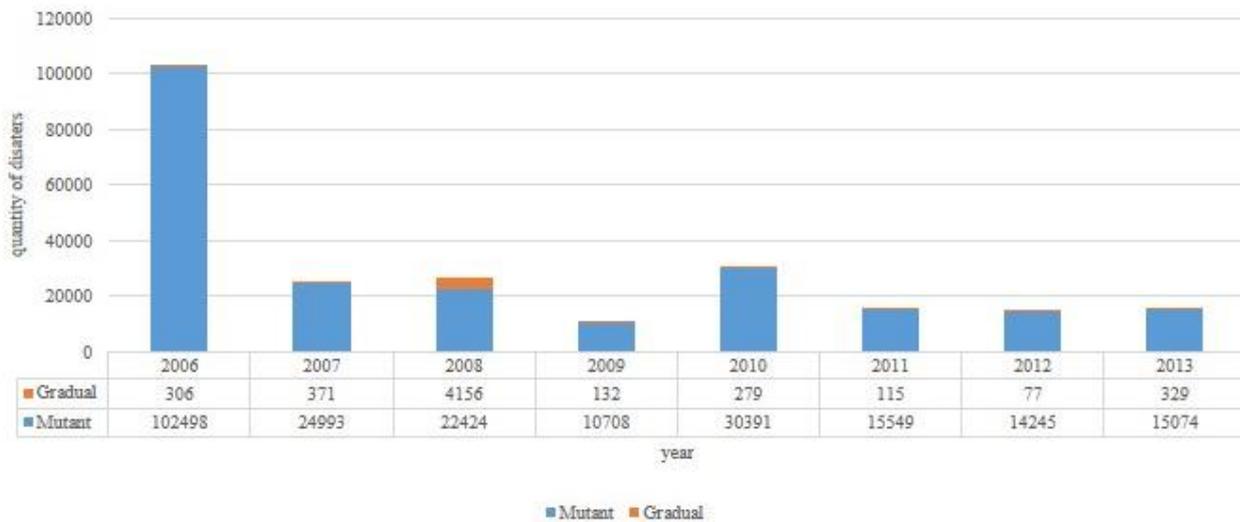


FIGURE 3 Classification chart of geological disasters in china from 2006 to 2013 [14]

Here are equations to calculate damage area of three mutant geological disasters [15].

$$L_{max} = \frac{V_{max}^2 \cos^2 \partial}{2g(f - \sin \theta)} + d, \tag{1}$$

where

$$V_{max} = \sqrt{2g(H - f \cdot L) + V_0^2}. \tag{2}$$

The damage area of landslide can be ascertained by the maximum distance of movement ( $L_{max}$ ). It can be calculated by Equation (1), where  $V_{max}$  is the maximum slip velocity;  $\partial$  is the angle between the maximum sliding direction and the bottom slip layer;  $g$  is the gravitational acceleration;  $f$  is the dynamic friction coefficient of the bottom slip layer;  $\theta$  is the gradient of the bottom slip layer;  $d$  is the distance from centre of gravity to the front of the debris.  $V_{max}$  can be calculated by Equation (2), where  $H$  is the level difference of bottom slip layer;  $L$  is the horizontal distance of the bottom slip layer;  $V_0$  is the initial velocity of the slip mass.

$$S = \frac{2}{3} L \cdot B - \frac{1}{12} B^2 ctg \frac{1}{2} R, \tag{3}$$

where

$$L = 0.7523 + 0.0060A + 0.1261H + 0.0607D - 0.0192G, \tag{4}$$

where

$$B = 0.2331 - 0.0091A + 0.1960H + 0.0983D + 0.0048G \tag{5}$$

and

$$R = 47.8296 + 8.8876H - 1.3085D. \tag{6}$$

The damage area of debris flow can be ascertained by the proportion, the maximum length and the maximum breadth of the deposition fan of debris ( $S$ ,  $L$ ,  $B$  and  $R$  respectively) by Equation (3). Besides, in Equations (4-6),  $A$  is the drainage area;  $H$  is the relative height of it;  $D$  is the length of the main gully;  $G$  is the average gradient of it.

$$L = \eta \frac{QS_0}{M} t. \tag{7}$$

The damage area of ground collapse of ground collapse can be ascertained by the radius of subsidence ( $L$ ). It can be calculated by Equation (7), where  $\eta$  is the coefficient related to the soil properties;  $Q$  is the exploitation quantity of groundwater;  $S_0$  is the effective drawdown under surface of basement rock;  $M$  is the overburden thickness;  $t$  is the duration of pumping.

#### 4.2 PURPORT AND CONTENT

##### 4.2.1 Different solutions to mutant and gradual geological disasters respectively

1) The layout of lands in cities and towns must avoid areas influenced by mutant geological disasters. For areas influenced by gradual geological disasters, comprehensive controls should be taken when social economic and technological conditions permit.

2) Divide dangerousness of mutant geological disasters into 4 grades: low, middle, high and sky-high. This needs achievement of geological survey as important basis.

3) Due to the individual differences in geological disasters, related detail regulations should be established in the master planning and detailed planning respectively.

##### 4.2.2 The geological hazards assessment

The geological hazards assessment is an important basis

of city planning for geological disaster defence and it plays a significant part in each stages of city planning. See Figures 4 and 5.

The geological hazards assessment is influenced by the predictive probability index [16].

$$Y = 0.62D + 0.38R \tag{8}$$

The predictive probability index of geological disasters ( $Y$ ) can be calculated by Equation (8), where  $D$  is complexity index of geological environment and  $R$  is index of the precipitation.

### 4.3 DETAILED REGULATIONS OF MANAGEMENT AND CONTROL IN CITY PLANNING SYSTEM

#### 4.3.1 Detailed regulations of control in urban system planning

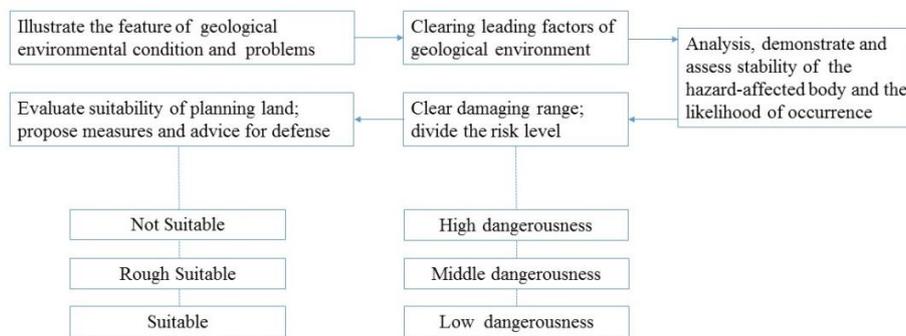


FIGURE 4 The process of geological hazards assessment

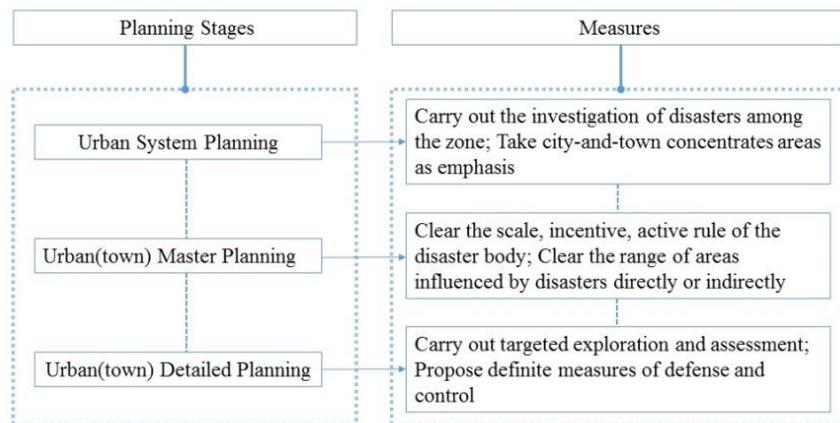


FIGURE 5 Key points of geological hazards assessment in different planning stage

#### 4.3.2 Detailed regulations of control in city (town) master planning

1) Based on the geological hazards assessment, we should evaluate the suitability of the urban (town) lands for construction, and propose planning measures for geological disasters defence on urban land choosing, layout of function, important infrastructure planning.

1) Put avoidance and defence as the main purpose. Predominating direction of city development, the layout of infrastructures and social facilities must avoid areas with high hazard, and choose areas for them with low or no geological disasters in prior.

2) The site selection of important projects should avoid areas with high dangerous ones, and avoid the possibility of geological disasters induced by the project construction.

3) In areas with high dangerous ones, the development scale of existing city or town must be controlled and it is not suitable for expanding the scale of land.

4) The existing city, town and housing estate in areas with sky-high dangerous mutant geological disasters should be removed when conditions permits.

2) Areas harmed by mutant geological disasters directly must be ranked as forbidding constructing areas.

3) The lifeline engineering should avoid high hazardous areas. Or effective control measures must be taken.

4) The detailed regulations of six kinds of geological disasters defence are shown in Table 2.

TABLE 2 Detailed regulations of geological disasters defence in master planning

Type of Disasters	Areas with the Disaster	Body be Formalized	Detailed Regulations
<b>Mutant</b>	1) Rock Fall	cities with high or middle possibility of disasters	compile the special planning
		high hazardous areas	clear the disaster body and condition of threaten construction forbidden
		areas harmed and influenced by disasters assessed as low disaster incident areas after governing	can be used as green space or square
	2) Landslide	lands on upward and downward side of rock bodies	reserve protective safe distance in advance
		areas influenced by disasters	manage and control in zonal
		areas with incident disasters	restrict development intensity, be suitable for green space and forestland
3) Debris Flow	urban lands for construction	safe avoiding distance (20~800m)	
	high hazardous areas	control the scale and density of cities and towns	
	existing cities(towns) in high hazardous areas	control the scale and form; remove housing estates and important architectures in direct hazardous areas	
<b>Gradual</b>	4) Ground collapse	route of regional infrastructure	avoid or decrease the possibility of inducing disaster
		channel of disasters and their influencing scopes	clear the mode of avoiding, arrange lands for construction in safe areas
	1) Ground fissure	infrastructure construction	avoid channel of the mud, reserve space for drainage in advance
		the choose of urban constructive lands	advise its suitability degree based on the distance from the boundary of disasters
2) Land Subsidence	lands choosing for new constructive areas	control the scale, avoid high hazardous areas, step back a distance for 100~200m	
	existing districts in high hazardous areas	restrict high-rise and high dense building groups	
		main traffic facilities	arrange in middle and low hazardous areas; need preventive measures when through high hazardous areas
		high hazardous areas	suitable for arranging green space and forestland; consider self-restraint and recovery of underground water

TABLE 3 Detailed regulations of geological disasters defence in detailed planning

Type of Disasters	Areas with the Disaster	Body be Formalized	Detailed Regulations
<b>Mutant</b>	1) Rock Fall	buildings (structures)	avoiding areas threatened with rocks; reserve 20~30m for defence in advance
		dangerous rock body	no buildings(structures) on areas influenced by disasters; green space and forestland can be ranged under it
		dangerous rock body that can be governed	carry out project and biological measures; then can be land for construction after passing the assessment
	2) Landslide	planning areas with landslide body	Constructions and project lifelines avoid the scope and influencing area
		areas must be used for construction	carry out reasonable governs to the safe standard
	3)Debris Flow	constructions in city	stay away from the scope; clear the distance for avoiding according to the specific condition
measures of defence and governing		stay back from it for 5% to 20% of its width according to the type of disasters	
<b>Gradual</b>	4) Ground Collapse	existing buildings	housing estate and important buildings should be removed; strengthen the power of buildings resistance to disasters
		constructions in city	stay back from the area at least 500m
	the area with incident disasters	forbid planning land for construction except green areas	
	areas influenced by mined deformation collapse	control the development intensity	
1) Ground Fissure	2) Land Subsidence	main traffic facilities and important projects	stay back from the incident areas above 500m
		existing housing estate in high hazardous areas	remove and avoid
	areas with karst collapse	avoid or govern it until reach the standard of assessment; mainly buildings stay back from it above 150m	
	constructive lands with disasters	stay away from the disaster for 6~40m according the Importance of lands or buildings	
		municipal engineering pipelines	avoid spanning the ground fissure; govern it and check it at regular interval if it can't be avoided.
		high hazardous areas	need to have assessment; control its development capacity
		important traffic facilities	stay back from high hazardous areas and areas influenced by it for 200~500m.
		important projects and projects, which have 7m above excavation of foundation pit	must consider the influence to land subsidence; propose governing measures to uneven subsidence
		projects with deep excavation by using drainage in high areas	clear the governing measures according to the risk assessment report

4.3.3 Detailed regulations of control in city (town) detailed planning

1) In the regulatory plan, measures of prevention and control should be taken in aspects of lands layout, constructive intensity, traffic, infrastructure planning etc.

While in the site plan, measures should be taken in aspects of architectural layout, ground vertical, municipal project pipeline.

2) The planning lands that located in the area with dense constructions and assessed as high hazardous ones,

can be used reasonably after disasters control and reaching the level of middle or low.

3) In the regulatory plan, the site and predicted scope of hazards must be labelled in planning map clearly. And planning control guideline and advice for defence or governing should be definitely proposed.

4) The detailed regulations of six kinds of geological disasters defence are shown in Table 3.

#### 4.3.4 Detailed regulations of management in city planning system

1) Consummate the system of laws and regulations to safeguard the requirement of coerciveness in city planning system.

2) According to the degree of geological disasters comprehensive hazards and different percentages of the mutant geological disaster in planning areas, clear the different requirements in specialized planning about the geological disaster defence in the planning achievement.

## References

- [1] Hanqing Zhang, Shenzhi Dai 2005 Literature review of safety research in urban planning *Urban Planning Forum* 156(2) 38-44 (In Chinese)
- [2] Guo Zaifu 2012 Study on connotation of safe city and its continual improvement process *Journal of Safety Science and Technology* 8(12) 53-7 (In Chinese)
- [3] Zhongmin Jin 2011 Safe city concept based mega-city disaster prevention *Planners* 27(8) 10-3 (In Chinese)
- [4] Zhang Hanqing 2011 Theoretical framework for safe city planning *Planners* 27(8) 5-9 (In Chinese)
- [5] Dai Shenzhi 2002 Urban safety strategy and system *Planners* 18(1) 9-11 (In Chinese)
- [6] Wei Wang 2007 *Study on disaster-prevention space planning of urban and its experience* Doctor Degree Thesis of Central South University: Guangzhou (In Chinese)
- [7] Hu Zhiliang, Gao Xiangduo 2012 The Construction and Planning Application of City Public Security Infrastructure under the Comprehensive Disaster Prevention Concept *Areal Research and Development* 31(2) 49-53 (In Chinese)
- [8] Lv Yuan 2004 *Study on the planning strategy of urban disaster-prevention space system* Master Degree Thesis of Beijing University of Technology Beijing (In Chinese)
- [9] Zhang Fenglin 2011 *Research on the prevention of urban geological disasters of urban planning* Master Degree Thesis of Lanzhou University Lanzhou (In Chinese)

## 5 Conclusions

In this research, we propose the frame of city planning for geological disasters defence based on the model of safe city planning. The frame owns two important control elements. 1) Take control by different kinds of geological disasters and the mutant geological disasters are emphasized. 2) Take the geological hazards assessment as the important basis. In order to guarantee the effective implement of city planning for geological disasters defence in practice, further we makes the detailed regulations of the "specific control and management", which aims at different geological disasters in the city planning system respectively.

## Acknowledgement

The National Science and Technology Support Program, the study on the key techniques of high density space's efficiency optimization in City cluster (2012BAJ15B03)

- [10] Zhang Wangfeng, Zhang Fenglin 2011 Introspection of urban geological disasters defence from debris flow in Zhou Qu *Gansu Science and Technology* 27(10) 53-6 (In Chinese)
- [11] George D B, Kalliopi G P, Harikia D S, Dimitrios P, Konstantinos G C 2012 Potential suitability for urban planning and industry development using natural hazard maps and geological-geomorphological parameters *Environmental Earth Sciences* 66(2) 537-48
- [12] Wang Yugang 2011 Urban planning and geological disasters *Overview of Disaster Prevention* 1 40-3 (In Chinese)
- [13] The Central People's Government of the P.R China 2003 *Regulation on the prevention and control of geologic disasters* [http://www.gov.cn/gongbao/content/2004/content\\_63064.htm/](http://www.gov.cn/gongbao/content/2004/content_63064.htm/) 1 January 2014
- [14] China Geological environmental Monitoring Institute P.R. China 2014 *A bulletin of geological disasters in China* <http://www.cigem.gov.cn/auto/db/explorer.html?db=1006&type=1&fd=16&fv=49&uni=0&md=15&pd=210&mdd=11&pdd=5&msd=11&psd=5&start=0&count=20/> 1 October 2013
- [15] Shipiao Chang, Sumin Zhang 2007 *Project Geology Handbook* China Architecture & Building Press Beijing
- [16] Ministry of Land Resources P R China 2008 *Technical Code for Assessment of Geological Hazard* <http://wenku.baidu.com/view/52b9d602e87101f69f319507.html/> 10 January 2014

## Authors



**Yuhui Xu, born in August, 1971, Yunnan Province, China**

**Current position:** Professor, PhD Supervisor, Head of Urban Planning Department and Key Laboratory of New Technology for Construction of Cities in Mountainous Area, Chongqing University, China, Department Head Faculty of Architecture and Urban Planning, Chongqing University, China  
**University studies:** D.Eng. at Chongqing University in China.

**Scientific interest:** Urban planning theories of mountainous cities and residential development.

**Publications:** 37 papers published in various journals. 1 teaching material and 5 books.

**Experience:** Teaching experience of 14 years. Researcher in 1 National Science And Technology Support Program, 2 programs of Natural Science Fund Subject and 1 program of Urban Sustainable Development.



**Qiuyue Luo, born in August, 1988, Sichuan Province, China**

**Current position:** Master Student in Grade 2 in Faculty of Architecture and Urban Planning, Chongqing University, China

**University studies:** B.Eng. in City Planning from Xihua University in China

**Scientific interest:** Urban planning theories of mountainous cities and residential development.

**Experience:** Researcher in 1 National Science And Technology Support Program.

# A forecasting model of the economic efficiency of data centre construction project

V Grekul\*, L Baydalina

National Research University Higher School of Economics, 20 Myasnitskaya Ulitsa, Moscow, 101000, Russia

Received 1 June 2014, www.tsi.lv

## Abstract

Today, a data centre is the main instrument for providing flexible, scalable IT services to business on the basis of distributed or cloud computing technologies. Building a data centre is always expensive and resource-intensive project. Therefore, during the development of the concept of this project it is extremely important to estimate accurately its economic efficiency. This article represents the model for the analysis of the effectiveness of investment in the data centre construction. Our model comprises several regressions that show correlations between main characteristics of the project (capital and operational expenditures, Net Present Value) and parameters of the data centre under construction (data centre area and the number of racks). The model is based on the results of the analysis of the current state and trends in the data centre market in Russia.

*Keywords:* datacentre, price, project, CAPEX, OPEX

## 1 Introduction

The problem of forecasting a capital expense to build and an operating expense to run and maintain a data centre attracts attention of many experts and organizations. In recent years a variety of activities were aimed at developing methodologies and tools for solving this problem [1-5]. However, they do not take into account the specifics of the Russian market, and usually do not estimate potential revenue of the data centre. In our work we collected and systematized the information about the construction and operation of more than 70 data centres in Russia. On this basis we have developed the model for preliminary evaluation of the data centre creation project.

We presume that any evaluation of the effectiveness of investment consists of comparing the income from the use of the system with the cost of its implementation and support.

Integrated view on the costs of creating and supporting data centre can be obtained on the basis of

valuation technique Total Cost of Ownership (TCO) [6], which highlights all costs that arise both during the data centre creation and while providing services to consumers.

Assessing the income of data centre one should take into account features of pricing, as well as the current status and trends in the structure of services provided.

The final step in the assessment is the discounting of possible cash flows of the data centre project that allows to determine the effectiveness of investment by calculating the most common indicators of Net Present Value (NPV) and Internal Rate of Return (IRR).

Initial data for this model were collected from publications, materials of analytical studies and from database "SPARK-Interfax» (SPARK – Professional market and company analysis system). A fragment of original table that contains data about 77 projects is listed in Table 1. Missing project data are marked by "-".

TABLE 1 Data of projects of datacentres construction

Company	Year	City	Costs	Area, sq. m.	Technical area, sq. m.	Number of racks	Power MW	Tier level
Irkutskenergosvyaz	2014	Irkutsk	2.5 BN RUR	10000	3200	1300	-	3
Government of Chelyabinsk region	2014	Chelyabinsk	27.269 M RUR	12000	-	1600	16	-
Vimpelcom	2013	Yaroslavl	4 BN RUR	15000	3000	1200	10	3
Rostelecom	2013	Moscow	30 M USD	11500	10000	-	40	3 sert
Inoventica	2012	Vladimir region	90 M RUR	-	300	60	0,45	3
Rostelecom	2012	Stavropol	-	280	250	20	-	-
Elektronnaya Moscva	2012	Moscow	114.5 M RUR	530	250	93	1	-
UTK	2008	-	320 M RUR	1000	300	-	1,5	-
M1, Stack	2007	Moscow	15 M RUR	2500	-	-	-	-
Technogorod	2007	Moscow	NCA = 10 M RUR	1500	-	-	1	3

\* Corresponding author e-mail: grekoul@hse.ru

As can be seen from this fragment, project data are often incomplete: for example, information can be available about capital expenses and area of datacentre, but there can be no data about quantity of racks. In this case it is possible to use relations identified from the market research performed by iKS-Consulting company (2012): the ratio of the area of datacentre and quantity of racks is usually 3.5 sq. m. per rack.

In some projects in the "cost" column we can meet acronym NCA. This means that the cost of the project was not disclosed, but the analysis of the data from the SPARK-Interfax allowed to detect an increase in the cost of Non-Current Assets (NCA) during construction of the data centre. As a result, we used different sets of projects to develop different models presented in the article.

Prices were transferred to a common currency (U.S. dollars), and to the level of prices of 2013 using the index of prices for machinery and equipment used in the construction [7].

**2 Estimation of the cost of data centre construction**

The expenditure part of the project includes the capital and operating costs.

Summarizing research materials about the cost structure [3, 8-11] we recorded the following components of the data centre construction cost:

1. Construction of the building (~ 10-15%).
2. Electrical power input (~ 20-25%).
3. Diesel Generating Set (DGS) (~ 0-5%).
4. Optical cable (~ 0-5%).
5. Service equipment (~ 60-65%). (Uninterruptible power supply, cooling system, raised floor, networking equipment, etc.).

To construct a regression model of the data centre construction cost we have collected and processed data of 77 data centres.

This data included the following parameters: data centre location (Moscow / St. Petersburg /other Regions of the Russian Federation), year of construction, data centre space, number of racks, total cost of construction, Tier level.

In the analysis were included data centres of the most popular in Russia Tier 3 level.

To estimate the cost of the construction we selected the following indicators as the basic parameters of a data centre:

- Planned space of a data centre (*S*)
- Expected number of racks (*N*)

Then the value of CAPEX (CAPital EXpenditure) can be defined by the following relations obtained in the form of regression models (for all projects in Russia). Dependence of CAPEX on the data centre area:  $CAPEX(S) = -2856583 + 22136 * S$ .

Characteristics of the regression model:  $R^2 = 0.72$ , P-value for *S* is  $4.73E-11$ . P -value for the free variable 0.76. This means that, despite the fact that the cost of construction (CAPEX) is negative when datacentre area is smaller than 129 sq. m., this coefficient is not significant statistically. At the same time, such a small P-value for the coefficient *S* allows us to state with confidence that the cost of construction per sq. m. lies between 19.797 – 24.475 dollars (with a standard error of \$ 2339). This result correlates well with the experts estimations of CAPEX: 15-25 thousand dollars per 1 sq. m. (iKS-Consulting company (2012)). Figure 1 shows the model in a graphical form.

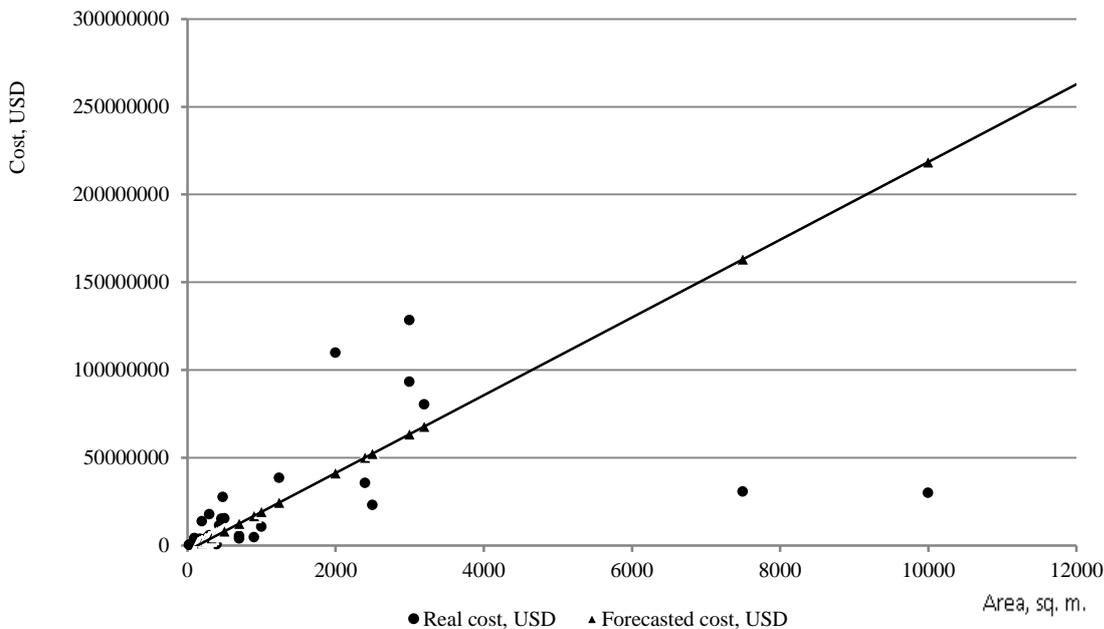


FIGURE 1 The regression model of dependence of CAPEX on the area of data centre

Dependence of CAPEX on the on the quantity of racks in the data centre:

$$CAPEX(N) = -3375063 + 78751 * N.$$

Characteristics of regression model: (R2 = 0.8, P-value for N is 2.01E-13. P-value for the free variable is 0.67). The cost of construction per one rack lies between

71 994 - 85 508 USD (with the standard error equals to \$6757). The resulting cost of construction per one rack is about 3.5 times more than the cost of construction per one sq. m. of a data centre area. It is practically equal to the ratio adopted in marketing research. Figure 2 shows the model in a graphical form.

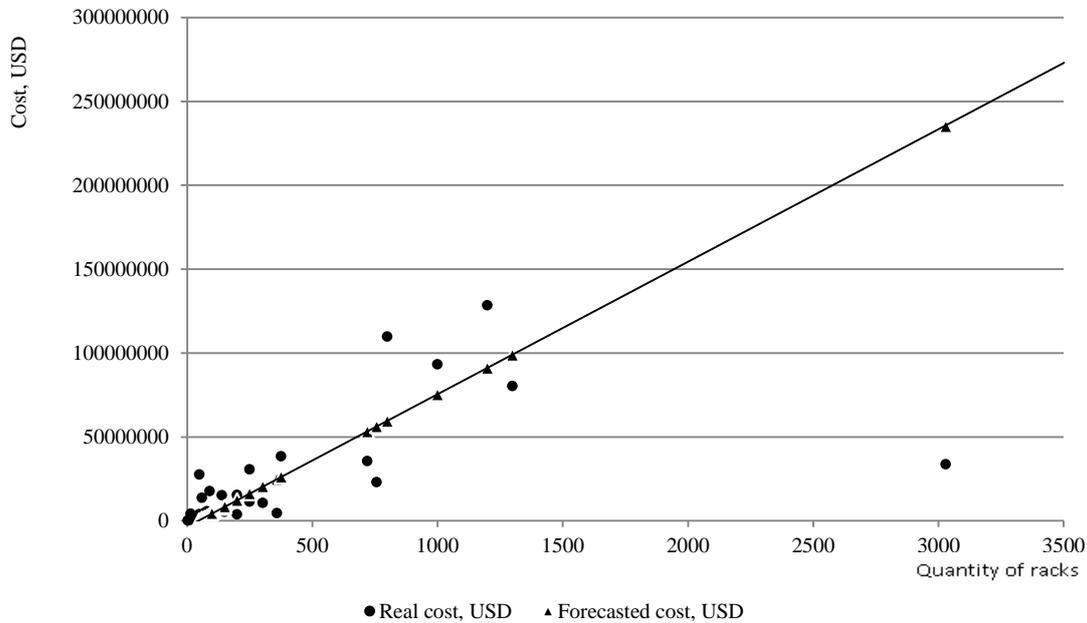


FIGURE 2 The regression model of dependence of CAPEX on the quantity of racks

Separate regression models for groups of data, classified by location of a data centre, were also formed.

For Moscow city we have got:

$$CAPEX_{Moscow}(S) = -2651754 + 22612 * S.$$

Characteristics of regression model: R2 = 0.7, P-value for S is 0.002.

$$CAPEX_{Moscow}(N) = -3315038 + 73616 * N.$$

Characteristics of regression model: R2 = 0.77, P-value for N is 0.0009.

For other regions of Russian Federation we have got:

$$CAPEX_{Reg}(S) = -8077885 + 26586 * S.$$

Characteristics of regression model: R2 = 0.88, P-value for S is 7.93E-09.

$$CAPEX_{Reg}(N) = -6171183 + 95935 * N.$$

Characteristics of regression model: R2 = 0.99, P-value for N is 5.38 E-17.

Deviations of calculated data from the average values of the entire sample are shown in Table 2.

TABLE 2 Cost of building data centre in geographical segmentation

Location of a datacentre	Moscow	Regions	Russia
Average cost of construction of 1 sq. m. (total CAPEX of sample / total S)	19 686	22 890	22 291
Average cost of construction of 1 rack (total CAPEX of sample / total N)	62 080	85 400	80 407
The forecasted cost of construction per 1 sq. m.	22 612	26 586	22 136
The forecasted cost of construction per 1 rack	73 616	95 935	78 751
Confidence interval for the forecasted cost of construction per 1 sq. m.	17 388 – 27 836	24 153 – 29 019	19 797 – 24 475
Confidence interval for the forecasted cost of construction per 1 rack	59 201 – 88 031	93 366 – 98 502	71 994 – 85 508
Average datacentre area	1 583	1041	1258
Average quantity of racks	509	242	349
Deviation of the cost of construction per 1 sq. m.	13%	14%	-1%
Deviation of the cost of construction per 1 rack	16%	11%	-2%

Due to the lack of representativeness of the sample it was not possible to include in the regression model such parameters as the Tier level and creation time.

The influence of the Tier level on the cost of a data centre construction can be represented by multiplying the value of obtained CAPEX by the correction coefficient

( $C_c$ ) which values were obtained on the basis of research [12]:

- $C_c = 0.8$  for the Tier 2 level,
- $C_c = 1$  for the Tier 3 level,
- $C_c = 1.8$  for the Tier 4 level.

To take into account the effect of time on the cost it is possible to use the research results presented in [21]. It shows that the construction cost of 1 square meter increases by about 30% per year.

Thus, the cost of building of a data centre in the year  $G$  may be defined by the relation:

$$CAPEX_G = CAPEX * C_c * 1.3^{(G-2013)}.$$

### 3 Estimation of the cost of data centre operation

There is a considerable difference in the structure of operating costs of Russian [3, 13-15] and overseas data centres [2, 12] (see Figure 3).

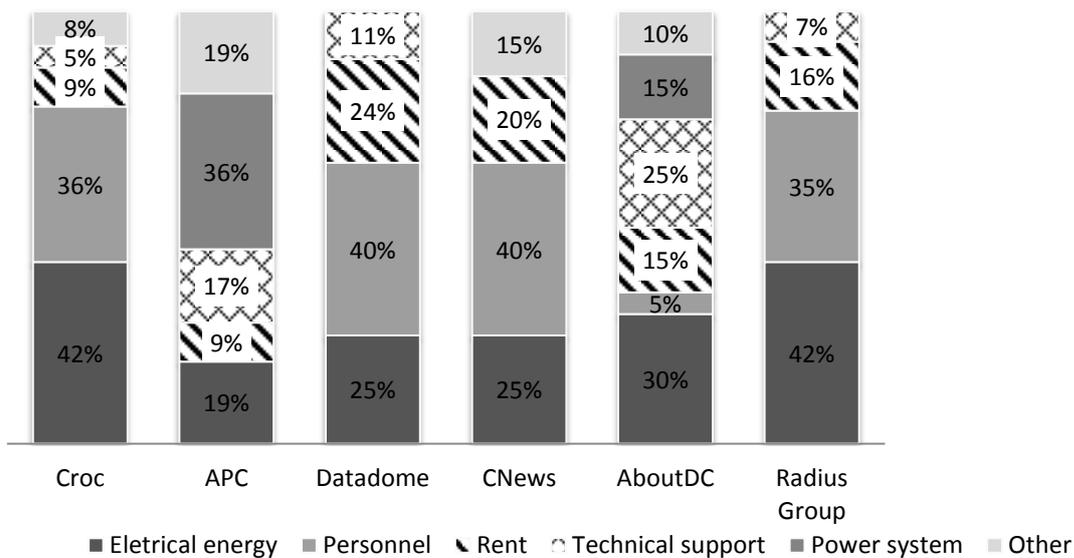


FIGURE 3 Structure of a data centre operational expenditures

The differences in OPEX (Operational Expenditure) structure are caused by different approaches to clustering subgroups of operating costs, as well as the specifics of the Russian economy (in particular, the difference in wages, energy costs, etc.). However, every research includes energy costs, rent, staff and technical support in the cost of data centre operation. Other subgroups of operating costs can vary considerably and for our purposes they are integrated into the group of "other expenses".

We assume that the operating costs are divided into 5 groups:

1. Electricity payments (~ 40-45%).
2. Rent (~ 15-20%).
3. The salary fund (~ 25-30%).
4. Technical support (~ 10-15%).
5. Other costs (~ 10-15%) (equipment depreciation, insurance, etc.)

Now operating costs can be calculated on the basis of such indicator as data centre power: we can estimate the

operating costs by defining the cost of electricity consumed. Analytical materials include very different estimations of the share of electricity in operating costs: from 25% to 45%. We have analysed the financial statements of a number of data centres and have chosen the value 42% as the most acceptable estimate.

If the project determines the number and capacity of racks in the data centre, the calculation of the consumed energy is trivial. But in this case it is necessary to take into account the distribution of energy consumption in data centres (Figure 4).

The analysis of the information about the structure of the electricity consumption in some data centres [2, 12, 14] showed that all data centre racks consume about half of the total electricity. This means that all data centre racks consume electricity, which makes about 21% of all operating costs. At the moment racks for 42 U with 5 kW power are the most popular but the cost of 1 kW of energy in different regions is very different.

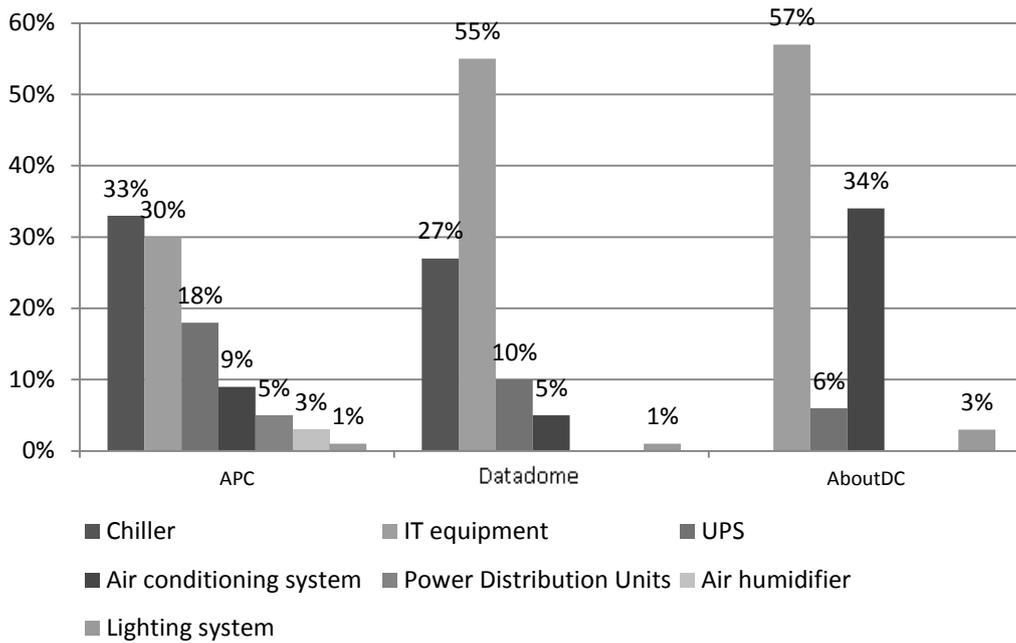


FIGURE 4 Distribution of energy consumption in the data centres

However, if the early stage of the project does not include definitions of equipment specifications, estimation of the power should be performed on the basis of statistical data. To forecast the data centre capacity ( $M$ , MW / h) we built a regression model identifying the relationship between this index and basic characteristics of a data centre (space and the number of racks):

$$M(N) = -0.18 + 0.012 * N$$

Characteristics of regression model:  $R_2 = 0.93$ ,  $P$ -value for  $N$  is  $6.61E-16$ .

$$M(S) = 0.24 + 0.003 * S$$

Characteristics of regression model:  $R_2 = 0.66$ ,  $P$ -value for  $S$  is  $2.28E-07$ .

Figures 5 and 6 present these models in a graphical form.

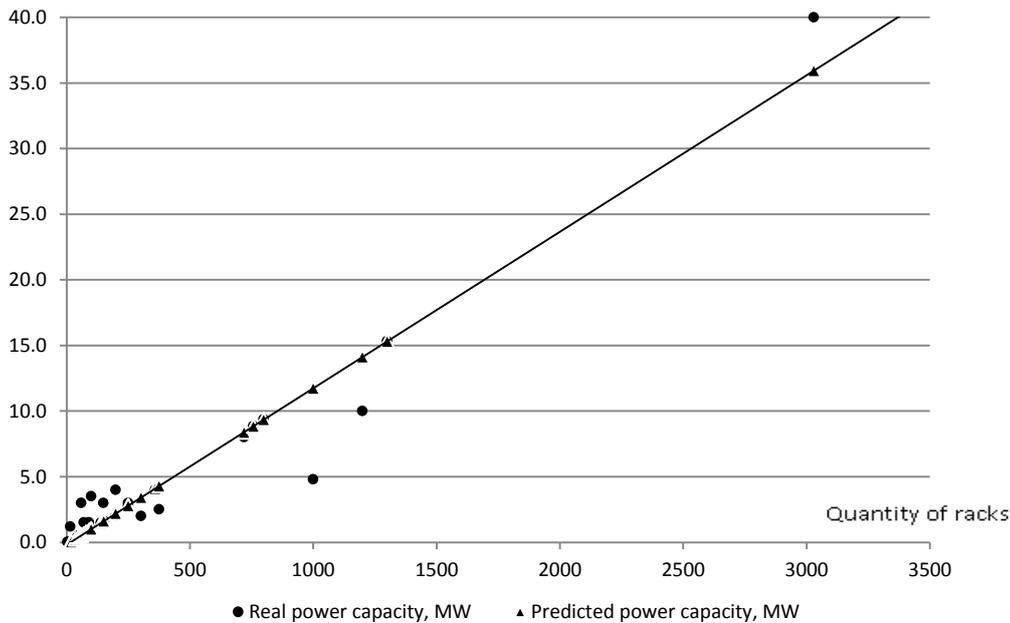


FIGURE 5 Regression model of dependence of datacentre power consumption on the amount racks

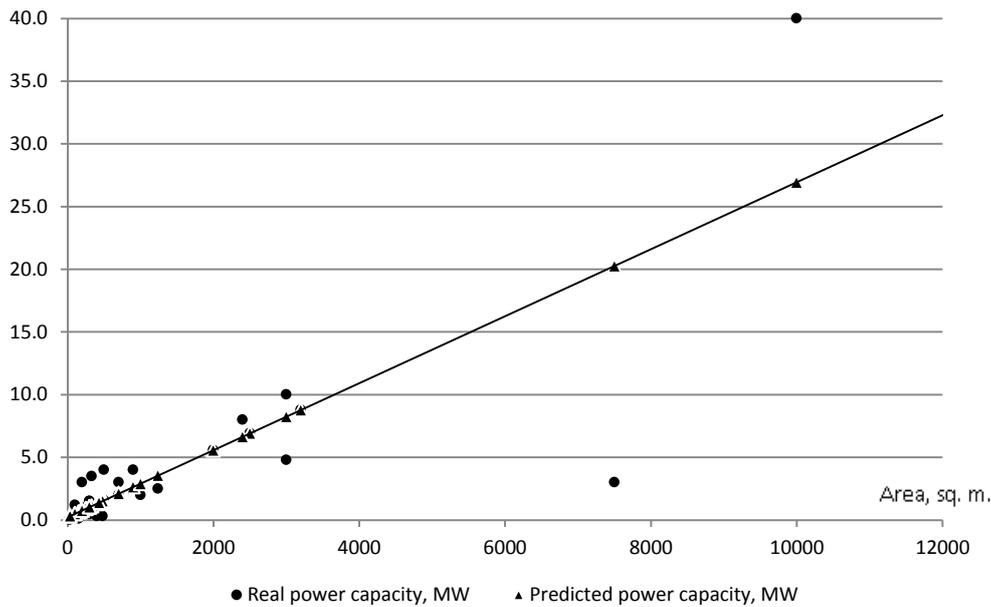


FIGURE 6 Regression model of dependence of the datacentre power consumption on the area

To obtain more accurate results it is better to use the number of racks (N).

Proceeding from the previous analysis of the structure of the operating costs, an estimation of the value of OPEX (during the year) may be performed using the following relations:

$$OPEX = M * [365 \text{ days}] * [24 \text{ hours}] * e / 0.42,$$

$$OPEX(N) = (-0.18 + 0.012 * N) * 20857 * e,$$

$$OPEX(S) = (0.24 + 0.003 * S) * 20857 * e,$$

where  $e$  is the cost of electricity, \$ / kW / h.

It is necessary to note that the price of energy differs in different regions of Russia is and this will lead to a certain error in estimation.

Degree of conformity between calculated and actual values of operating costs is illustrated for several Russian data centres in Table 3.

TABLE 3 Comparison of actual and calculated data centre operating costs

Data centre	Safedata	OBIT	Parking.ru
OPEX/racks deviation	31%	34%	14%
OPEX/sq.m. deviation	9%	31%	14%

For modelling future cash flows we should take into account changes in the prices of the main components of operating costs. In particular, in case of constant parameters of capacity, electricity payment costs will rise according to electricity prices, payroll - according to the real wages, etc.

Deep analysis requires not only consideration of the cost structure, but also assessment of the dynamics of

changes in the future. This requires understanding of how prices will vary for each subgroup of costs (as the number in each group will have the same characteristics if data centre doesn't change).

So, we use the following forecasts of the Ministry of Economic Development [16] (Figure 7).

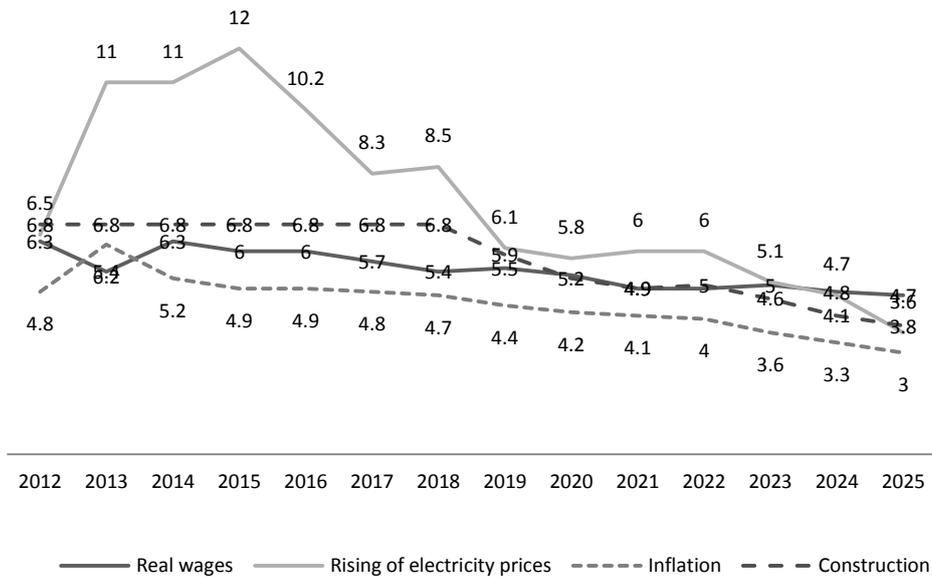


FIGURE 7 Forecast of macro economical indicators, %

We constructed the forecast of the structure of operating costs, using the following assumptions:

- The base year (average year in used researches) - 2011
- The original structure of operating costs is calculated as the average result of used researches
- Personnel costs increase in proportion to the growth of real wages
- The cost of electricity is rising in proportion to the price of electricity for all consumers
- The cost of renting the building increases in proportion to the growth of prices in construction market

- All other costs are rising in proportion to the average annual inflation rate.

The result is the forecast presented in Figure 8.

As can be seen, the overall structure of operating costs does not change significantly, but we can note the increase of the share of energy costs and reducing of the share of personnel costs in the total cost.

These data can be used in the process of specification of forecasted costs. As a generalized estimation of future changes in operating costs, it is possible to use the weighted average of the growth rate of operating costs equal to 6%.

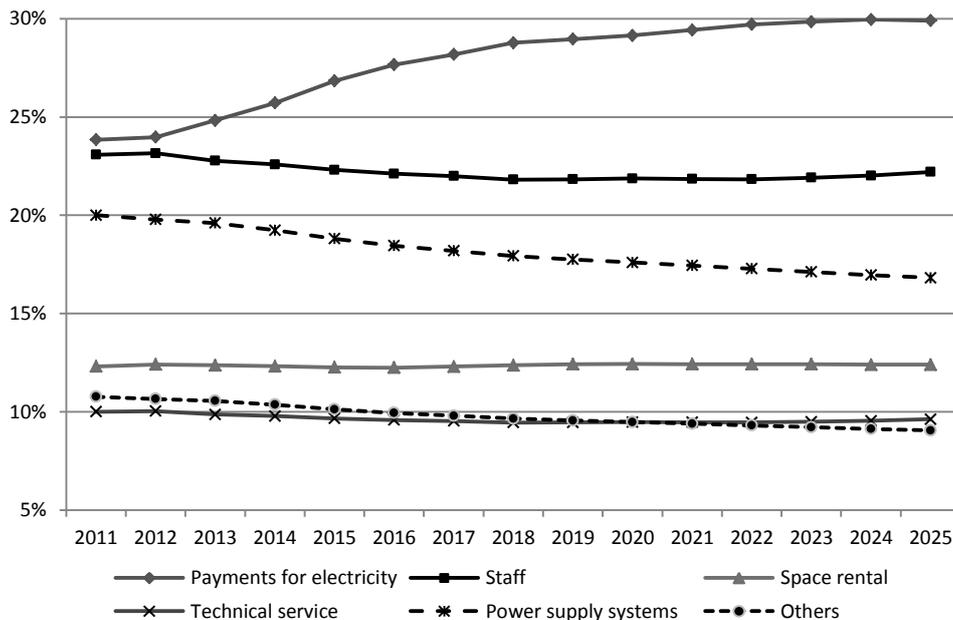


FIGURE 8 Forecast of changes in the structure of operating costs

The average annual growth of the components of operating costs was determined using forecasts of macroeconomic indicators of Russian economy, published by the Ministry of Economic Development of Russian Federation [16]. Weighted average annual growth of operating costs (including the share of each group in operating expenses) is 6%.

Now operating costs of a data centre in the year G can be determined by the following relation:

$$OPEX_G = OPEX * 1.06^{(G-2013)}$$

#### 4 Forecasting the revenue of a data centre

To calculate cash flows such indicators as revenue per rack and revenue per square meter of the data centre space are usually used.

As a basis for our analysis we have used the results of revenue forecasts from the data centre market research of iKS-Consulting company (2012) (Figure 9, 10). Absolute values of revenue presented in these research significantly differ from our results because iKS-Consulting company used generalized data (total number of racks, total area of data centres in Russia, summarized declared income etc.), not the data of concrete projects. But this data are quite enough to analyse the behaviour of revenue indicators. We have used trend estimation technique to make and justify some statements about tendencies in these data. As we can see, the values of the coefficients of determination for the linear trends are within 0.00 – 0.41. This indicates the absence of explicit changes in the data over time. The variability of indicators is mainly determined by other factors.

It allows us to conclude that these indicators are quite stable in nature and will not change significantly in the foreseeable future.

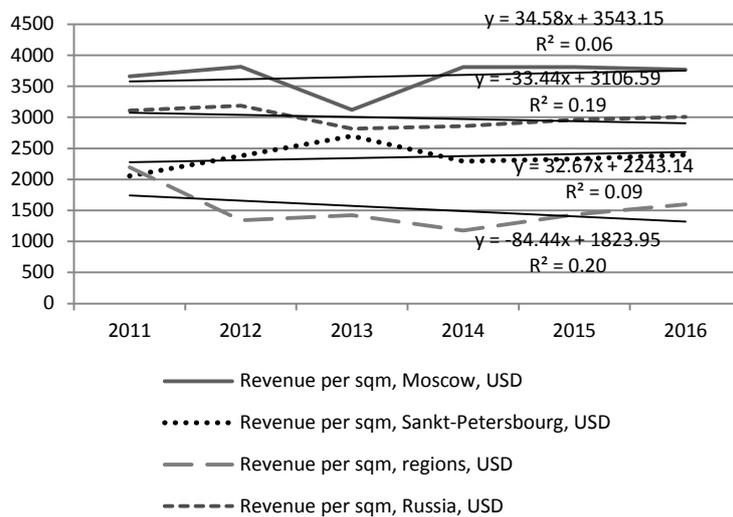


FIGURE 9 Forecast of revenue per sq. m in geographical terms

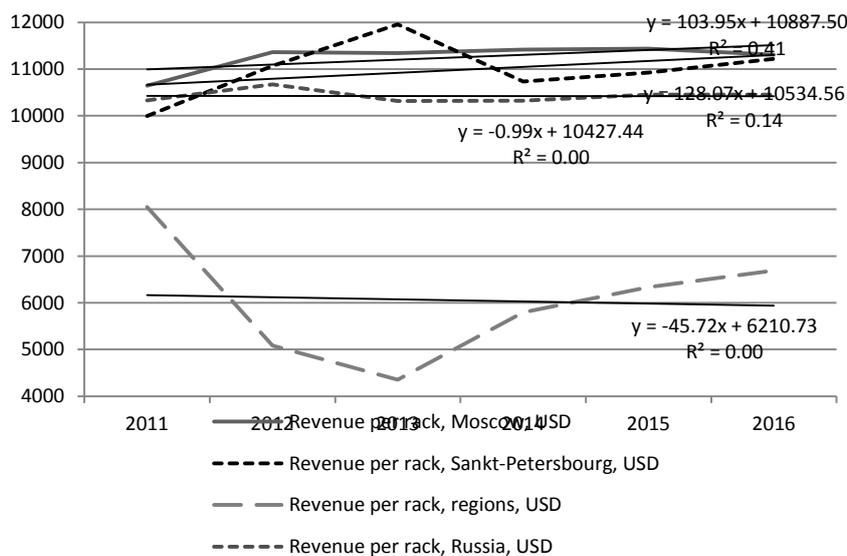


FIGURE 10 Forecast of revenue per one rack in geographical terms

Due to significant differences in accounting policies in different companies information about revenue of data centres is very heterogeneous and often contradictory (for example, stated in the reports values of revenue per sq. m differ by 37-48 times for data centres with the same location). Therefore, for the analysis we selected the data of several Moscow typical data centres with relatively clear accounting system. Finally, we have got the following average indices: revenue in Moscow and Moscow region is 5 700 \$ / sq. m or 12 430 \$ / rack. For St. Petersburg and other Russian regions these values can be adjusted with the help of the above mentioned forecasts (report of iKS-Consulting company (2012)).

It is necessary to consider the impact of the degree of occupancy of a data centre on revenues. We assume that the average time to achieve the level of normal occupancy of a commercial data centre is two years [17]. We believe that the filling of a data centre proceeds linearly. It means that at the end of the first year the level of occupancy will reach 50% and the average level of occupancy for the first year will be 25%. At the end of the second year occupancy will reach its normal level,

and the average level for this period will be 75% of normal.

As a result the revenue (Rev) of future periods (since the moment of commissioning data centre) can be represented by the following relations:

$$Rev = 0,25 * N * RPR + 0,75 * N * RPR + \sum_{q=3}^{Per} (N * RPR)_q,$$

$$Rev = 0,25 * S * RPM + 0,75 * S * RPM + \sum_{q=3}^{Per} (S * RPM)_q,$$

where:

- *RPR* (Revenue per rack) - index of revenue per rack in every given year
- *RPM* (Revenue per square meter) - index of revenue per square meter of data centre space in every given year
- *Per* - period for which the analysis is performed (number of years).

The degree of compliance of real and calculated revenues of several data centres [18] is shown in Table 4.

TABLE 4 Comparison of actual and calculated data centre revenue

Data centre	Safedata	Dataline	DEMOS-Internet	OBIT
Revenue per rack, deviation	12%	14%	17%	-6%
Revenue per sq. m, deviation	12%	17%	24%	-11%

### 5 Estimation of financial performance of a project

The final step in the evaluation of the effectiveness of investment in a data centre is to build a discounted cash flows model. It includes income and expenses of future periods, loans needed for starting the project and the future payment of interest on them. As a result the net cash flows will be obtained and finally they must be discounted. The estimations of *NPV* and *ROI* allow us to receive a final evaluation of the effectiveness of the project.

During the preplanning of the project the following parameters should be defined:

- Space of a data centre (in the ideal case - the number and power of racks, as well).
- Period of construction of a data centre (usually – about 2 years).
- Duration of the data centre lifecycle.

The Net Present Value of a data centre creation project, launched in the year *G*, can be determined (using data about technical space) by the relation:

$$NPV = -0.5 * CAPEX_G - 0.5 * CAPEX_{G+1} * i + 0.25 * [-OPEX_{G+2} + S * RPM_{G+2}] * i^2 + 0.75 * [-OPEX_{G+3} + S * RPM_{G+3}] * i^3 + \sum_{q=4}^D [-OPEX_{G+q} + S * RPM_{G+q}] * i^q$$

where:

- *q* - serial number of the year of project,
- *i* - discounting coefficient,
- *D* - duration of the data centre lifecycle.

### 6 Conclusion

Of course, every data centre is unique, so there can't be a universal tool that would exactly forecast cash flows. Our model provides a reasonably accurate estimate based on the most aggregated data centre characteristics – space and number of racks. In comparison with the data of real projects the errors of our model lie within following ranges:

- COPEX(N)* - 2%-16%
- COPEX(S)* - 1%-14%
- for the operational expenditures
- OPEX/racks* - 14%-34%,
- OPEX/sq. m.* - 9% -31%,
- for the revenue
- RPR* - 6%-17%
- RPM* - 11%-24%.

for the capital expenditures

Parameters of the model depend considerably on the place of the data centre construction (Moscow, St. Petersburg or other regions of the Russian Federation).

As an example, let us consider the project of typical data centre in the Moscow area (of the Tier 3 level, with the area of 1 000 square meters). Using the proposed model we obtain the following results:

- $CAPEX = -2651754 + 22612 * S = 620\ 007$  thousand rubles = 19 960 thousand dollars.

## References

- [1] *Datacentre downtime: How much does it really costs?* 2012 <http://www.stratus.com/~media/Stratus/Files/Library/AnalystReports/AberdeenDatacentreDowntimeCost.pdf> 14.03.2013
- [2] *Determining Total Cost of Ownership for Data Centre and Network Room Infrastructure White paper # 6 Revision 2005* <http://www.apc.com/go/promo/whitepapers/index.cfm> 20.03.2012
- [3] *Economics of data centre// Radius Group* 2013 <http://www.rvip.ru/1065/document1071.shtml> 25.07.2013
- [4] Hamilton J 2010 *Overall Data Centre Costs.Perspectives* James Hamilton's Blog <http://perspectives.mvdirona.com/2010/09/18/OverallDataCentreCosts.aspx> 22.03.2012
- [5] Patel C, Shah A 2005 *Cost Model for Planning, Development and Operation of a Data Centre* Internet Systems and Storage Laboratory - HP Laboratories Palo Alto, <http://www.hpl.hp.com/techreports/2005/HPL-2005-107R1.pdf> 22.03.2012
- [6] *Total Cost of Ownership (TCO): Definition, Meaning and Use* Encyclopedia of Business Terms and Methods 2013 ISBN 978-1-929500-10-9
- [7] *Information about prices of acquired basic construction materials, components and structures* <http://www.fedstat.ru/indicator/data.do> 15.02.2013
- [8] *Data centre: expenditures structure* Radius Group <http://www.rvip.ru/document999.shtml>
- [9] Pavlov A, Basistiy D, Kusakin D 2012 *Business plan and financial assessment of creating a data centre with high level of reliability* <http://www.iks-media.ru/issue/2012/7/4430815.html> 18.02.2013
- [10] Samoylov Y 2010 *What does it take to build a data centre?!* <http://www.dtlr.ru/press-tsentr/sobytiya/cto-nam-stoit-tsod-postroit> 22.03.2013
- [11] *System of professional analysis of markets and companies SPARK-Interfax.* <http://www.spark-interfax.ru/Front/Index.aspx> 15.05.2013
- [12] *Data centre statistics* 2011 AboutDC.ru <http://aboutdc.ru/page/64.php> 20.08.2012
- [13] Khoroshikh D 2010 *All advantages of external data centre* <http://library.croc.ru/document/3336/> 20.08.2012
- [14] Pavlov A 2007 *Engineering solutions for reducing the operating cost of the data centre* <http://www.osp.ru/data/670/942/1238/10.pdf>
- [15] Volchaninov L, Ilyin S 2008 *The crisis will force introducing technologies of optimization data centre* <http://www.r-style.com/presscentre/publications/cnews-cod-08/> 28.02.2013
- [16] Ministry of Economical Development <http://www.economy.gov.ru/> 28.02.2013
- [17] Gabrielyan V 2009 *How to build private data centre* Byte Russia <http://www.bytemag.ru/articles/detail.php?ID=15495> 2.09.2012
- [18] Soukhov R 2008 *How to estimate expenditures of data centre?* <http://www.osp.ru/nets/2008/04/4944787/> 18.09.2012
- [19] *Data-centres are growing as mushrooms* 2013 <http://www.it-weekly.ru/analytics/tech/45540.html> 18.09.2012
- [20] Pavlov A 2011 *Data centre Checklist* <http://www.outsourcing.ru/content/rus/327/3274-article.asp> 15.10.2012
- [21] Orlov S 2008 *Data centre: Russian realities* <http://www.osp.ru/lan/2008/09/5535759/> 15.10.2012

## Authors



**Vladimir Grekul, born on September 15, 1949, Sakhalin, Russia**

**Current position, grades:** Professor, Head of the Department of Corporate Information Systems National Research University Higher School of Economics, Moscow.

**University studies:** Kiev Higher Engineer Aviation Training School of the Air Force

**Scientific interest:** IT management

**Publications:** 90

**Experience:** 40 years experience in design and implementation of information systems.



**Liubov Baydalina, born on January 4, 1990, Chelyabinsk, Russia**

**Current position, grades:** Junior Business Analyst, Schneider Electric, Business Development department.

**University studies:** MA degree, Magna Cum Laude (Honors degree).

**Scientific interest:** business processes modelling and optimization, modelling of market development, estimate of NPV of IT project.

**Publications:** 6.

**Experience:** Datacentres market research

# An improved energy-efficient distributed storage system

Hongyan Li\*

*School of Information Management, Hubei University of Economics, Wuhan, China*

*Received 12 June 2014, www.tsi.lv*

---

## Abstract

Energy consumption has increasingly become a serious problem in contemporary data centres. The electricity bill contributes a significant fraction of the Total Cost of Ownership (TCO), and it is predicted to increase at an even faster pace in the following years as extremely large volume of data are being generated on a daily basis which would necessitate corresponding storage capacity to hold them. As a profitable work-around step toward the energy problem within the cloud infrastructure, in this paper, we propose REST, an energy-efficient cloud storage, which is built upon a cluster-based object store similar to GFS. It achieves high energy-efficiency by cleverly exploiting the redundancy already present in the system without compromising the inherently well-established schemes for consistency, fault-tolerance, reliability, availability, etc., while maintaining a reasonable performance level. By modifying slightly the data-layout policy, REST can safely keep a large amount of the storage nodes in standby mode or even powered off entirely most of the time. Deploying a sophisticated monitor, it also provides the flexibility to power up sleeping or powered down nodes when necessary to accommodate to the variations in workloads. Trade-offs between energy efficiency and performance can be conveniently made by simply adjusting a trade-off metric in REST. The FileBench and real world workload experimental results demonstrate that power savings can reach 29% and 33%, respectively, while still providing comparable or even surprisingly better performance.

*Keywords:* cloud storage, data centre, distributed storage, energy efficient, power saving

---

## 1 Introduction

With more and more internet services, outsourced storage services being concentrated in data centres and cloud computing infrastructures, added by a variety of data-intensive applications, like Google search engine, genetic projects, satellites images, data centres are increasingly getting filled with extremely large amount of data. Such huge storage requirements pose a lot of challenges to the IT management in terms of privacy, security, efficiency, energy consumption, etc. Even worse, the storage requirements have been reported to be rising by 60% annually. The phenomenal amounts of data in data centres not only call for tremendous investment on hardware, e.g. disks, to provide the corresponding storage capacity, but also need continuing power supply to feed the hardware. In large data centres, the energy cost consumed by the IT equipment over their lifetime is comparable to the hardware investment and occupies a significant portion of the TCO [1]. To make the situation more complicated, the power consumed never comes alone, but with many accompanying negative side effects, such as environmental impacts, noises, health disturbance. For example, according to the data from EPA, generating 1 kwh of electricity in the United States gives birth to an average of 1.55 pounds of carbon dioxide (CO<sub>2</sub>) emissions and consuming the same amount of electricity would further incur more emissions and heat which needs other additional electricity to keep

data centres temperature from getting too high [2]. What's more, cutting the electricity bill is compelled by external factors and especially critical in certain situations. For example, in major cities the electricity prices are extraordinarily high and the requirements of increasing power supply may not be possible to be fulfilled at all [14].

Fortunately, many researchers from academic and industrial background have extensively investigated the power consumption problem and put forward a number of fruitful techniques to attack the problem over the past several years [3-8]. Generally speaking, those techniques can be classified into two broad subcategories: component-based solutions [6, 12] and system-level solutions [4, 5, 7, 8]. Since the data centres must be designed to account for the peak workloads, they are most of the time relatively over-provisioned due to the wide temporal variations exhibited by the workloads, e.g. diurnal peaks and troughs, which enable those techniques to be effective. However, while power-proportionality can be achieved relatively easier for some kinds of components, e.g. using dynamic voltage and frequency scaling (DVFS [13]) for CPUs, non-CPU components, especially like disks, are not power-proportional. Thus, to conserve energy consumption in storage subsystem, taking advantage of the observed idle periods between successive disk accesses is a common practice. For example, put some disks into standby mode under light or moderate workloads and try to manufacture and prolong

---

\* *Corresponding author* e-mail: hongyanli78@aliyun.com

idle periods [4, 5, 7], both of which can be also perceived as forms of power-proportional.

Still, there remains a relatively less-explored spectrum of large-scale storage system power-saving space, which is from a high-level system design perspective. Distributed storage system systems such as GFS [9], HDFS [10], and KFS [11] are widely deployed as the backend storage infrastructure in large data centres and cloud computing infrastructures due to the aggregate high I/O performance and cost advantages over conventional SAN and NAS solutions. However, they were originally designed with little if not absolutely no power considerations. They were established on the assumption that instead of on enterprise-grade disks, they would be running on clusters consisting of hundreds of thousands of commodity servers, for which the unpredictable and sporadic fault or failure happenings should be considered as norm rather than exception [9]. Thus, they must be designed to be able to gracefully tolerate large numbers of component faults with little or no impact on service level performance and availability. Facing this hard situation, replication technique, a method widely thought to be able to provide high system reliability, better performance and high availability, had become a natural technique candidate to be deployed. As a result, each data block is replicated to a user defined level, typically three, replicas in those systems, resulting in large amount of redundancy. However, such redundancy at the same time introduces a lot of overheads to the system in aspects of storage capacity, replica consistency, networking bandwidth requirements, power consumption, etc. For the energy conservation consideration, an obvious and simple ideal occurs to us: is it possible to put some of the redundant nodes into power-saving mode or entirely power them down under light or moderate workloads while maintaining the existing sophisticated built-in mechanisms?

In the remaining of this paper, we present REST, a new cloud storage scheme based on a replicated, distributed file system KFS [11]. It aims to improve energy-efficiency without incurring severe performance degradation from the perspective of applications. The central point is that it turns down some redundant nodes under light or moderate workloads to conserve energy and also keeps the capability to power them up again in response to variations in workloads to prevent the performance from degrading too far. Our main contributions are: an energy efficient cloud storage scheme capable of reacting gracefully to variations in workloads and an architecture deploying new technology to manage cache and consistency under energy-saving mode which could be also potentially applied under disconnected conditions that happen frequently within large-scale systems. The design specifics are detailed in section 3. Experimental results have showed that REST has lived up to our expectations very well.

The remainder of the paper is structured as follows: Section 2 introduces the background and motivation.

Section 3 details the design principles and implementation specifics. Section 4 presents our evaluation methodology and results. Section 5 makes a conclusive remark.

## 2 Background and motivation

Basically, many widely deployed distributed file systems share some common design and implementation strategies and tactics. They mainly consist of three parties assuming respective responsibilities: the client library, metadata server (MDS) and chunk-servers. The client provides API facilities to access the file system; MDS is the central component taking the responsibility of managing the whole file system name space; chunk servers are physical nodes where the data are actually stored. The objects stored in the systems are partitioned into chunks. Each chunk is replicated on multiple chunk servers to guard against disk or machine failures and to provide high performance and availability. The central MDS is implemented as an in-memory data structure so as to provide fast access speed, as it is visited much more frequently by normal operations and scanning checks [9]. It keeps all the metadata information. It is checked periodically to persistent storage to guarantee reliability and fast recovery in the event of MDS failure. Chunk servers communicate with MDS through frequent heartbeat messages reporting their storage status and receiving corresponding acknowledgement from MDS to maintain high availability. If the MDS notices some replicas are unavailable, it initiates re-replication process to prevent the data block getting under-replicated. From a high-level perspective, read and write requests are handled similarly. They both firstly go forward to the MDS to get the necessary information, like chunk servers' location and lease information, then directly contact the corresponding chunk servers to complete the data transfer, rendering the MDS out of the data transfer path. By doing this way, the involvement of MDS is minimized and the chances of its becoming potential bottleneck would thus be minimized.

However, there exists a significant difference between the phases of data transfer for read and write requests. For writes, all the nodes hosting the replicas should remain powered on for the purpose of strong consistency during the entire write process. For example, GFS [9] uses lease mechanism to define the write order of the replicas and applies pipelining technique to propagate the content. But for reads, after getting a list of chunk servers hosting the replica, the client would contact (may use an optimal algorithm, e.g. least distance first, to minimize the latency) one of them to fetch the replica, leaving the other ones unvisited if the firstly chosen chunk server succeeds in servicing the request. To get a holistic view of the read and write distribution, we modified the file access APIs of dbench4.0 [15] to those APIs exported by KFS, and then executed dbench4.0 on KFS. We configured KFS to consist of a MDS and 60 chunk servers. After the running

process finished, we analysed the statistics from MDS and all chunk servers. The overall number of reads is what MDS has recorded, while the total number of writes is the multiplicative result of the MDS recorded number and the replication factor. The read and write requests' cumulative distribution function (CDF) figure is shown in Figure 1.

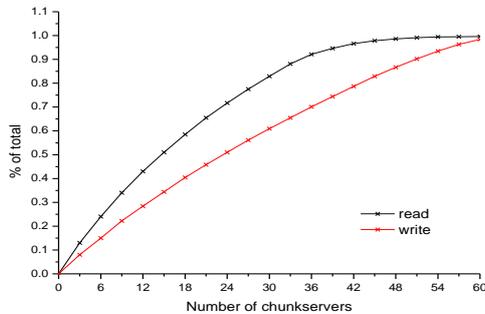


FIGURE 1 CDF of read/write among all the chunk servers

As shown in Figure 1, it shows a cumulative distribution function across the numbers of the chunk servers versus the percentage of read/write that the chunk servers experienced over the course of the running process. As pictured in the figure, read distribution embodies a wide discrepancy from write, i.e. read is more skewed than write, whereas write is approximately linear, which results from the inherent different handling processes. For example, nearly 83% of the reads were concentrated on 50% of the chunk servers, while 50% of the chunk servers had absorbed 61% of the total write operations. That's because, for read, every time the request for the same data block would with a strong likelihood be routed to the same chunk server hosting a copy (either primary or replicas) of the data block due to the same internal decision policy, e.g. judging by IP addresses, with the exception of the cases of chunk server having failed or the system topology having changed. But for write, because of the load-balance hinting data placing policy, especially for newly written data block, the write operations are more likely to be distributed uniformly among all the chunk servers. Such observed skewness in the read/write distribution motivates our work: is it possible to remove the write traffic from some chunk servers to make them presented with more read-dominated access patterns, which would lead to more skewness and exploit such skewness to conserve energy?

It is worth noting that due to hardware limitations, the experiment was not conducted on 60 physical chunk servers, but with each server hosting multiple chunk servers. Chunk servers are represented by different processes running on different ports with respective dedicated directories providing storage capacity. Such configurations have the following important implications: the chunk servers bear great homogeneity, have relatively flat network topology and are rendered to be equally subject to the network conditions. By contrast, in a genuine environment of large-scale storage system

installation, like data centres, the situation is far away from here. Usually, it embraces a wide range of heterogeneity resulting from its hundreds of thousands of commodity components possibly differing widely in aspects of storage capacity, computing capability, network bandwidth, etc. Furthermore, it is typically constructed in a hierarchical form using different levels of switches to connect racks and chunk servers, which would make different constituents subject to various network conditions. However, under such a complicated situation, it is reasonable to assume the existence of similar or even better observations that we had experienced with our relatively simple and flat experimental setup. The assumption is based on the failure preference phenomenon observed in [16, 17], stating that failures tend to happen again to where they had happened before at a much higher probability. This phenomenon would potentially translate into more skewness in read/write distribution, which would present us a great opportunity to conserve energy as discussed in the following sections. Another thing that should be pointed out is that with no doubt we would rather like to admit dbench can never be representative of all read-world workloads, but we hope it would be reasonable to claim that though simple, it would shed some light on the problem we are discussing in this paper.

### 3 Design and implementation

As discussed in the preceding section, the skewness in I/O behaviour, especially in read distribution, may provide us with great potential to conserve energy. That is because skewness implies that some redundant chunk servers would remain idle even in the presence of I/O workloads. Thus, the node's storage subsystem or even the whole node if the node's only role is providing storage function can be transitioned into power-saving mode in a manner that is oblivious to the applications running on the system. This section is devoted to discuss how the energy-conservation potential can be realized. It deals with the details of the design and implementation of our prototype system REST. Specifically, it first outlines the design principles and goals that guide our design, and then presents the challenges and problems that should be resolved, followed by the overview architecture of REST. Finally, it discusses in depth the strategies and tactics we have deployed to achieve energy-efficiency purpose.

#### 3.1 GOALS AND CHALLENGES

The ultimate goal of our work is to build an energy efficient storage system through exploiting the redundancy inherently existing in a replicated, cluster storage system. We choose the distributed file system KFS [11] as our basic architecture and develop our prototype on that. During the process of development, we strive to fulfil the following design disciplines and principles, which also act as our design guidelines:

- Changes made to the original system should be minimized.
- Energy efficiency should be obtained not at the expense of severe system performance degradation.
- Energy efficiency should be obtained without compromising reliability, availability, consistency and failure resilience.
- Trade-offs between energy efficiency and high performance should possibly be made by users and the system should be reasonably flexible to automatically respond to the workload variations.
- Failures of the components should be handled gracefully, ideally transparently, without adversely impacting the applications.

KFS was initially designed with almost no power-awareness considerations, but with its focus on building a high reliable, high available, high aggregate performance, scalable and fault-tolerant storage system on commodity-level components. Its salient feature is that reliability and availability can be well guaranteed even in the presence of failure occurrences to some components. The rationale behind our design is to conserve as much energy as possible while maintaining basic functions and desirable excellent features. To achieve this goal, the following challenges and problems should be well addressed:

- Data placement policy should be power-aware. The policy deployed by KFS places data blocks in a random way which would potentially cause all chunk servers to be highly correlated to each other. Using that policy, for example, at most  $N-1$  chunk servers can be transitioned into power-saving mode to conserve energy if the system's replica factor is  $N$ .
- How many redundant chunk servers and when are they going to be put into power-saving mode while maintaining a reasonable performance level? And how to differentiate those deliberately powered-down servers from those that actually malfunction?
- Under what circumstances should the sleep servers be woken up? And in what manner are they woken up?
- What measurements should the system to take if failures occur?
- How to guarantee consistency among all replicas when some are temporarily not unreachable for power-saving purpose?

We present our approach to addressing those challenges and problems in the following subsections. At first, we give an overall description of the architecture of REST, and then we dive into the details of the design strategies and tactics integrated in REST.

### 3.2 SYSTEM OVERVIEW

As mentioned previously, our simple strategy is to exploit the redundancy present in the system and the skewness in

IO distribution to realize energy-efficiency purpose. In order to do so, we have slightly changed the architecture of the original system and additionally integrated some functional modules into it. As shown in Figure 2, there are three main roles in REST: MDS as in the basic infrastructure with a newly added functional module instructor (not shown in the figure), loggers that are performed by dedicated high-performance servers and chunk servers. Note that the constant heartbeat messages among them are omitted for simplicity.

Compared with its basic architecture, REST differs itself in the following aspects: the back-end chunk servers are partitioned into several subsets, the new entering of loggers and an instructor that decides when and how many chunk servers are going to be powered down and up. Since the details are given in the subsequent subsections, we only give a brief description in the remain of this subsection. As shown in the figure, we strategically partition the entire space of chunk servers into several subsets. The most important one is named kernel subset coloured in red, and the others are named backup subsets coloured in green and light green, respectively. Ideally, the kernel is expected to remain powered on, while backups are to be kept powered down under light and moderate workloads. It is the responsibility of instructor that determines when to power backups down and when to up based on a number of factors. The up/down commands are piggybacked in the acknowledgements to periodic heartbeat messages from chunk servers. The loggers are designated the functions of providing temporary storage space to hold the data destined to the chunk servers that are powered down for that period of time, forwarding them to the corresponding chunk servers when they are powered on again and reclaiming the space. Read requests are sent to loggers as shown by read step: 1) (solid-line) if the requested data blocks exist there and the loggers are not too overloaded, otherwise they would be served by kernel or backup subsets as shown by read step; 2) (dash-line). Write requests are handled as usual in KFS if the kernel and backup subsets are all powered on, in other cases they are written to the powered-on chunk servers and loggers, and then immediately return to the clients to indicate write completion.

### 3.3 POWER-AWARE DATA LAYOUT

The inherent limitation that prevents the original system from powering down more than  $N-1$  chunk servers in an  $N$ -way replicated system lies in that the initial data block assignment policy logically imposes a strong tie between every pairwise, despite they are physically separate. For example, if  $N$  or more chunk servers were powered down, the data block of which all  $N$  replicas unluckily happen to reside on the  $N$  powered-down chunk servers or on a subset of the powered-downs would have been rendered unavailable.

To eliminate this limitation, we divide the whole chunk servers into several independent subsets named

kernel and backups, and for each data block, it is guaranteed that there would be at least one replica in each of the subset. This is achieved by the new data placement policy. For the convenience of our discussion, we define the following symbols:  $N$  for the total number of chunk servers in the system;  $r$  for replication factor;  $k$  ( $1 \leq k \leq \lfloor \frac{N}{r} \rfloor$ ) for the size of kernel. As a system parameter,  $k$  is critical to the system performance and energy efficiency, since it determines the kernel's performance and thus, how often backups are going to be powered down and up, as discussed in the next following subsection. Fortunately, its value determination can be hinted by the individual performance of the chunk servers and operator's well-understandings of the characteristics of the expected workloads or it can be gradually adapted to the most suitable value using test-and-tune method. Backup subsets are obtained by equally partitioning the rest of the chunk servers, making its size  $\lfloor \frac{N-k}{r-1} \rfloor$ . The

definitions of kernel and backups can be specified in advance in a system configuration file, and are all maintained in MDS. The chunk servers that belong to kernel and backups are typically chosen to have the same fault-tolerance properties, like a rack, respectively, for they are highly logically related, i.e. the failing of any one of them would render the data blocks residing on it unavailable within the corresponding subset. Interestingly, such strong relationship among physically separate nodes is exactly the reason that motivates our new data placement policy. However, doing so has several advantages, for example, utilizing less precious networking bandwidth and creating opportunity for power savings for switches, which would otherwise remain powered on, but also bears potential disadvantages, like being easily bottlenecked and forming high-temperature spots in large data centres.

When allocating a data block, MDS firstly allocates one data block for it from kernel, and then allocates the remaining replicas from backups, one from each. Within each of the subset, we also balance the data blocks among the chunk servers. With this kind of data placement policy, the minimum availability would safely be guaranteed by kernel, and backups can be powered down freely and independently when necessary without compromising system availability. MDS dynamically tracks the status of kernel and backups and differentiates the failed chunk servers and those powered down chunk servers. The placement policy imposes several changes to the read/write processes as discussed previously. When the kernel fails down, one of or all of the backups can be powered up to take the role of the kernel and rebuild/recovery the kernel, based on the trade-offs between greater power savings and higher rebuild/recovery speed.

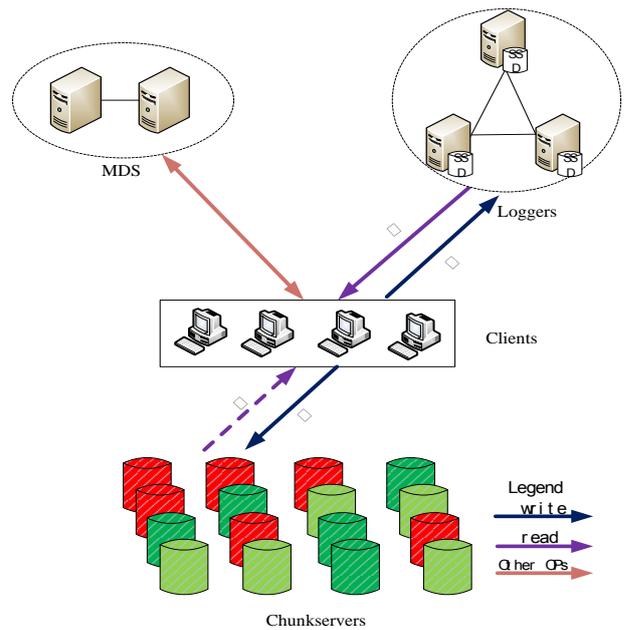


FIGURE 2 REST architecture

### 3.4 INSTRUCTOR

In reality, due to wide variations in workloads, it is improper to keep the number of powered down chunk servers constant. In REST, there is a central role named instructor that determines when and how many chunk servers are required to be powered down or up using WOL [18] technique. It periodically makes the decision based on a complex combination of multiple factors. Formally, the decision-making can be expressed as:

$$n = f(\lambda P, \beta R, \mu F, \eta C) \tag{1}$$

$$\lambda + \beta + \mu + \eta = 1 \tag{2}$$

In the above equations,  $n$  is the result of the decision making and indicates the number of the to-be-powered down or up chunk servers.  $P$  is defined as the ratio of performance to power, both of which are monitored by instructor, it reflects the system power efficiency;  $R$  is the required performance level, usually assigned as the minimum requirements that must be met;  $F$  represents the instructor's predication of the near-future workloads. It is obtained by analysing past workload characteristics and is used to instruct the decision. If it is predicated that the forthcoming write requests are very large or the writes would last for a long time, it will generate a larger value of  $F$ , indicating it is desirable to power more chunk servers up. The heuristic is that it tries to directly write large writes to all the chunk servers as possible as it can be, while redirect as many small writes as possible to the loggers, which is decided by two reasons: the loggers' capacity is limited and we want small write requests to be returned faster, since their latencies are more sensitive to the applications. If it is predicated that there are

enormous read-intensive access patterns in the near future, it would pro-actively prepare to power up more chunk servers to mask the relative long time, usually minutes, to power chunk servers up and to avoid degrading read performance too much.  $C$  is the health conditions of the kernel. Since the kernel is expected to be working all the time and unexpected failures in the kernel would be expensive due to the minimum availability being compromised, kernel health conditions [17] should be constantly monitored and reported to the instructor to take corresponding precautions, like replace new devices, if the health conditions are not so good.  $\lambda, \beta, \mu$  and  $\mu$  are their corresponding coefficients and can be assigned by the operator. They represent their relative importance in the process of decision making. For example, if the operator gives higher priority for performance than power-saving consideration, she/he can easily achieve that by assigning  $\beta$  bigger than  $\lambda$ .

In addition to the above power-down and up scheme, there are some other scenarios that the chunk servers should be powered up in REST. For example, in our current prototype implementation, the loggers and chunk servers' statuses are periodically reported to MDS through heartbeat messages. When MDS notices that if the overall space utilization of loggers has reached certain threshold, e.g. 80% or some members of kernel are unavailable, it would force all the chunk servers to be woken up immediately.

### 3.5 LOGGERS

In REST, unlike other techniques using dedicatedly reserved space on the existing devices, e.g. Erased [19], DIV [7], write-offloading, we have used dedicated servers equipped with solid-state drives (SSD) to log the data destined to the temporarily powered down chunk servers. From the point of view of our considerations, there are four reasons for doing so. At first, we want the logged data to be persistently stored more reliably, and SSDs can fulfil our requirements, additionally, SSDs themselves consume much less power than conventional disks counterparts. Secondly, it is desirable to place logged data in different fault domains, typically different racks in data centres. Thus, even if the kernel subset fails down due to whole rack or switch failing down [26], it is still achievable to restore the data to the latest status with the logged data and the newly powered up servers. Thirdly, with the technology drastically advanced, the shortcomings formerly associated with SSD have already been well overcome [20] and their prices are not that high as before making them increasingly become acceptable and practical to be deployed in production systems. Lastly, SSD flash drives have blazingly fast read speed, including both of sequential and random patterns. We deploy this attractive characteristic to provide high performance for read requests by diverting read requests to loggers firstly if they are not overwhelmingly overloaded with traffics.

To take maximum advantage of SSDs and avoid their excruciating slow write shortcoming, we deploy a log-structured [21, 22] store engine to record the logged data. For each logged data block, there is corresponding information specifying whether it is logged and logged on which logger in its in-memory metadata entry in MDS namespace. In addition, each logger maintains a hash table in memory for itself. The table maps the unique 64-bit chunkID to the location where the on-disk entry resides within the store engine. The on-disk entry of logged data block is self-contained. It is composed of a logger header and a logger body. The logger header contains the following information: chunkID, version number, destinations, logger factor (meaning how many copies the chunk should be propagated to newly powered up chunk servers to in order to maintain the required replication factor) and chunk checksum. The logger body is the content of the data block.

For every data block write request, it would be firstly tried to be forwarded to the powered on chunk servers and if any of the destination chunk servers is powered down temporarily, it would be logged to one of the loggers before returning to the application. Specifically, MDS would choose a logger and send it a log request containing the necessary information and update its corresponding metadata entry. Receiving the log request, the logger appends an on-disk entry to the store engine, and then inserts a mapping entry into the local hash table, or replaces the mapping entry corresponding to the chunk if it has been inserted previously. That implies, for each logged chunk, there is only one mapping entry, which points to the most recent version of that chunk. For every read request, MDS would look the chunks in the namespace and preferably forward it to the loggers if it has been logged in any of the loggers, otherwise forward it to the powered on chunk servers. From the data paths of write and read requests, it would be expected that loggers would improve the system performance due to the log-structured design and SSDs' superior random read performance, which is actually proved by our experiments.

Considering the limited space of the loggers, we have designed each logger to initiate propagating and reclaiming processes at regular intervals, i.e. one minute. The propagating processes scans through the entire local hash table, and for each entry, it tries to contact those corresponding destination chunk servers to see whether they are reachable and if so it sends the logged data to them and update the entry information. The reclaiming process scans the on-disk entries from the beginning, and reclaims the space occupied by those entries whose logger factor is zero. It works in a similar way to the cleaner in [21]. Thanks to the two periodically-run processes, the loggers' utilization of space is reasonably prevented from going high quickly.

3.6 ENERGY MODELLING

To estimate the energy efficiency of REST, in this subsection, we analyse the energy models for both of the original system and REST. The following table summarizes the parameters in our analysis:

TABLE 1 Energy modelling parameters

Symbol	Description
$N_c$	The total number of chunk servers
$N_l$	Number of loggers
$T_i^k$	Length of the $k^{\text{th}}$ active interval of the $i^{\text{th}}$ chunk server
$A_i$	Number of active intervals of the $i^{\text{th}}$ chunk server
$P_i^k$	Power of the $i^{\text{th}}$ chunk server at time $t$ within the $k^{\text{th}}$ interval $0 < t \leq T_i^k$
$E_i^u$	Energy consumed by the $i^{\text{th}}$ chunk server to transition up
$E_i^d$	Energy consumed by the $i^{\text{th}}$ chunk server to transition down
$C_i^u$	Count of transition ups of the $i^{\text{th}}$ chunk server
$C_i^d$	Count of transition downs of the $i^{\text{th}}$ chunk server
$P_j$	Power of the $j^{\text{th}}$ chunk server when power up
$T_j$	Active time of the $j^{\text{th}}$ logger. Loggers have no up-and-downs

Two points should be noted about Table 1. One is that we have not outlined the energy consumed by MDS, since our main purpose is to compare the energy efficiency of the original system and REST. And we assume the amounts of MDS energy of them are approximately equal. We denote it as  $E_m$  in the following discussion. The other one is that we treat the instantaneous powers of loggers and chunk servers differently, i.e. we consider the power of loggers to be stable, while the power of chunk servers to be varying over time. Because we expect that individual chunk servers would experience bigger power gaps between peak power and the lowest power than loggers. Now we can calculate the total energy consumed by the original system using Equation (3).

$$E_b = E_m + \sum_{i=1}^{N_c} \sum_{k=1}^{A_i} \left( \int_0^{T_i^k} P_i^k dt \right). \tag{3}$$

Since the original system has no power up-and-downs, the numbers of active intervals of all the chunk servers are the same and equal to 1. Thus, Equation (3) can be translated into Equation (4).

$$E_b = E_m + \sum_{i=1}^{N_c} \left( \int_0^{T_i^1} P_i^1 dt \right). \tag{4}$$

Involving with the chunk servers power up-and-downs, REST has a more complex energy consumption formula:

$$E_R = E_m + \sum_{j=1}^{N_l} (T_j \times P_j) + \sum_{i=1}^{N_c} (E_i^u \times C_i^u) + \sum_{i=1}^{N_c} (E_i^d \times C_i^d) + \sum_{i=1}^{N_c} \sum_{k=1}^{A_i} \left( \int_0^{T_i^k} P_i^k dt \right). \tag{5}$$

It says that the total energy is the sum of individual components' energy: MDS, loggers, chunk servers. The differences between the two energy models lie in that REST divides the energy of chunk servers into active status energy, transition energy and powered down energy, which is zero, and has additional energy consumed by loggers, while the original system keeps the chunk servers up all the time. The active energy is the sum of all the energy consumed by all the chunk servers over all their respective active intervals. The power up-and-down overheads are the sum of all the energy consumed by transitioning. Comparing with the original system, REST's potential gaining power savings stem directly from how many and how long chunk servers are powered down.

4 System evaluation

In this section, we evaluate REST from various aspects using benchmarking method and realistic workload. We also test the original system, which is referred to NO-REST later as baseline for comparison reason. Section 4.1 describes our test environment. Experimental results are discussed in the subsequent subsections.

4.1 EXPERIMENTAL SETUP

Our test bed consists of one MDS server and a number of chunk servers composed of 4 servers and 32 commodity PCs belonging to our lab's graduate students. The hardware configuration of the MDS server is characterized by a quad-core 2GHz CPU 16GB RAM and 16 1TB hard disks. One of the 4 servers is equipped with a quad-core 2GHz CPU, 4GB RAM and a Kingston 128 GB SSD disk drive and functions as a logger. The remaining three servers have the same configurations: a quad-core 2GHz CPU, 4GB RAM and 8 1TB hard disks structured as RAID5. The other 32 chunk servers bear a wide variety of configurations and performance, since they were purchased in the different years when their respective owners were enrolled in our lab.

We conducted our test using both benchmarking and realistic workload. The benchmark is FileBench [23], an application level workload generator that enables the users to emulate various workloads. Its Workload Model Language (WML) provides users with the capability to flexibly define the workload bearing different characteristics. A WML workload description is called a personality and it typically contains the following information about the workload: average file size, directory depth, the total number of files, and alpha parameters governing the file and directory sizes that are

based on a gamma random distribution [24]. FileBench can define the period of time for which the personalities are going to be run, and report the total number of performed operations at the end of each run. We selected four personalities included in FileBench to drive our testing. They are Web, File, Mail and Database servers, and their workload characteristics are specifically described in [24]. Each of them was run for a period of 1 hour. We deployed fuse support of REST and NO-REST to access the storage like conventional file systems. The realistic workload is a shared server workload in our lab. The server is shared by all lab staff doing their own jobs, e.g. upload/download files, doing backups, visiting CVS source code repository, etc., it is also the backup server of B-cloud [25] system developed at our lab which provides on-line backup services. We configured REST and NO-REST as the backend storage infrastructures of the server, respectively and monitored the server for 48 hours dating from 8:00 am GMT on Nov. 10 to 8:00 am GMT on Nov. 12 to observe the workload characteristics and how REST performs.

We emulated 76 chunk servers for both REST and NO-REST for fair comparison. Each of the 3 servers hosted four chunk servers, and each of those PCs hosted 2 chunk servers. The logger server emulated 2 loggers in REST. Both of REST and NO-REST were configured as three-way replicated systems. The kernel was set to include all those chunk servers hosted on the 3 high performance servers and its size was set to 24. For each run, we initially powered on all the chunk servers.

#### 4.2 PERFORMANCE IMPACTS

We compare the performance of different workloads in terms of the performance metric reported by FileBench. FileBench reports file system performance under different workloads in units of operations per second (ops/sec). Due to the peculiar characteristics of the Web server workload, FileBench would report much higher ops/sec than the other workloads. In order to avoid the results figures being too skewed, we present Web result in units of 100 ops/sec and ops/sec for the others. Figure 3 shows the performance results.

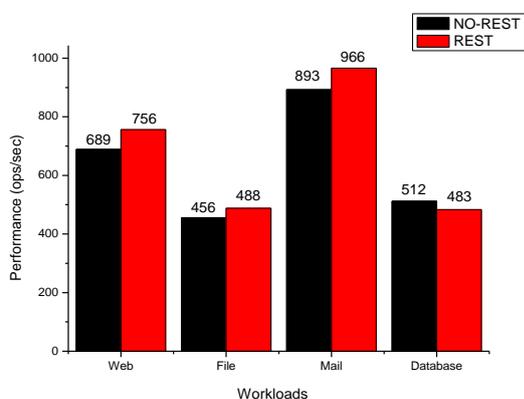


FIGURE 3 Performance of different workloads with REST and NO-REST

It is shown in the figure that REST well outperforms NO-REST for Web, Mail, File, achieving 9.7%, 8.17% and 7% performance gains respectively, while consuming less energy as discussed in the next subsection. And Database slightly lags behind NO-REST by a margin of 5.7%. Generally, the performance gains can be attributed to the newly added loggers in REST. Because reads are preferentially considered to be routed to loggers, which can provide excellent read performance, including both random and sequential reads and writes are redirected to loggers when the destination chunk servers are powered down temporarily. Due the log-structure, writes can be returned to applications much sooner, thus enhancing performance.

It is interesting to note that the amounts of percentage performance gains are closely related to the read/write ratio. And it's surprising to know that the performance gains are proportional to the energy savings. For example, Web conserves the most energy but with the most performance improvement, which are 17% and 9.7% of the NO-REST counterpart respectively, due to its highest read/write ratio and its sequential reading entire files patterns. Analysing the workloads, we know Web, File, Mail and Database's R/W ratios are 10:1, 1:2 and 20:1 respectively. It reveals that workloads with higher percentage of reads can get more energy savings and the chosen kernel can reasonably satisfy the read requests in most scenarios. However, there is an exception to that: Database has the biggest R/W ratio, but exhibits degrading performance. It is partly because, besides 200 readers, it launches 10 asynchronous writers and a log writers. In addition, they perform extensive concurrency and random read/writes, which would adversely prevent chunk servers from powering down. Another reason is that the writers' extensive writing operations would quickly cause the loggers' space utilization threshold to be reached and chunk servers to be waken up more frequently.

#### 4.3 TRANSITIONS AND POWER SAVINGS

In this section, we discuss the chunk servers transitions and power savings of REST. Power savings were calculated by substituting the parameters in our energy models with corresponding dynamically monitored numbers and representative real world values. As pointed out previously, the chunk servers' transitions are determined by the instructor. We define a trade-off metric  $T$  as  $\lambda / \beta$ , i.e.  $T = \lambda / \beta$ . It reflects the preference degree for power efficiency, meaning the bigger  $T$  is, the more power savings are desired. To see how instructor affects the power up-and-downs, we collected all chunk servers up-and-downs for the four workloads with varying  $T$  (as a side note, the preceding section's results are from experiments conducted with  $T = 1$  for REST) value while keeping  $\eta$  and  $\mu$  invariable and let their sum equal to 0.3. The statistics are summarized in Table 2.

TABLE 2 Transition counts summary

Workload	T=0.5	T=1	T=1.5	T=2
Web	126	113	110	95
File	243	220	201	198
Mail	189	165	146	134
Database	74	87	102	121

The table demonstrates that for Web, File and Mail, the number of transitions decrease with  $T$  increasing. Bigger value of  $T$  implies trying to achieve better energy efficiency, which means once chunk servers are powered down, they would be powered up under more serious conditions. Again, Database is the exception with transitions increasing with  $T$  increasing. As explained before, the nature of Database prevents chunk servers transitioning, but bigger value of  $T$  tries to force transitioning, causing REST struggling powering down and up.

Since the energy savings and performance of all the four workloads share similar characteristics, we take Mail as our example to discuss power savings and performance with varying value of  $T$ . Figure 4 shows the relationship between them. The energy savings and performance are percentages relative to its NO-REST counterpart. It is apparent that power saving and performance can be traded off with different  $T$  values. For example, we can get 29% power savings at 83% performance level, or alternatively, we can enjoy 112% performance level at the cost of less potential power savings, which is 10%. This has important practical implications for applications, i.e. performance and power efficiency can be flexibly traded off in REST by simply changing  $T$ .

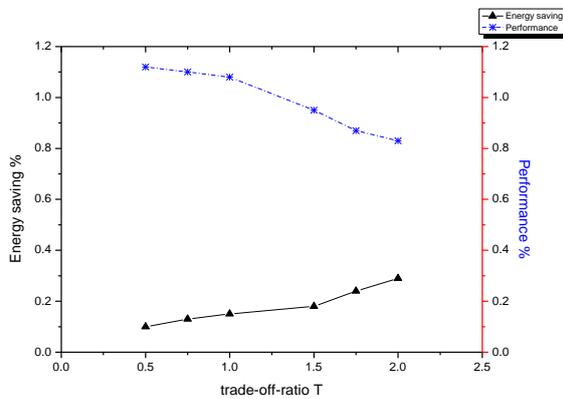


FIGURE 4 Energy saving of REST relative to NO-REST for Mail at different  $T$  value

4.4 REALISTIC WORKLOAD EXPERIMENT

For our realistic lab workload, we obtained similar energy efficiency and performance results. The power saving surprisingly reaches as high as 33% due to the wide variations in the workload, while maintains comparable performance level. In addition to that, we

sampled the REST dynamical number of powered on chunk servers every 2 hours over the 48 hours experiment period. The result is portrayed in Figure 5.

It shows that the workload exhibits periodicity, and REST responded to that in a power-proportional way. For each of the two days, we observed that at 12:00, 18:00, the up chunk servers are more than other times. We find that is because our lab members often saved their work on to the server before leaving the lab, resulting in peak time in workload. And for the spike points at 24:00 each day, we assume that is caused by our B-cloud server services. Several small and medium size companies are using B-cloud for their daily backup tasks, and they usually do backups at that point of time. We also notice there is a spike point at 04:00 on Nov 12 due to a kernel network partition occurrence. When failures occur to the kernel, REST would wake up chunk servers swiftly to rebuild/recovery the kernel.

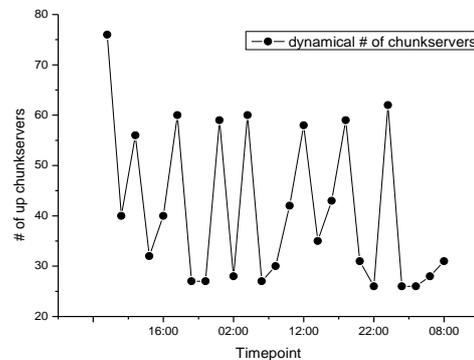


FIGURE 5 Dynamical number of powered on chunk servers over the experiment period References

Our real world lab workload experiment has revealed that REST has practical applicability. It can save energy by exploiting both the redundancy in the storage system and real workload characteristics, like periodicity and burstness in the workload.

5 Conclusions

In this paper, we present REST, a redundancy-based energy efficient cloud storage system. Motivated by the observations of workloads' periodicity and asymmetric phenomenon in read and write requests, we suggest in REST powering down the whole redundant chunk servers to achieve energy efficiency. We explicitly explain the techniques that we deployed in REST, including power-aware data layout, instructor, loggers and the energy models. Our experimental results show that a reasonable amount of energy savings can be achieved at comparable or unexpected better performance level, especially for realistic workload.

## References

- [1] Belady C 2010 In the data centre, power and cooling costs more than the IT equipment it supports *Electronics Cooling*
- [2] Zong Z, Briggs M, O'Conner N, Qin X 2007 An energy-Efficient Framework for Large-Scale Parallel Storage Systems *Proc. Int'l Parallel and Distributed Processing Symp., March 2007*
- [3] Benini L, Bogliolo A, de Micheli 1999 A survey of design techniques for system-level dynamic power management *IEEE Transactions on VLSI Systems* 8(3) 813-33
- [4] Papathanasiou A E, Scott M L 2004 Energy efficient prefetching and caching *In Proc. of the USENIX Annual Technical Conference, June 2004*
- [5] Zhu Q, David F M, Devaaraj C F, Li Z, Zhou Y, Cao P 2004 Reducing Energy Consumption of Disk Storage Using Power-Aware Cache Management *Proc. High-Performance Computer Architecture, 2004*
- [6] Zhu Q, Chen Z, Tan L, Zhor Y, Keeton K, Wikes J 2005 Hibernator Helping Disk Arrays Sleep Through The Winter *Proc. ACM Symp. Operating Sys. Principles, October, 2005*
- [7] Pinheiro E, Bianchini R, Dubnicki C 2006 Exploiting redundancy to conserve energy in storage systems *In Proceedings of the 2006 SIGMETRICS Conference on Measurement and Modeling of Computer Systems. Saint Malo, France, June 2006*
- [8] Kaushik R T, Bhandarkar M 2010 GreenHDFS: Towards an energy-Conserving, Storage-Efficient, Hybrid Hadoop Compute Cluste *In HotPower*
- [9] Ghemawat S, Gobiuff H, Leung S-T 2003 The Google File System *In the Proceedings of the 9<sup>th</sup> Symp. on Operating Systems Principles, Oct. 2003*
- [10] Konstantin S, Kuang H, Radia S, Chansler R 2010 The hadoop distributed file system *MSST*
- [11] <http://kosmosfs.sourceforge.net/>
- [12] Barroso L A, Hözlze U 2007 The case for energy-proportional computing *IEEE Computer* 40(12) 33-37
- [13] Chen Y, Das A, Qin W, Sivasubramaniam A, Wang Q, Gautam N 2005 Managing server energy and operational costs in hosting centres *In SIGMETRICS '05: Proceedings of ACM SIGMETRICS International Conference on Measurement and Modelling of Computer Systems* 33 303-14
- [14] Rakesh Kumar 2006 Gartner: A message from data centre managers to CIOs: Floor space, power and cooling will limit our growth, August 2006
- [15] <http://dbench.samba.org/>
- [16] Clement A, Kapritsos M, Lee S, Wang Y, Alvisi L, Dahlin M, T Riche M 2009 UpRight cluster services *In SOSp*
- [17] Pinheiro E, Weber W-D, Barroso L A 2007 Failure trends in a large disk drive population *In Proc. USENIX Conference on File and Storage Technologies (FAST'07), San Jose, CA, Feb. 2007*
- [18] <http://en.wikipedia.org/wiki/Wake-on-LAN>
- [19] Li D, Wang J 2005 Conserving Energy in RAID Systems with Conventional Disks *In Proceedings of the International Workshop on Storage Network Architecture and Parallel I/Os, Sept 2005*
- [20] <http://www.samsungssd.com/>
- [21] Rosenblum M, Ousterhout J 1991 The design and implementation of a log-structured file system *In Proc. ACM Symposium on Operating Systems Principles (SOSP'91), Pacific Grove, CA, Oct. 1991*
- [22] Ganesh L, Weatherspoon H, Balakrishnan M, Birman K 2007 Optimizing power consumption in large scale storage systems *In Proc. Workshop on Hot Topics in Operating Systems (HotOS'07), San Diego, CA, May 2007*
- [23] FileBench, July 2008  
[www.solarisinternals.com/wiki/index.php/FileBench](http://www.solarisinternals.com/wiki/index.php/FileBench).
- [24] Sehgal P, Tarasov V, Zadok E 2010 Evaluating performance and energy in file system server workloads *In Proceedings of the 8<sup>th</sup> USENIX Conference on File and Storage Technologies (FAST '10), February 2010*
- [25] Wei J, Jiang H, Zhou K, Feng D 2010 MAD2: a scalable high-throughput exact deduplication approach for network backup services *MSST*
- [26] Ford D, Labelle F, Popovici F, Stokely M 2010 Availability in Globally Distributed Storage Systems *In proceedings of the 9th USENIX Symposium on Operating Systems Design and Implementation, 2010*

## Authors



**Hongyan Li, born on October 12, 1978, Wuhan, China**

**Current positions, grades:** Now she is a PhD candidate in school of Computer Science & Technology of Huazhong University of Science & Technology. She is an associate professor at School of Information management, Hubei University of Economics. Since 2008 she is Member of IEEE.

**University studies:** received her M.Sc. in Computer Science (2005) from Central China Normal University.

**Scientific interests:** different aspects of storage Systems.

**Publications:** (co-)authored more than 20 papers.

# Research on ontology-based literature retrieval model

**Zhijun Zhang<sup>1, 2, 3</sup>, Hong Liu<sup>1, 2\*</sup>**

<sup>1</sup> School of Information Science and Engineering, Shandong Normal University, Jinan 250014, China

<sup>2</sup> Shandong Provincial Key Laboratory for Novel Distributed Computer Software, Jinan 250014, China

<sup>3</sup> School of Computer Science and Technology, Shandong Jianzhu University, Jinan 250101, China

Received 12 June 2014, www.tsi.lv

---

## Abstract

Proper understanding of textual data requires the exploitation and integration of unstructured and heterogeneous scientific literature, which are fundamental aspects in literature retrieval research. The traditional literature retrieval is based on keyword matching, and the retrieval results often deviating from the users' needs. In this paper from the perspective of ontology, we built shareable and relatively perfect medical enzyme ontology, which is the foundation of the study of domain ontology constructing method. The ontology-based full text retrieval algorithm is put forward, and a document retrieval system based on medical enzyme semantics is designed and implemented, which can not only implement intelligent literature retrieval, but also improve the recall significantly while keeping high precision. This system can employ in particular area moreover it can be used in different areas of the semantic retrieval, which can provide intelligent foundation for the expert systems in medical enzymes field, information retrieval and natural language understanding, etc. The experimental results on the public medical enzyme domain dataset show that our approach performs better than the state-of-the-art methods.

*Keywords:* Ontology, Literature Retrieval, Semantic Web, Ontology Construction, similarity computation

---

## 1 Introduction

With the rapid development of computer network technology, the demand for information storage, transmission and processing power increases rapidly, and the retrieval and use of mass information has become an important research and application domain in computer information retrieval technology. Information retrieval is mainly implemented through search engines on the Internet, and its query function is to crawl on the Internet to retrieve resources by an automatic processing program using web spiders, and to visit public sites for resource collection and to organize and process the information correspondingly so as to provide users with convenient retrieval service.

Search engine is an indispensable tool for people to surf the Internet now, but with the wide use of search engines, users' satisfaction degree becomes increasingly low. Many results of information retrieval cannot meet the demands of users, which are either retrieval insufficiency or irrelevant. That is mainly because the current search engine generally adopts full-text retrieval technology based on keyword matching, which returns too much useless information and cannot reveal the semantic level of user queries. Semantic retrieval, which broke through the defects of mechanical matching confined to the surface, can understand and deal with users' retrieval request from the semantic level of words expression.

Tim Berners-lee put forward the concept of Semantic Web [1]. Its basic idea is that the data on the Web can be understood by machine through embedding the mark, which can be read by machine and represents some kind of knowledge in the creation and release of Web information. Semantic web [2] is considered to be the next generation of network technology, whose core is to use metadata to describe resources on the network, and it has been widely applied to knowledge retrieval in the area of library information. Knowledge retrieval emphasizes the semantic matching based on knowledge, while the ontology is just making a standard description and organization of knowledge from the semantic level with a good concept hierarchy. Full text retrieval algorithm based on ontology combined with the ontology knowledge logic can further improve the correlation between the information retrieval results and target, to make the retrieval results comply better with the needs of users [3].

With the improvement of medical scientific research level, when faced with more and more medical information resources, it is very difficult for people to understand them in time and apply them reasonably. In order to reduce the difficulty of finding effective information for the medical workers, as well as to improve work efficiency and avoid duplication of effort, we must use scientific methods to organize and manage medical information resources effectively [4]. The application of ontology is mainly used in the field of information retrieval and knowledge organization of

---

\* *Corresponding author* e-mail: lhsdcn@jn-public.sd.cninfo.net

medical enzyme literature retrieval. Ontology can reflect restrictions between the mapping relationships of the lexical semantic and semantic which supports intelligent retrieval. The retrieval results will not be complete if the retrieval is done according to search terms provided by users only. When retrieving information, users usually hope that the results is something that they are interested in, and at the same time the engines could filter out irrelevant information, so that they could get the most valuable information. And when using ontology to retrieve information, researchers can use the ontology to map search terms to a set of specifications concept set automatically. Ontology can make a structured organization of information resources based on some knowledge organization system and can show links between content of information and knowledge organization system, and can connect domain knowledge base of ontology with information systems, so that in the process of using information, users can utilize ontology to understand specific concept and link the related resources more conveniently [5].

The remainder of this paper is organized as follows. In section 2, we provide an overview of ontology application in literature retrieval and related work. Section 3 introduces the construction of medical enzyme ontology and the overall framework of full-text retrieval system. Section 4 verifies the superior performance of full-text retrieval algorithm based on ontology in recall, precision and F-measure by the experimental data and results, followed by the conclusion and future work in section 5.

## 2 Related work

The traditional information retrieval is based on keyword matching, with the retrieval results often deviating from users' needs. At present semantic research in information retrieval mainly includes three aspects: natural language processing, method based on ontology and method based on concept. Voorhees first suggested using the concept of ontology to do query and expansion [6], and its basic idea is to use subclasses relationship of the ontology and synonyms. Ontoseek [7], developed by Guarino, is a retrieval system based on collaborative intelligent Agent. It can accurately describe the products or services in web pages, combining a content match mechanism driven by ontology with a formalization representation system with moderate expression ability, which tries to integrate ontology with big dictionary library, and provides users with a system in which interactive semantic query can be made with any words in the field. Although Ontoseek has quite effectively realized semantic functions, its degree of using ontology is not very high due to it is a content match mechanism. Swoogle [8] is a semantic web retrieval system based on the spider web concept, which extracts ontology from each searched text, comparing relationships between texts based on their ontology relevancy, but Swoogle method cannot search the

associated location ontology, as a result the ontology collection is not accurate. Maki puts forward the semantic retrieval methods based on the structure of ontology [9], effectively using the path in ontology to extend a user's query request. Navigli put forward a kind of query expansion method based on ontology annotation [10]. Literature [11] introduces an information retrieval model based on ontology: MELISA, which is used to retrieve literatures in the medical field. The Intelligent Information Processing Laboratory in the Institute of Computing Technology of Chinese Academy of Sciences established a kind of information retrieval server according to the theory of ontology and multiple intelligent agents [12], which can reflect dynamic changes of network information timely, and has good ability of information guidance.

At present, the discussion about the semantic web mainly concentrates on the research and development of the ontology. The concept of ontology initially originated in the field of philosophy, which was used to explain and illustrate the objective existence of things. With the continuous development of science and technology, ontology has been widely used in artificial intelligence, information retrieval, the semantic Web, natural language processing, and so on. In 1993 Gruber, who worked in Stanford University Knowledge System Laboratory (KSL), first presented a widely accepted definition of ontology: ontology is a specification of a conceptualization [13]. In 1998 Guarino proposed a concise definition of ontology and pointed out that ontology is a logical theory and is used to indicate a normal intended meaning of the vocabulary. Ontology is language-related; while concept is language-irrelevant. The concept of ontology has four layers of meaning: conceptualization, explicit, formal and share. Generally speaking, the ontology describes the relationships between the concepts in an application domain, which makes them to have uniquely definite meaning. With the help of ontology, we can obtain relevant knowledge of the field, and provide a shared understanding of the domain knowledge to facilitate communication between users and computers.

Ontology has become one of the effective tools for obtaining query expansion words. When using the method based on ontology for query expansion, we can select just a few extension words that are most closely related with the query expansion words by using the synonymous relationship, semantic entailment relationship, semantic extension relationship and semantic related relationship [14] between concepts. Association relationship between concepts is expressed by semantic similarity. By means of controlling the similarity threshold, we can adjust the scope of the extension concept set. Traditional models of semantic similarity calculation based on domain ontology between the concepts mainly has three types [15]: the semantic similarity calculation model based on semantic distance, the semantic similarity calculation model based on

content, the semantic similarity calculation model based on attribute. Domain-specific Ontology depicts both categories and instances in the field and their hierarchical relationships, inducts and abstracts the domain knowledge by defining elements such as categories, instances, attribute, relations, axioms and so on [16]. So far, many areas have emerged a large number of Ontology, such as medical ontology UMLS [17]. Ontologies that had been implemented mainly include: CYC, En-terprise, SENSUS, NKI, the massive knowledge system and medical knowledge library [18] presided by professor Cao Cungen, who works in the Institute of Computing Technology of the Chinese Academy of Sciences, and the research of software requirements elicitation method based on ontology developed by professor Jin Zhi at Beijing University and so on.

The above researches discussed the ontology retrieval model, but none of them involve ontology learning and inference problem, nor did they build their ontology models through formal ontology description language, which lead to neither fine nor precise use of ontology. At present there exists no large, shared, reusable, extensible medical enzyme ontology, thus the building of a good medical ontology is of vital significance. In this paper the medical enzyme ontology and its knowledge representation are built up based on the formal ontology theory. The ontology construction tools-Protégé is used to help building the ontology, which is the foundation through which knowledge acquisition and knowledge analysis are conducted based on medical enzyme ontology. Relevant function module has been realized in medical enzyme semantic literature retrieval system – MESLRS, which will provide intelligence foundation in the field of medical enzyme expert system, information retrieval and natural language processing.

### 3 The construction of ontology and framework design

#### 3.1 ONTOLOGY

The ontology formalization description is as Equation (1):

$$O = C, R, H_c, Rel, A_o, \quad (1)$$

where  $C$  is the set of concepts.  $R$  is the set of relations.  $H_c$  shows the concept hierarchy, that is, the taxonomy relation between concepts.  $Rel$  shows the non-taxonomy relation.  $A_o$  is the ontology axiom. And here  $C$  and  $R$  is two disjoint sets. It can be seen from the structure of the ontology that the task of ontology learning includes the acquisition of concept, the acquisition of relations between concepts and the acquisition of axiom. These three kinds of ontology learning objects form the levels from simple to complex.

#### 3.1.1 Ontology construction

Currently there is no unified method of the ontology construction and different methods are used in different research areas, so is the process of ontology construction. Now the way of ontology construction includes the following methods such as SEVEN, METHONTOLOGY, IDEFS, TOVE, FRAMEWORK, SENSUS, KACTUS and so on. At present many research fields have set up their own standard ontology, which indicates that the study of ontology model has entered a new stage. Generally speaking, the process of ontology construction can be divided into the following several stages: specification, conceptualization, integration, application and maintenance, in which knowledge representation, assessment and document management usually run throughout the entire process. Both knowledge representation and standardization are executed simultaneously in the first stage and evaluation is a stage of vital importance. In practice, ontology is usually constructed through method of accumulation, that is, a fundamental ontology is firstly constructed and then developed further. Many ontology constructions use a specific task as a starting point, which is easy for knowledge acquisition and descriptions of ontology function.

The early establishment and edition of ontology can only be conducted by experts and professionals of the field, while with the in-depth research and promotion of ontology, a series of ontology editing tools have been developed and each of which has specific advantages and disadvantages. Ontology development tools can be divided into six categories according to their different applications: ontology construction tools, ontology integration tools, ontology evaluation tools, ontology storage tools, ontology query tools and ontology annotations tools, and some of them might have multiple functions at the same time. Commonly used ontology development tools mainly include Protégé3.3.1, OntoEdit, OilEd, WebOnto.

#### 3.1.2 Ontology learning

At present manual way of ontology construction is a kind of main ontology construction method, which is not only slow, but less efficient. In order to improve the efficiency of ontology construction, the concept of ontology learning is put forward.

Ontology learning [19] refers to the procedure during which the desired ontology is obtained from the existing data resource automatically or semi-automatically by means of machine learning and statistical techniques. It is still difficult to realize the completely automatic knowledge acquisition, and ontology learning can only be semi-automatically done under the guidance of users. Ontology learning includes the learning of concept, the learning of relationship and the learning of axiom. Ontology learning can be divided into the following three

categories: ontology learning based on structured data, ontology learning based on unstructured data and ontology learning based on semi-structured data. The current evaluation of ontology learning system has not yet formed a unified evaluation criterion.

### 3.2 ENZYME ONTOLOGY CONSTRUCTION

We take the industrial enzymes for instance to introduce the establishment of a medical enzyme. One purpose of establishing industrial enzymes ontology is to assist the teaching and scientific research, which can automatically show students the contents of knowledge to learn according to the process of the teaching, to help students to get more accurate query results and solve perplexed problems because it can provide a certain amount of semantic search on these resources. Therefore, we need to provide the semantic relations between concepts as much as possible. Industry enzymes ontology has not only stored the classification knowledge of the industry enzyme, the basic method of enzyme preparation but also stored the relationship between them.

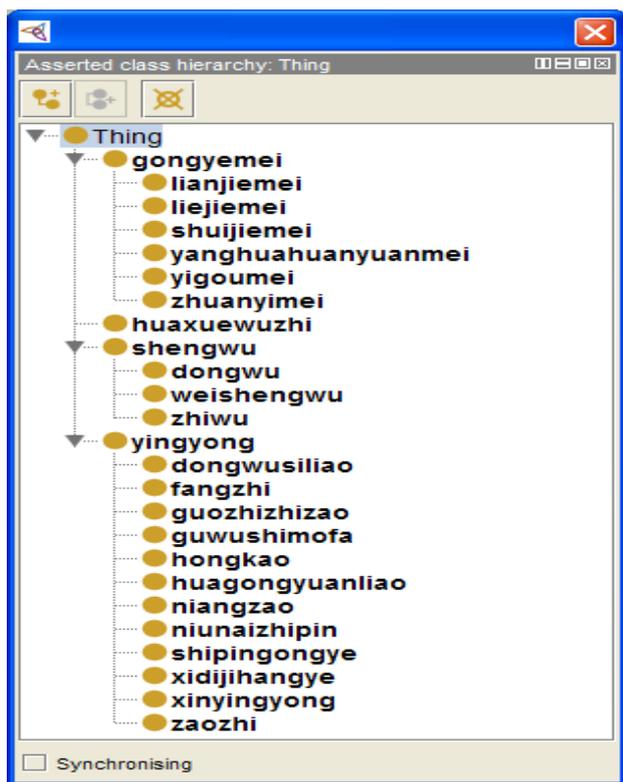


FIGURE 1 The graph of industrial enzyme hierarchical structure

The most important application of industrial enzymes ontology is the realization of data mining technology and the intelligent retrieval and mining of user requirement in literature systems. In this process, the validity and reusability of the industrial enzymes ontology model are also verified as well as the applicability of ontology as the knowledge organization system in the literature retrieval.

Zhang Zhijun, Liu Hong

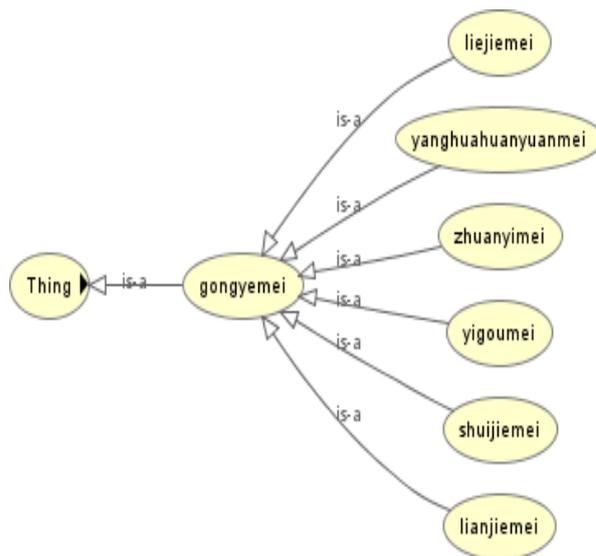


FIGURE 2 The graph of industrial enzyme classification tree

TABLE 1 OWL documents of industrial enzymes

```
<Ontology xmlns="http://www.w3.org/2006/12/owl2-xml#"
  xml:base="http://www.w3.org/2006/12/owl2-xml#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl2xml="http://www.w3.org/2006/12/owl2-xml#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:Ontology="http://www.semanticweb.org/ontologies/2013/11/Ontology.owl#"
  URI="http://www.semanticweb.org/ontologies/2013/11/Ontology.owl">
  <SubClassOf>
    <Class URI="&Ontology;dongwu"/>
    <Class URI="&Ontology;shengwu"/>
  </SubClassOf>
  <Declaration>
    <Class URI="&Ontology;dongwu"/>
  </Declaration>
  <SubClassOf>
    <Class URI="&Ontology;dongwusiliao"/>
    <Class URI="&Ontology;yingyong"/>
  </SubClassOf>
  <Declaration>
    <Class URI="&Ontology;dongwusiliao"/>
  </Declaration>
```

With the help of plot plug-in graphviz-2.28 of Protégé4.0 a class relation diagram is automatically generated, as shown in Figure 1, 2, which shows the class hierarchy of industrial enzymes ontology. Owl: Thing is the superclass of all the classes. The four class of “yingyong”, “shengwu”, “huaxuewuzhi” and “gongyemei” are the subclass of the superclass-thing.

Part of the OWL documents described as Table 1.

3.3 FULL-TEXT RETRIEVAL ALGORITHM

3.3.1 Design of full-text retrieval algorithm

The query of users is often made up of a set of keywords, which are either elements of ontology library such as class, instance, attribute and attribute values, or other kind of common keywords. Therefore, the user’s input must be analysed firstly, which is done by using the analyser component in Figure 4. We use the same analysis algorithm for both indexing and retrieval, which can reach the optimal matching retrieval results. The design of full-text retrieval algorithm is as follows:

Algorithm 1: Full-text retrieval algorithm

Step 1: For each keyword in query, scanning the input keywords;

Step1.1: If the keyword is ontology element, then adding it to the ontology annotations stack  $h_2$ ;

Step1.2: Else if keyword is other kind of keyword, then adding to the ordinary stack  $h_1$ ;

Step 2: For  $h_2$  to create space vector model  $v_2$ ,  $v_2$  is given the higher weight value  $w_2$ ;

Step 3: For  $h_1$  to create space vector model  $v_1$ ,  $v_1$  is given the lower weight value  $w_1$ ;

Step 4: Integrating of  $v_1$  and  $v_2$ , creating user input space vector model  $v$ .

3.3.2 Basis of ranking algorithm design

Ranking of full-text retrieval algorithm is calculated based on the vector space model of information retrieval in Lucene system. For a collection of documents  $D$ , the closer between document  $d$  and query conditions  $q$ , the higher score of document  $d$ . The computation formula is as in Equation (2).

$$rank(q, d) = \sum_{t \in q} \frac{tf_{t,q} \cdot idf_t}{norm_q} \cdot \frac{tf_{t,d} \cdot idf_t}{norm_d} \cdot coord_{q,d} \cdot weight_t, \quad (2)$$

where  $tf_{t,x} = \sqrt{termFrequency\ t, X}$ ,

$$idf_t = 1 + \log \frac{|D|}{documentFrequency(t, D)}$$

$$norm_q = \sqrt{\sum_{t \in q} tf_{t,q} \cdot idf_t^2}, \quad norm_d = \sqrt{|d|}$$

$$coord_{q,d} = \frac{|q \cap d|}{|q|}$$

3.3.3 Similarity measures method

In this section, we introduce some methods for similarity computation, with details of their adaptation to the ontology domain. We firstly exploit the geometrical model provided by concept hierarchies. The taxonomy

graph is shown in Figure 3. Wu and Palmer [20] (W&P) proposed a path-based method, which consider the distance between the concepts.

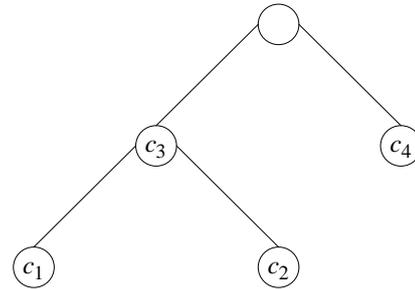


FIGURE 3 The graph of taxonomy

$$sim_{W\&P}(c_1, c_2) = \frac{2 \times N_3}{N_1 + N_2 + N_3}, \quad (3)$$

where  $N_1$  is the number of is-a links from  $c_1$  to its LCS (least common subsumer),  $N_2$  is the number of is-a links from  $c_2$  to its LCS, and  $N_3$  is the number of is-a links from the root of the ontology to the LCS.

Li et al. [21] proposed a similarity method that compounds the depth of the ontology evaluated and the shortest path length in a non-linear fashion.

$$sim_{Li}(c_1, c_2) = e^{-\alpha path(c_1, c_2)} \cdot \frac{e^{\beta h} - e^{-\beta h}}{e^{\beta h} + e^{-\beta h}}, \quad (4)$$

where  $\alpha \geq 0$  and  $\beta \geq 0$  are parameters,  $path(c_1, c_2)$  is the shortest path length between concept  $c_1$  and  $c_2$ , and  $h$  is the minimum depth of the LCS in the hierarchy.

Choi and Kim [22] also proposed a similarity measure method, which is calculated according to the difference on the distance of the shortest path of two concepts and the depth levels between them.

$$sim_{CK}(c_1, c_2) = \frac{MAX\_PATH - path(c_1, c_2)}{MAX\_PATH} \times \frac{MAX\_LEVEL - diff\_level(c_1, c_2)}{MAX\_LEVEL}$$

4 Experimental and data analysis

4.1 DATA SET

To test and verify the effectiveness of the full-text retrieval algorithm based on ontology, we developed medical enzyme semantic literature retrieval system – MESLRS. The purpose is to verify the correctness and reuse of the medical enzyme domain ontology model. MESLRS is constructed on .NET 2003 platform with SQL Server 2003 as the backstage database and deploying in the windows 2003 operating system. We selected 100 pieces of article about medical enzyme from cnki.net as text datasets of the domain, and used 12

categories of corpus in TanCorpV1.0 that is the corpus of Chinese text classification as background data sets.

### 4.2 METRICS

The test in this paper was based on the established medical enzyme ontology library. The evaluation metrics to evaluate the experiment result is *precision*, *recall* and *F*-measure, which are widely used in the field of information retrieval. The bigger of the value of the *precision*, *recall* and *F*-measure, the better of the result. Relevant concepts are shown in Table 2. Evaluation metrics are defined as in Equation (6) (7) (8).

$$Precision = \frac{A}{A+B}, \tag{6}$$

$$Recall = \frac{A}{A+C}, \tag{7}$$

$$F_j = \frac{2}{\frac{1}{R(j)} + \frac{1}{P(j)}}, \tag{8}$$

where  $R_j$  and  $P_j$  are the recall and precision of the document  $j$  respectively.

TABLE 2 Collection of documents measured by precision and recall

Full document collection		
	Retrieved documents	Not retrieved documents
Related documents	A (number of documents being retrieved correctly)	B (number of documents not being retrieve)
Unrelated documents	C (number of documents being retrieved wrongly)	Number of documents refused correctly

### 4.3 EXPERIMENTAL PROCEDURE

#### 4.3.1 Full-text retrieval framework design

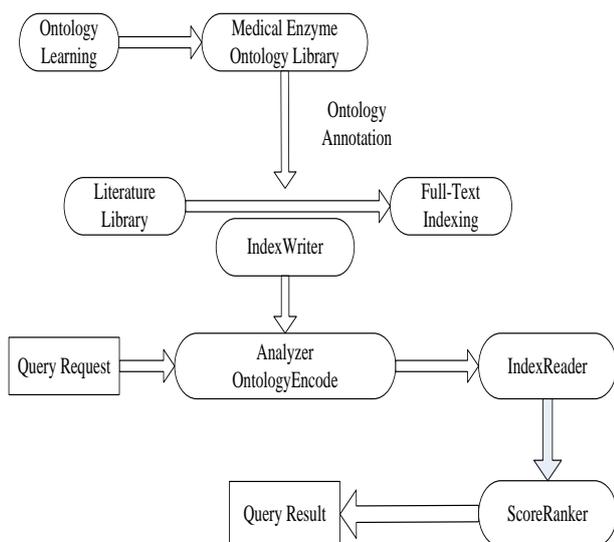


FIGURE 4 The retrieval model based on medical enzyme ontology

The full-text retrieval framework based on medical enzyme ontology model is divided into two parts: full-text index based on ontology and full-text retrieval based on ontology. Full-text index is the basis of full-text retrieval, because the indexed medical literatures are the objects of full-text retrieval. The reason for creating index is due to the vast data of the full-text, therefore in order to improve retrieval efficiency and integrate retrieval algorithm we must first create a full-text index and retrieval the full-text afterwards. In the full-text index mechanism based on ontology knowledge of domain ontology is integrated and indexed content is expanded by using the ontology annotations technology, which can s both recall and precision ratio of the retrieval. Similarly,

the full-text retrieval uses ontology library to analyse query request, which can grasp users' query request more accurately and provide them with a better retrieval ranking result. The full-text retrieval model is shown in Figure 4.

#### 4.3.2 The implementation process of full-text retrieval algorithm

The implementation process of full-text retrieval framework base on medical enzyme ontology is as follows:

After reading a document from medical enzyme ontology library, the analyser component firstly executes annotations pre-treatment to content of the document, which will then deliver the annotated content the IndexWriter component to be indexed. The analyser component analyses the query request, matches its ontology elements with the ontology library, and returns the related literature. The ontologyEncoder component as a sub-component of the analyser component, switches the various elements in ontology library into a more efficient multi-way tree, which can not only be used for ontology annotations of IndexWriter component, but also be the scanning object when the IndexRead component queries the ontology elements. The ScoreRanker component is used to rank the literature of the query and list the ones that most conform to users' query request at the top to facilitate users to find the required documents quickly.

### 4.4 EXPERIMENTAL ANALYSIS

Although precision of traditional literature retrieval based on keyword matching is higher, it paid little attention to meanings and associations of words in different contexts for the reason that it did not consider the concepts that keywords might represent. However, the literature

retrieval algorithm based on ontology executes query expansions based on ontology, which compares the original query keywords submitted by users with the terms in the ontology library and finds the relevant words from ontology, then forms new query vectors so as to use

to full-text retrieval by adding the corresponding context relationship of keywords to the query keywords. So when evaluating the retrieval effect in this study we ought to consider the evaluation of semantic relevance.

TABLE 3 The comparison of experimental results

Semantic Relevancy	Recall (%)		Precision (%)		F-measure (%)	
	General	Ontology	General	Ontology	General	Ontology
0,1	2,73	4,79	99,33	100	0,05	0,09
0,2	4,9	8,87	98,12	95,15	0,09	0,16
0,3	12,47	19,92	99,56	95,14	0,22	0,33
0,4	16,63	35,05	98,01	94,99	0,28	0,51
0,5	21,98	47,36	91,83	94,03	0,35	0,63
0,6	30,59	69,93	89,38	93,33	0,46	0,8
0,7	36,65	81,85	84,95	88,2	0,51	0,85
0,8	53,18	90,25	71,84	76,01	0,61	0,83
0,9	64,18	98,96	58,67	56,93	0,61	0,72

In general query algorithm the recall and precision of the initial query term are calculated and we put the average value of several queries as the query results. Similarly in literature retrieval algorithm based on ontology the recall and precision ratio of the expand query terms are calculated and we put the average value of several queries as the results. Lastly, we calculate F-measure value of the general query algorithm and query algorithm based on ontology. The experimental results are shown in Table 3.

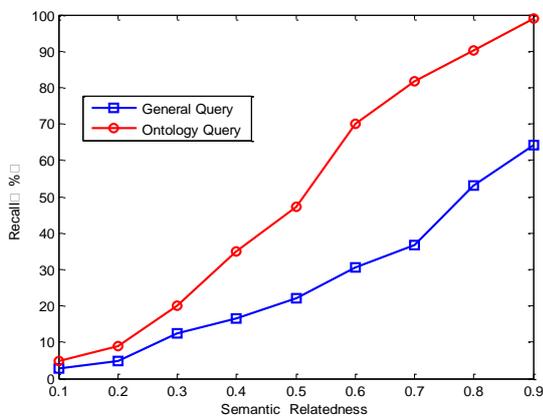


FIGURE 5 Recall comparison

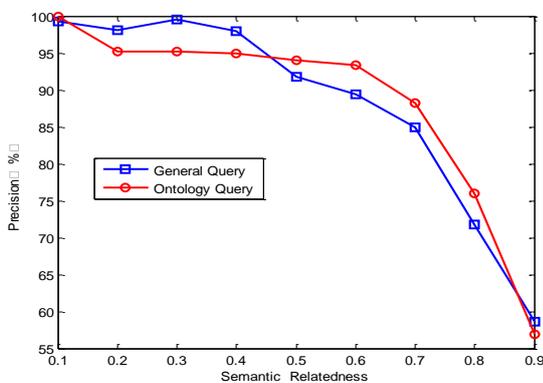


FIGURE 6 Precision comparison

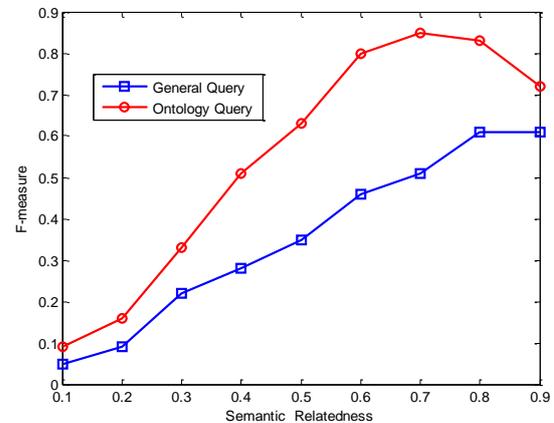


FIGURE 7 F-measure comparison

From Figure 5, 6, 7, we can draw the following conclusions.

Recall and precision has a reciprocal relationship in Figure 5, 6. When recall is high, precision is low. On the contrary, when precision is high, recall is low. A literature retrieval system can be compromised between them. In extreme cases, when retrieval system returns all documents of the system its recall is 100%, but the precision is very low. On the other hand, if the literature retrieval system can just return to the unique document, there will be a very low recall, but its precision could be 100%.

It can be seen from the Figure 6 that as far as precision is concerned the retrieval algorithm based on ontology has no special advantage. However, it can be seen from the Figure 5 that the recall based on ontology query is significantly higher than that of the general query. For example, when semantic relatedness is 0.6 the recall based on ontology query is 39% higher than that of the general query. Thus, the quality of query algorithm based on ontology is higher than general query based on keywords.

From Figure 7 we can find that F-measure of ontology query is higher than that of general query so it shows the advantages of ontology query.

Compared with the general query, the full-text retrieval algorithm based on ontology has the following advantages:

The full-text retrieval algorithm based on ontology can extract implicit knowledge. Using ontology knowledge library to extract potential keyword can retrieve keywords information that the user does not input, however the results might be very useful to users. From Figure 8 we can find that the number of retrieval document is less than 100 by checking with general retrieval containing keywords "gongyemei", but more than 900 documents can be retrieved by an ontology semantic extension.

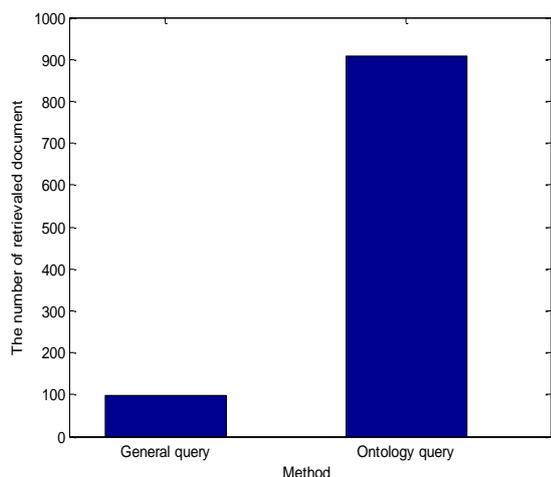


FIGURE 8 The number of retrieval document comparison with keywords "gongyemei"

The full-text retrieval algorithm based on ontology has intelligent query function. In the case of laccase we query the literature about functions of laccase. In medical enzyme ontology knowledge library, keywords and concepts are one-to-one correspondence, which guarantees that the query uses concepts instead of text. When the literature retrieval system based on ontology gets the keywords "qimei" and "gongneng" it firstly queries values of attribute "gongneng" of instances "qimei" in ontology knowledge library, with the returning values including the pulp bleaching, papermaking, hair dye, the effect of lignin, industrial wastewater treatment and so on. From Figure 9 we can find that the query result is 312 literatures on the application of laccase and the most appropriate literature is at the top by applying the ranking mechanism. On the contrary, the general query only find out less than 80 related literature when entering keyword "qimeidegongneng". There is no way of finding the literature corresponding to subordinate concept of "qimei" and "gongneng" no matter what kind of query retrieval strategies are used.

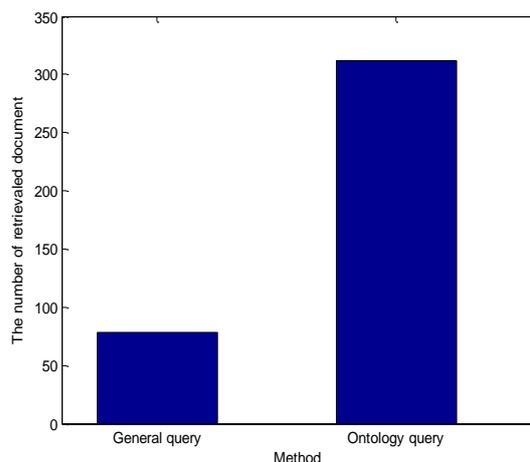


FIGURE 9 The number The number of retrieval document comparison with keywords "qimei"

### 5 Conclusions

This paper studies the construction method of domain ontology. We have constructed a reusable, shared, extensible, relatively perfect and practical medical enzyme ontology by using ontology to integrate and standardized medical enzyme knowledge system. The modification, query and storage of ontology library are carried out based on some applications of semantic web supported by Jena, and the full-text retrieval algorithm based on ontology is put forward. We have designed and implemented a literature retrieval system based on medical enzyme semantic, which have realized intelligent ontology learning about medical enzyme literature knowledge, as well as intelligent literature retrieval. In the premise of precision ratio, recall ration have significantly increased, what's more, F-measure of ontology query is higher than that of general query. It provides intelligent basis for medical enzyme literature related expert system, information retrieval, the education system, the natural language understanding and so on. This system is of strong reusability, because it cannot only focus on a particular area, but also be used in different semantic retrieval areas. As long as the corresponding domain ontology is changed the system can be used for information retrieval of the new domain. The efficiency of information retrieval is improved by using the advantage of ontology semantic expression.

In this article, we have just extracted part of the enzyme domain knowledge to construct the medical ontology model. Currently formalization of medical enzyme ontology is relatively rare and at the same time, there is no visual reasoning system, and the reasoning function of reasoning layer also needs to be strengthened. Therefore, it is very necessary for us to go on expanding and constructing enzyme ontology model, which is also our further research direction about ontology.

## Acknowledgments

This paper is supported by the National Natural Science Foundation of China (No. 61272094), Natural Science Foundation of Shandong Province (ZR2010QL01,

ZR2012GQ010), A Project of Shandong Province Higher Educational Science and Technology Program (J12LN31, J13LN11), Jinan Higher Educational Innovation Plan (201303001) and Shandong Provincial Key Laboratory Project.

## References

- [1] Berners-Lee T, Hendler J, Lassila O 2001 The semantic web *Scientific american* **284**(5) 28-37
- [2] Spanos D E, Stavrou P, Mitrou N 2012 Bringing relational databases into the semantic web: A survey *Semantic Web* **3**(2) 169-209
- [3] Tsakonas G, Mitrelis A, Papachristopoulos L 2013 An exploration of the digital library evaluation literature based on an ontological representation *Journal of the American Society for Information Science and Technology* **64**(9) 1914-26
- [4] Jimenez-Castellanos A, Fernandez I, Perez-Rey D 2013 Biomedical Literature Retrieval Based on Patient Information *Biomedical Engineering Systems and Technologies* Springer Berlin Heidelberg 312-23
- [5] Lee C S, Jiang C C, Hsieh T C 2006 A genetic fuzzy agent using ontology model for meeting scheduling system *Information Sciences* **176**(9) 1131-55
- [6] Voorhees E M 1994 Query expansion using lexical-semantic relations *SIGIR '94* Springer London 61-9
- [7] Guarino N, Masolo C, Vetere G 1999 Ontoseek: Content-based access to the web *Intelligent Systems and Their Applications, IEEE* **14**(3) 70-80
- [8] Aleksovski Z, Klein M, Ten Kate W, et al 2006 Matching unstructured vocabularies using a background ontology *Managing Knowledge in a World of Networks*. Springer Berlin Heidelberg 182-97
- [9] Maki W, McKinley L, Thompson A 2004 Semantic distance norms computer from an electronic dictionary (wordnet) *Behaviour Research Methods, Instruments & Computers* **36** 421-31
- [10] Navigli R, Velardi P 2003 An analysis of ontology-based query expansion strategies *Proceedings of the 14th European Conference on Machine Learning, Workshop on Adaptive Text Extraction and Mining, Cavtat-Dubrovnik, Croatia* 42-9
- [11] Abasolo J M, MELISA G M 2000 An ontology based agent for information retrieval in medicine *Proceedings of the First International Workshop on the Semantic Web Lisbon, Portugal* 73-82
- [12] Wu C G, Jiao W P, Tian Q J 2001 An information retrieval server based on ontology and multi-agent *Journal of Computer Research and Development* **38**(6) 641-7
- [13] Gruber T R 1993 A translation approach to portable ontology specifications *Knowledge acquisition* **5**(2) 199-220
- [14] Wang H, Sun R Z 2010 Research of semantic retrieval system based on domain-ontology and lucene *Journal of Computer Application* **30**(6) 1655-57
- [15] Resnik P 2011 Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language *arXiv preprint arXiv 1105.5444*
- [16] Zhong X Q, Fu H G, Yu L 2010 Geometry knowledge acquisition and representation on ontology *Chinese Journal of Computers* **33**(1) 167-74
- [17] Bodenreider O 2004 The unified medical language system (UMLS): integrating biomedical terminology *Nucleic acids research* **32** 267-70
- [18] Zhou X B, Cao C G 2003 Medical Knowledge Acquisition: An Ontology-Based Approach *Computer Science* **30**(10) 35-9
- [19] Du X Y, Li M, Wang S 2006 A survey on ontology learning research *Journal of Software* **17**(9) 1837-47
- [20] Batet M, Sánchez D, Valls A 2011 An ontology-based measure to compute semantic similarity in biomedicine *Journal of biomedical informatics* **44**(1) 118-25
- [21] Li Y, Bandar Z A, McLean D 2003 An approach for measuring semantic similarity between words using multiple information sources *Knowledge and Data Engineering, IEEE Transactions on* **15**(4) 871-82
- [22] Choi I, Kim M 2003 Topic distillation using hierarchy concept tree. *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval* ACM 371-2

Authors	
	<p><b>Zhijun Zhang, born in 1973, in Shandong, China</b></p> <p><b>Current position, grades:</b> Computer software and theory Specialty, Associate professor.</p> <p><b>University studies:</b> He received his M. S. degree in School of Computer Science and Technology, Shandong University in 2006. Currently, he is a Ph.D. candidate in School of Information Science and Engineering, Shandong Normal University.</p> <p><b>Scientific interest:</b> information retrieval and recommender systems.</p> <p><b>Publications:</b> having issued more than twenty academic dissertations.</p>
	<p><b>Hong Liu, born in 1955, in Shandong, China</b></p> <p><b>Current position, grades:</b> Professor. Doctoral supervisor. She is currently the dean of School of Information Science and Engineering, Shandong Normal University, Jinan, China. She is also the director of Shandong Provincial Key Laboratory for Novel Distributed Computer Software Technology. <b>Publications:</b> over 100 articles to professional journals.</p> <p><b>Scientific interests:</b> Distributed Artificial Intelligence and Computer Aided Design.</p>

# Case-based reasoning intelligent prediction model of rotary kiln temperature

Gongfa Li<sup>1, 2\*</sup>, Jia Liu<sup>1</sup>, Guozhang Jiang<sup>1</sup>, Honghai Liu<sup>2</sup>, Wentao Xiao<sup>1</sup>

<sup>1</sup>College of Machinery and Automation, Wuhan University of Science and Technology, Wuhan 430081 China

<sup>2</sup>Intelligent Systems and Biomedical Robotics Group, School of Creative Technologies, University of Portsmouth, PO1 2DJ, United Kingdom

Received 1 March 2014, www.tsi.lv

## Abstract

Temperature is a key technical index in rotary kiln combustion process, which is so difficult to measure directly online. The offline analysis has large-time delay and poor precision. An intelligent prediction model of rotary kiln temperature based on case-based reasoning was developed, which consists of four modules: data collection and pre-treatment, prediction, online modification and effect estimate. The practical data of some rotary kiln were simulated. The industrial application results show that the prediction model can reflect the actual operation condition and meet the requirement of real-time control. Its effectiveness is proved evidently.

**Keywords:** case-based reasoning, intelligent prediction, temperature control, rotary kiln

## 1 Introduction

Rotary kiln is being widely used in many industrial departments. But the biggest shortcoming of it is its high energy consumption and low thermal efficiency. The backward method to test and control it is the main cause. At present, the estimation of its thermal state is still dependent on the fire workers keeping observing the "ring of fire" of the rotary kiln. The workers' mental state, technological literacy, responsibility and many other factors would affect them. The large randomness make it hard for the rotary kiln to save energy. Besides, the rotary kiln is a typical multivariable, time varying and distributed parameters nonlinear system. The thermo technical process is so complex that it's very difficult to build a mathematical model for it. Using soft-sensing technique to online test the temperature of rotary kiln is of great significance to control the combustion process of rotary kiln. At present, as an effective way to estimate the uncertain variables of industrial process, soft-sensing technique is being more and more widely used. It is mainly aimed at building mathematical model for process variables. It can be named prediction model according to its characteristic and function.

## 2 Modelling method based on case-based reasoning technology

CBR is a methodology using past experience to simulate human brain judging things. It expresses and stores a large number of problems and their solutions in the form of case. When meet a new problem (case), the system will match similar cases from its case library and retrieve the most similar one, then adjust their solutions to solve

it. The new case with high typicality will be stored. And in that way, the CBR system is improved.

The Figure 1 shows an intelligent prediction model structure. It can forecast key variables of a complex industrial process.  $\hat{X}$  is the output of the case-based reasoning prediction module,  $\bar{X}$  is the correction output,  $\Sigma$  is the process data set from the distributed control system (DCS),  $\Theta$  is the artificial measurement data set obtained by the measure model,  $e$  is the online correction parameter from the online correction module,  $u$  is the control input,  $y$  is the output of the controlled object.

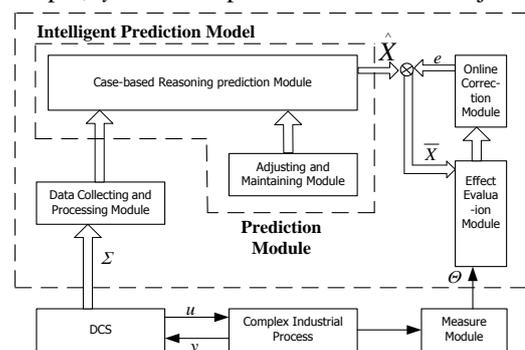


FIGURE 1 The structure of the intelligent prediction model

### 2.1 DATA COLLECTING AND PROCESSING MODULE

Data collected from the working site always accompanied with various kinds of interference noise. To furthest avert them, we need converse the data and deal with the errors. Under some conditions, we need also deal with the output of the prediction model appropriately. As the model is built on the premise of a series of hypotheses, it can't be

\* Corresponding author e-mail: ligongfa@wust.edu.cn

all the fours with the practical situation. There are some model errors.

2.2 PREDICTION MODULE

Firstly, prediction model reads the current working condition and retrieves similar cases from the case library. Then the retrieved cases will be matched and reused according to their similarity threshold to get the solution, which is the dominant variable soft measurement value that needs to be estimated. Analyse the error between the actual measured value and the soft measured value, assess the precision of the soft measuring model precision. If a case cannot reach the prospective accuracy, adjust it. Else store it into the case library by corresponding rules.

When the case-based reasoning module is working, as the object’s technological parameters and other condition changing, the original useable cases may not be appropriate any more. In order to make sure the case-based reasoning module can get the object’s changing information and obtain the right result, the case-based reasoning system need to adjust and maintain. That is the function of the adjusting and maintaining module.

2.3 ONLINE CORRECTION MODULE

After the prediction module come into using, the output of the module may drift if the object’s situation and working location changed. To ensure the prediction value’s veracity, it needs correction.

$$e = \frac{1}{n} \sum_{i=1}^n (X_i^* - \hat{X}_i), \tag{1}$$

$e$  is the online correction parameter,  $\hat{X}_i$  is the output of the prediction model,  $X_i^*$  is the actual measured value,  $n$  is the sample size. The corrected output is

$$\bar{X} = \hat{X} + e. \tag{2}$$

This method is easy to realize. We can adjust the output of the prediction module with type (1) and type (2) to facilitate the output result drifting and ensure the accuracy.

2.4 EFFECT EVALUATION MODULE

This module compares the output of the case-based reasoning prediction module with the artificial measured data from the measure module to evaluate the prediction accuracy.

TABLE 1 The case of rotary kiln temperature

Time	Working Condition			Solution
	Heating Gas Flow	Calorific Value of Gas	Heating Gas Pressure	Rotary Kiln Temperature Prediction Value
			Combustion Chamber Draft	

3 Rotary kiln temperature intelligent prediction model based on case-based reasoning

When building a temperature prediction model of rotary kiln, the temperature of rotary kiln should be analysed in consideration of the periodically statistics of rotary kiln temperature, the combustion chamber draft, the heating gas flow, the heating gas pressure and the calorific value of gas. Intelligent prediction will be operated based on case-based reasoning technology. Figure 2 is the model structure. The rotary kiln temperature intelligent prediction model consists of case-based reasoning prediction model, self-adjusting model and so on. There is the measure data set of rotary kiln combustion process. There is the output of the case-based reasoning prediction model, There is the measure value of rotary kiln during a time interval. There is the statistic of the artificial measured temperature values. There is the error between the output of the case-based reasoning prediction model and the artificial measured temperature value, ( $>0$ ) there is the presupposed error limitation. There is the adjusted output of rotary kiln temperature value. After the model selector obtain the test data from rotary kiln combustion process, the case-based reasoning prediction model will accomplish the prediction and get the value. The result will be adjusted according to the artificial testing temperature statistic obtained by the self-adjusting model. Then we can get the desired value.

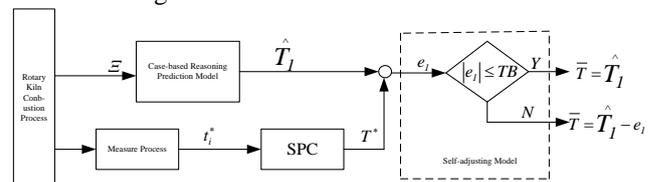


FIGURE 2 The complex intelligent prediction model of rotary kiln

3.1 CASE-BASED REASONING ALGORITHM

3.1.1 Presentation of the case

The prediction model of rotary kiln temperature is based on case-based reasoning prediction algorithm. Analyse the operating parameters with visible controllable analysis and use correlation analysis to compare them with other variables. In consideration of variable simplification, choose the parameters below as the auxiliary variables of the soft-sensing model of rotary kiln temperature: the heating gas flow  $u$ , the calorific value of gas  $h$ , the heating gas pressure  $p$  and the combustion chamber draft  $n$ . The cases of rotary kiln are stored into computer in form of data base. The data base is composed of several case records, presented as the Table 1.

3.1.2 Case retrieval

Case retrieval is of great importance to case-based reasoning. Retrieve the case library according to the new case and find the best solution. If the current working state is described as  $X = (x_1, x_2, \dots, x_n)$ , but the case in the case library is  $X_k = (x_{1,k}, x_{2,k}, \dots, x_{n,k})$ ,  $k = 1, \dots, m$ ,  $m$  is the number of the cases. The similarity between  $x_i$  ( $1 \leq i \leq n$ ) and  $x_{i,k}$  can be defined as:

$$Sim(x_i, x_{i,k}) = 1 - |x_i - x_{i,k}| / Max(x_i, x_{i,k}). \tag{3}$$

And the similarity between current working state  $C_c$  and the existing case  $C_k$  is:

$$Sim(C_r, C_k) = \sum_{i=1}^n w_i Sim(x_i, x_{i,k}). \tag{4}$$

The  $w_i$  in Equation (4) is characteristic weight parameter,  $\sum_{i=1}^n w_i = 1$ .

Then write the similarity value into corresponding case library.

Assume the similarity threshold is  $Sim_{max} = max(sim(c_r, c_k))$ .

$$Sim_v = \begin{cases} J_v, & Sim_{max} \geq J_v \\ Sim_{max}, & Sim_{max} < J_v \end{cases} \tag{5}$$

The threshold  $J_v$  is confirmed by engineers from the practical situation.

If similarity value between a retrieved case with the practical situation is  $\geq Sim_v$ , then the case is a matched case.

3.1.3 Case retrieval and matching

Case retrieval and matching is the key to case-based reasoning. Its main purpose is retrieving cases according to the description of new problems and finding out their solutions. Any case whose similarity value with the current practical situation is over the threshold  $Sim_v$  will be retrieved as matched case.

3.1.4 Case reusing

In general, the solutions of the retrieved matched cases can't be directly used as the solution of the current working situation. Assume the retrieved matched case set is  $C_k = \{T_k, X_k, Y_k, Sim_k\}$ ,  $k = 1, 2, \dots, l$ ,  $l < m$ .  $k$  is the

number of matched cases,  $Sim_k$  is the similarity value between the case set  $C_k$  and the working situation.

Suppose  $\tilde{C}$  as a case set with biggest similarity value  $Sim_{max}$  and the solution as  $\tilde{J}$ .  $J_u$  is calculated solution reused according to the case set.

$$J_u = \sum_{k=1}^l (Sim_k \times Y_k) / \sum_{k=1}^l Sim_k. \tag{6}$$

The variables above are characterizing attributes. Their feature weights are determined to be 0.25, 0.25, 0.25, 0.25 according to expert experience. Based on the past cases from the temperature case library, the rotary kiln temperature can be predicted with case-based reasoning method.

3.2 SELF-ADJUSTING ALGORITHM

To ensure prediction accuracy, the initial predicted result needs to be self-adjusted. Suppose the artificial measuring data set as  $\{t_i^*, i = 1, 2, \dots, k\}$ . The data can be conducted with statistical process control (SPC) method.

$$T^* = \frac{\sum_{i=1}^k t_i^*}{k}. \tag{7}$$

$\hat{T}_1$  is the initial rotary temperature value obtained by case-based reasoning prediction model. The temperature prediction effect is evaluated with the type below.

$$e_i = \hat{T}_1 - T^*, \tag{8}$$

$e_i$  is the error between the initial rotary temperature value obtained by case-based reasoning prediction model and the artificial test data. If  $|e_i| > TB$ , it means the output of the prediction model needs to be adjusted,  $e_i$  is the adjusting parameter. Else, the output does not need to be adjusted and  $e_i$  can be supposed to be 0. The adjusted output  $\bar{T}$  is:

$$\bar{T} = \hat{T}_1 - e_i. \tag{9}$$

4 Industrial application

Apply the intelligent prediction model of rotary kiln into an iron and steel complex's mineral processing intelligent control system. The temperature prediction is shown in Figure 3. Statistic suggests, if the odds, which the rotary kiln temperature prediction errors stay within  $\pm 10^\circ C$  can reach 91.8%, it will meet the industrial production requirement.

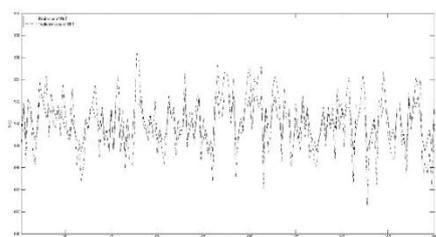


FIGURE 3 The prediction effective of the rotary kiln

## 5 Conclusions

The key parameters of a complex industrial production process can't be measured directly online. The intelligent prediction model based on CBR is aimed at solving the problem. In consideration of the rotary kiln temperature cannot be measured in real time, an intelligent prediction

## References

- [1] Jianyi Kong, Gongfa Li, Hegen Xiong 2007 Research on Soft-sensing Modeling Methods and its Application in Industrial Production Machine Tool and Hydraulics *35*(6) 149-51
- [2] Yoo C K, Lee I B 2004 Soft Sensor and Adaptive Model-based Dissolved Oxygen Control for Biological Wastewater Treatment Processes *Environment Engineering Science* **21**(3) 331-40
- [3] Fortuna L, Graziani S, Xibulia M G 2005 Soft Sensors for Product Quality Monitoring in Debutanizer Distillation Columns *Control Engineering Practice* **13**(2) 2499-508
- [4] Armaghan Negar, Renaud Jean 2012 An application of multi-criteria decision aids models for Case-Based Reasoning *Information Sciences* 210 55-66
- [5] Jain Pooja, Dahiya Deepak 2012 Knowledgeable multi agent system for ecommerce (KMASE) using case based reasoning and knowledge beads *International Journal of Computer Science Issues* **9**(2) 395-403
- [6] Dendani-Hadiby, Nadjette Khadir, Mohamed Tarek 2012 A case based reasoning system based on domain ontology for fault diagnosis of steam turbines *International Journal of Hybrid Information Technology* **5**(3) 89-104
- [7] Jagannathan R, Petrovic S, McKenna A, Newton L 2012 A Novel two phase retrieval mechanism for a clinical case based reasoning system for radiotherapy treatment planning *International Journal on Artificial Intelligence Tools* **21**(4) 1121-29
- [8] Chang Pei-Chann, Lin Jyun-Jie Dzan, Wei-Yuan 2012 Forecasting of manufacturing cost in mobile phone products by case-based reasoning and artificial neural network models *Journal of Intelligent Manufacturing* **23**(3) 517-31
- [9] Khan Malik, Jahan Awais, Mian Muhammad, Shamil Shafay, Irfan Awan 2011 An empirical study of modeling self-management capabilities in autonomic systems using case-based reasoning *Simulation Modelling Practice and Theory* **19**(10) 2256-75
- [10] Gongfa Li, Jianyi Kong, Guozhang Jiang 2012 Air-fuel Ratio Intelligent Control in Coke Oven Combustion Process *Information* **15**(11) 4487-94

Authors	
	<p><b>Gongfa Li, born in 1979, Hubei, China</b></p> <p><b>Current position, grades:</b> Associate professor, college of Machinery and Automation, Wuhan University of Science and Technology.  <b>University studies:</b> Ph.D. degree in mechanical design and theory from Wuhan University of Science and Technology in China.  <b>Scientific interest:</b> modelling and optimal control of complex industrial process.  <b>Publications:</b> nearly twenty papers in related journals.</p>
	<p><b>Jia Liu, born in 1990, Hubei, China</b></p> <p><b>Current position, grades:</b> student M.S. degree in mechanical design and theory at Wuhan University of Science and Technology  <b>University studies:</b> B.S. degree in mechanical engineering and automation from Wuchang institute of Technology, Wuhan, China, 2012  <b>Scientific interest:</b> mechanical CAD/CAE, signal analysis and processing</p>
	<p><b>Guozhang Jiang, born on December 15, 1965, Tianmen, China</b></p> <p><b>Current position, grades:</b> Professor of Industrial Engineering, and the Assistant Dean of the college of machinery and automation, Wuhan University of Science and Technology.  <b>University studies:</b> the Ph.D. degree in mechanical design and theory from Wuhan University of Science and Technology, China, 2007.  <b>Scientific interest:</b> computer aided engineering, mechanical CAD/CAE and industrial engineering and management system.  <b>Publications:</b> 120.</p>
	<p><b>Honghai Liu, born in 1973, China</b></p> <p><b>Current position, grades:</b> Professor in Intelligent Systems, Head of Intelligent Systems and Biomedical Robotics, University of Portsmouth.  <b>University studies:</b> PhD in Intelligent Robotics in 2003 from Kings College, University of London, UK.  <b>Scientific interest:</b> approximate computation, pattern recognition, multi-sensor based information fusion and analytics, human machine systems, advanced control, intelligent robotics and their practical applications.  <b>Publications:</b> 300.  <b>Experience:</b> research appointments at King's College London, University of Aberdeen, and project leader appointments in large-scale industrial control and system integration industry</p>
	<p><b>Wentao Xiao, born in 1989, Hubei, China</b></p> <p><b>Current position, grades:</b> student M.S in mechanical design and theory at Wuhan University of Science and Technology  <b>University studies:</b> B.S. degree in mechanical engineering and automation from City College of Wuhan University of Science and Technology, Wuhan, China, in 2013  <b>Scientific interest:</b> mechanical CAD/CAE, signal analysis and processing.</p>

# Study on subsynchronous resonance problem in series-compensated transmission system

Chengbing He\*, Dongchao Chen

School of energy, power and mechanical engineering, North China Electric Power University, Beijing 102206, China

Received 10 May 2014, www.tsi.lv

## Abstract

The series capacitor compensation technique can improve the transmission capability of transmission system; however, it may cause subsynchronous resonance (SSR) which will seriously influence the safe operation of large turbo-generator shafts. Aiming at a series capacitor compensation transmission system, the transient torques of SSR is simulated in different compensation degrees and unit load levels. Based on the modified segmentation Prony algorithm (MSPA) proposed in this paper, the characteristic frequencies and damping parameters are identified and analysed. The MSPA, complex Morlet wavelet and time-frequency contour map are integrated and used in analysis of transient torques of SSR in time-frequency field. The results of study show that the transient torques develop towards divergence direction and the frequencies of SSR (converted to the rotor side) in negative damping area move towards into low frequency area with the increasing of the compensation degree. With the increasing of the power output, the torque values of each shaft section will increase, but the dominant oscillation mode will not change. The increasing of power output can make the damping effect slightly change and the trend of the change develops towards the direction that is not conducive to restrain SSR.

*Keywords:* series-compensated transmission, subsynchronous resonance, modified segmentation Prony algorithm, complex Morlet wavelet, time-frequency contour map

## 1 Introduction

Series compensation is an effective and economical mean to enhance the transmission capability and improve the stability of long transmission systems, but it may cause subsynchronous resonance (SSR). The electrical-mechanical resonance phenomenon would seriously affect the safety of large turbo-generator shafts and the stability of the power system. The SSR analysis methods mainly include eigenvalue analysis, complex torque coefficient approach, time domain simulation method, Prony algorithm (PA), and so on.

PA can obtain the characteristic parameters of sampling signal easily and is applicable for SSR analysis. Most of the existing PAs can only be applied to part of sampling signal every time, so the results are random and inaccurate [1-5]. An improved PA [6,7], which is called segmentation PA (SPA) in this paper, enhances the accuracy of the identified results by dividing the whole observation window into several segments and comparing the results of each segment obtained by PA to determine the optimal results, but it reduces the computational efficiency. More important, it is difficult to deal with and choose the results among different segments. Reference [8] reports a moving window PA (MWPA) which has stronger adaptability to noise and can be applied to long time range signals. However, it is raised on the basis of the modified PA (MPA) which uses singular value decomposition (SVD) to determine the rank of Prony model. Experimental results of this paper and reference [9]

demonstrate that the order of the model acquired by means of SVD is low, so some concerned oscillation modes sometimes cannot be obtained.

In view of the problems mentioned above, a modified SPA (MSPA) based on traditional PA (TPA) and SPA is proposed in the paper. The MSPA, complex Morlet wavelet and time-frequency contour map are integrated and used in analysis of transient torques of SSR in time-frequency field.

## 2 Modified segmentation Prony algorithm

PA is a method fitting a linear combination of exponential terms to a signal ( $N \geq 2p$ ). Supposing that a signal is composed by  $p$  exponential functions, the mathematical model in discrete time function form is:

$$x_n = \sum_{k=1}^p A_k \exp(j\theta_k) \exp[(\sigma_k + j2\pi f_k)n\Delta t], \quad (1)$$

where  $n=0,1,\dots,N-1$ ;  $k=1,2,\dots,p$ ;  $p$  is the order;  $\Delta t$  is the sampling time-interval;  $A_k$ ,  $f_k$ ,  $\sigma_k$  and  $\theta_k$  are magnitude, frequency, attenuation factor and primary phase.

Meanwhile, the expressions of residues and pole are defined as follows:

$$b_k = A_k \exp(j\theta_k) \quad z_k = \exp[(\sigma_k + j2\pi f_k)\Delta t] \quad (2)$$

\* Corresponding author e-mail: hcbyy@126.com

Assuming that  $N$  is the number of data points in the observation window, while the number of sampling data within each segment is same as  $m$ , and the identification order of each segment is same as  $p$  ( $p$  is great enough to ensure the identification accuracy for each segment). To ensure that it is completely applicable to effective signal, there must be several repetitive data points whose number is  $m-p$  between the adjacent sections. The serial number of each segment is  $i$  and the total number of the segments is  $j$  which can be calculated by the following formula.

$$j = [(N - m) / p]_- + 1, \tag{3}$$

where the operator  $[ ]_-$  means negative round number.

For the  $i^{\text{th}}$  segment similar with the derivation process of TPA, the forward differential equation can be obtained and expressed as:

$$\sum_{h=0}^p a_h^i x_{n+h}^i = 0, \tag{4}$$

where  $x_n^i$  are the data points in the  $i^{\text{th}}$  segment,  $i=1,2,\dots,j$ ;  $n=0,1,\dots,m-p-1$ ;  $a_p^i=1$ . Sum the two sides of equations shown as the above equation in all of the segments, and then the following formula can be obtained.

$$\sum_{i=1}^j \sum_{h=0}^p a_h^i x_{n+h}^i = 0. \tag{5}$$

Because  $a_h^i$  ( $i=1,2,\dots,j$ ;  $h=0,1,\dots,p$ ) should remain unchanged for each segment, let them be equal to  $a_h$ . Equation (5) can be transformed into Equation (6).

$$\frac{1}{j} \sum_{h=0}^p a_h \sum_{i=1}^j x_{n+h}^i = 0. \tag{6}$$

The Equation (7) can be constructed by Equation (6).

$$\begin{bmatrix} X(p) \\ X(p+1) \\ \vdots \\ X(m-1) \end{bmatrix} = \begin{bmatrix} X(0) & X(1) & \dots & X(p-1) \\ X(1) & X(2) & \dots & X(p) \\ \vdots & \vdots & \vdots & \vdots \\ X(m-p-1) & X(m-p) & \dots & X(m-2) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{p-1} \end{bmatrix}, \tag{7}$$

where  $X(n)=(x_n^1+x_n^2+\dots+x_n^j)/j$ ;  $n=0,1,2,\dots,m-1$ . As shown by Equation (7), the overall sample function matrix is constructed and the average error effect of the whole observation window is taken into account in the MSPA, which are different from the exiting ones.

Using singular value decomposition and total least square method (SVD-TLS) to solve the coefficients  $a_h, z_k$  can be obtained by solving the characteristic polynomial which is presented as Equation (8).

$$Z^p + a_{p-1}Z^{p-1} + \dots + a_1Z + a_0 = 0 \tag{8}$$

The MSPA residues can be computed by SVD-TLS according to the following formula.

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_p \\ \vdots & \vdots & \vdots & \vdots \\ z_1^{m-1} & z_2^{m-1} & \dots & z_p^{m-1} \end{bmatrix} \begin{bmatrix} b_1^1 \\ b_2^1 \\ \vdots \\ b_p^1 \end{bmatrix} = \begin{bmatrix} x_0^1 \\ x_1^1 \\ \vdots \\ x_{m-1}^1 \end{bmatrix}. \tag{9}$$

In term of  $z_k$  and  $b_k^1$  (residues in the first segment), magnitude, primary phase, damping coefficient and frequency can be obtained on basis of the following formula.

$$\begin{cases} A_k = |b_k^1| \\ \theta_k = \arctan [\text{Im}(b_k^1) / \text{Re}(b_k^1)] \\ \sigma_k = \ln|z_k| / \Delta t \\ f_k = \frac{\arctan [\text{Im}(z_k) / \text{Re}(z_k)]}{2\pi\Delta t} \end{cases}. \tag{10}$$

The resonance characteristic of the power system is negative damped oscillation when  $\sigma_k$  is positive, and vice versa.

According to the sampling theorem, the sampling frequency 1000 Hz is adopted in the analysis of SSR. The time length of each segment is set to 2 seconds and the time length of the window 4~5 seconds is adopted and the failure period should be avoided. With regard to the rank  $p$ , the rank 100 is temporarily adopted here. Lots of interference components will emerge when  $p$  is assigned to a big value directly. The sampling signal is analysed by KW-FFT method in order to get the frequencies ( $F_L, L=1,2,\dots$ ) of the peak points, then the oscillation modes can be preliminary screened by Equation (11).

$$|f_k - F_L| \leq C, \tag{11}$$

where  $C$  is a constant, which could be 0.5~1Hz according to practical needs.

By means of making changes to the energy class, the energy class of SSO modes is defined as:

$$E_k = A_k^2 \sum_{s=1}^m (e^{\sigma_k s \Delta t})^2 = \frac{A_k^2 (e^{2m\Delta t \sigma_k} - 1)}{1 - e^{-2\Delta t \sigma_k}}. \tag{12}$$

Both magnitude and attenuation factor are taken into account in Equation (12). Near each  $F_L$ , a group of  $f_k$  can be obtained on basis of Equation (11). Their energy class can be calculated by Equation (12). And then, the mode whose energy class is the biggest among the group can be found out. It is just one of the oscillation modes.

Simulation results based on the FBM show that the MSPA presented here is suitable for the analysis of SSO and superior to the traditional ones to some extent [10].

### 3 Complex morlet wavelet and time-frequency contour map

The traditional FFT method is not suit for non-stationary signal analysis. The complex Morlet wavelet adopts

Gauss function, which time-frequency window area is smallest, whose localization properties of time-frequency domain are good, and the symmetry is better. The transform result can reflect both the signal amplitude and phase relationship [11, 12]. The introduction of the complex Morlet wavelet transform can provide a new means to torsional vibration monitoring and analysis.

The time-domain expression of complex Morlet wavelet can be written as:

$$\psi(t) = \frac{1}{\sqrt{\pi f_b}} \exp(j2\pi f_c t) \exp(-t^2 / f_b), \quad (13)$$

where,  $f_b$  is bandwidth parameter.  $f_c$  is centre frequency. The values of  $f_b$  and  $f_c$  can be adjusted according to the time-frequency resolution need in practice. The analysis result of the complex Morlet wavelet can be represented by time-frequency contour map. After the specific form of Morlet wavelet is determined, observation of a scales wavelet transform is equivalent to that of time-frequency contour map. Frequency-domain information and time-domain information are contained in time-frequency contour map that can visually not only represent the original signal energy distribution but also show the relationship between each scale wavelet coefficient [13].

**4 Simulations and results**

The study system is according to the conversion design of the second benchmark model for computer simulation of SSR, whose wiring diagram and parameters are given in reference [14]. The generator and shaft are replaced. The parameters of the turbo-generator come from a 600MW unit, whose shaft is four spring-mass models, including HIP, LPA, LPB and GEN. The system exhibits three torsional modes at frequencies 12.512 Hz, 20.755 Hz, and 25.646 Hz. The auxiliary power is 4% of generating unit rating.

**4.1 ANALYSIS OF DIFFERENT SERIES COMPENSATION DEGREES**

The simulation conditions are: generator output of  $P_e=0.9$  p.u., terminal voltage of  $V_t=1.0$  p.u.. It is assumed that the system is experienced by an instantaneous single phase short-circuit from 1.5 s to 1.6 s in the middle of the line which contains series compensation. The simulation lasts 10 s. Series compensation degrees (described by  $X_C$ ) are respectively set as 30%, 50% and 70%.

The torque response curves in time-domain between LPB and GEN and their time-frequency contour maps in different compensation degrees are given in Figure 1, Figure 2 and Figure 3. Table1 shows the maximum torque response values in each section within ten seconds when compensation degrees are different. It is clearly indicated that the torque does not diverge when compensation degree is 30%. However, the torque curves become divergent when compensation degrees are 50%

and 70%. Based on the time-frequency contour maps, it should be noted that the former three modes are excited. The second order natural frequency of torsional vibration is the dominant oscillation mode when compensation degrees are 30% and 50% (the curves in the figures exhibit dark colour). As the compensation degree is 70%, the first step mode is the dominant oscillation mode. The results in Table1 show that the maximum torque response value appears between LPB and GEN when compensation degrees are 30% and 50%. As  $X_C$  equals to 70%, the maximum torque response value appears between LPA and LPB (the maximum torque between LPB and GEN is 3.498 p.u., however, it is 5.222 p.u. between LPA and LPB). Therefore, the effect on the position of the maximum torque should be considered in the process of safety analysis on torsional vibration of turbo-generator shafts to avoid locating the wrong dangerous section.

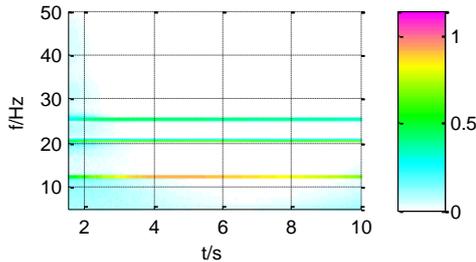
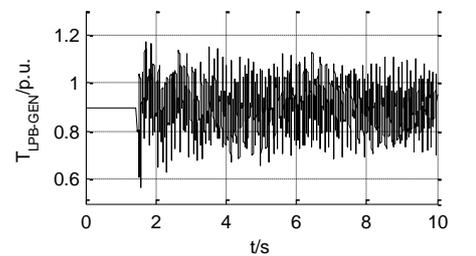


FIGURE 1 The torque response curve between LPB and GEN and its time-frequency contour map (XC=30%)

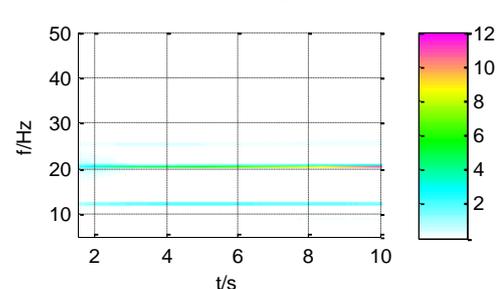
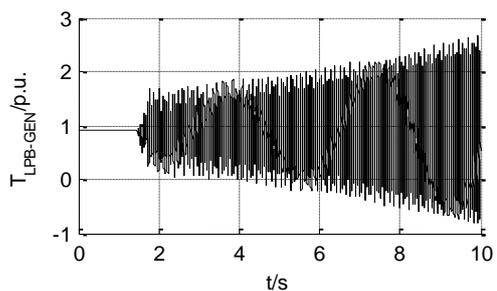


FIGURE 2 The torque response curve between LPB and GEN and its time-frequency contour map (XC=50%)

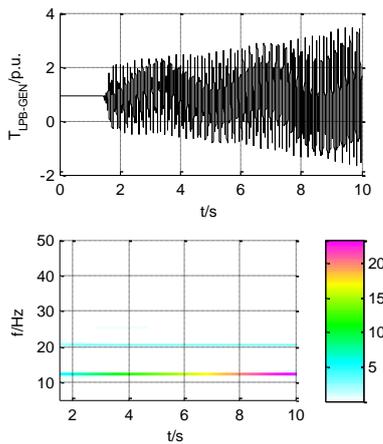


FIGURE 3 The torque response curve between LPB and GEN and its time-frequency contour map (XC=70%)

TABLE 1 The maximum torque response values within ten seconds when compensation degrees are different

	$X_c=30\%$	$X_c=50\%$	$X_c=70\%$
$T_{HIP-LPA}/p.u.$	0.928	1.122	1.609
$T_{LPA-LPB}/p.u.$	1.038	1.826	<b>5.222</b>
$T_{LPB-GEN}/p.u.$	1.172	<b>2.691</b>	3.498

Table 2 shows the results of the torques between LPB and GEN under the different series compensation degrees analyzed by MSPA. As shown in the table, all of the torsional vibration frequencies can be correctly recognized. When the series compensation degree is 30%, the modes at different frequencies have the positive damping, so the torque response curve is damped. As the series compensation degree is 50%, the modal damping at the second order frequency is negative and the torque curve appears to oscillation and divergence. As the series compensation degree is 70%, the modal damping at the first order frequency is negative and the torque curve appears to oscillation and divergence. The frequencies of SSR (converted to the rotor side) in negative damping area move towards into low frequency area with the increasing of the compensation degree. Therefore, in order to prevent the happening of SSR, the series compensation degrees must be designed reasonably in the new transmission projects to avoid torsional natural frequencies.

TABLE 2 The results of the torques (TLPB-GEN) under the different compensation degrees analysed by MSPA

$X_c$	$f/Hz$			$\sigma$		
	30%	50%	70%	30%	50%	70%
Mode1	12620	12628	12642	-0.0346	-0.0184	<b>0.1278</b>
Mode2	20841	20826	20822	-0.0009	<b>0.1314</b>	-0.0024
Mode3	25634	25635	25632	-0.0356	-0.0496	-0.0383

4.2 ANALYSIS OF VARIOUS POWER OUTPUTS

The simulation conditions are: series compensation degree of XC=50%, terminal voltage of  $V_t=1.0$  p.u.. It is assumed that the system is experienced by an instantaneous single phase short-circuit from 1.5 s to 1.6 s in the middle of the line which contains series compensation. The simulation lasts 10 s. Power outputs

(described by Pe) are respectively set as 0.3 p.u., 0.6 p.u. and 0.9 p.u.. The torque response curve and time-frequency contour map are illustrated in Figure 2.

The torque between LPB and GEN response curves and their time-frequency contour maps for Pe=0.3 p.u. and Pe=0.6 p.u. are given in Figure 4 and Figure 5. The maximum torque response values in each section within ten seconds when power outputs are different and the results of the torque between LPB and GEN based on MSPA are shown in Table 3 and Table 4. It can be seen from the figures that all of torque response curves are divergent and the second step mode is the dominant oscillation mode. With the increasing of power output, the maximum torque response values in each section are also getting higher. All of the maximum torque response values appear between LPB and GEN, at the same time, the dominant oscillation mode of the torque has no change. In other words, the dominant oscillation mode will not change when power output is changed. However, in view of the attenuation factors, the increasing of power output can make the damping effect slightly change and the trend of the change develops towards the direction that is not conducive to restrain SSR.

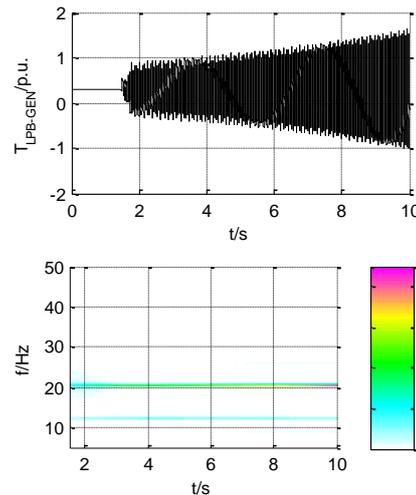


FIGURE 4 The torque response curve between LPB and GEN and its time-frequency contour map (Pe=0.3p.u.)

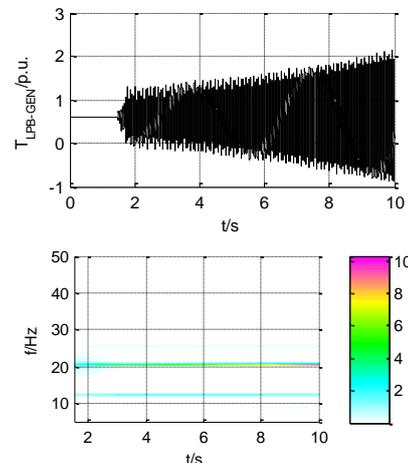


FIGURE 5 The torque response curve between LPB and GEN and its time-frequency contour map (Pe=0.6p.u.)

TABLE 3 The maximum torque response values within ten seconds when power outputs are different

	$P_c=0.3p.u.$	$P_c=0.6p.u.$	$P_c=0.9p.u.$
$T_{HIP-LPA}/p.u.$	0.523	0.820	1.122
$T_{LPA-LPB}/p.u.$	0.941	1.334	1.826
$T_{LPB-GEN}/p.u.$	1.615	2.131	2.691

TABLE 4 The results of the torques (TLPB-GEN) under the different power outputs analysed by MSPA

$P_c/p.u.$	$f/Hz$			$\sigma$		
	0.3	0.6	0.9	0.3	0.6	0.9
Mode1	1263	12635	12628	-0.0434	-0.0319	-0.0184
Mode2	2083	20832	20826	<b>0.1107</b>	<b>0.1199</b>	<b>0.1314</b>
Mode3	2563	25635	25635	-0.0607	-0.0530	-0.0496

## 5 Conclusions

The series capacitor compensation technique can cause SSR which will seriously influence the safe operation of large turbo-generator shafts. The transient torques of SSR are simulated in different compensation degrees and unit load levels in the paper. Based on the modified segmentation Prony algorithm proposed in this paper, the characteristic frequencies and damping parameters are identified and analysed. The MSPA, complex Morlet wavelet and time-frequency contour map are integrated and used in analysis of transient torques of SSR in time-frequency field. The achievements of this dissertation are

## References

- [1] Jaroslaw Z, Janusz M 2011 *Metrology and measurement systems* **18**(3) 371-8
- [2] Shim K S, Nam H K, Lim Y C 2011 *European transactions on electrical power* **21**(5) 1746-62
- [3] Maergoiz L S 2008 *Doklady mathematics* **78**(3) 822-4
- [4] Jaroslaw Z, Janusz M 2012 *Metrology and measurement systems* **19**(3) 489-98
- [5] Deng J X, Tu J, Chen W H 2007 *Power System Technology* **31**(7) 36-41
- [6] Dong H, Liu D C, Zou J F 2006 *High Voltage Engineering* **32**(6) 97-100
- [7] Nam S R, Lee D G, Kang S H 2011 *Journal of electrical engineering & technology* **6**(2) 154-60
- [8] Ding L, Xue A C, Li J 2010 *Automation of Electric Power Systems* **34**(22) 24-8
- [9] Barone P, Massaro E, Polichetti A 1989 *Astronomy and Astrophysics* **209**(1) 435-44
- [10] Gu Y J, Chen D C, Jin T Z 2011 *Lecture Notes in Electrical Engineering* **100**(4) 211-7
- [11] Kim J H, Park B Y, Akram F 2013 *Sensors* **13**(3) 3724-38
- [12] Jiang D X, Diao J H, Zhao G 2005 *Proceedings of the CSEE* **25**(6) 146-51
- [13] He P Li P, Sun H Q 2011 *Procedia Engineering* **15** 464-68
- [14] IEEE Subsynchronous Resonance Working Group 1985 IEEE Transactions on Power Apparatus and Systems **104**(5) 1565-72

mainly as follows: (1) The second order natural frequency of torsional vibration is the dominant oscillation mode when compensation degrees are 30% and 50%. As the compensation degree is 70%, the first step mode is the dominant oscillation mode. (2) The transient torques develops towards divergence direction and the frequencies of SSR (converted to the rotor side) in negative damping area move towards into low frequency area with the increasing of the compensation degree. (3) With the increasing of power output, the maximum torque response values in each section are also getting higher. All of the maximum torque response values appear between LPB and GEN. (4) The dominant oscillation mode will not change when power output is changed. However, the increasing of power output can make the damping effect slightly change and the trend of the change develops towards the direction that is not conducive to restrain SSR.

## Acknowledgments

The authors are grateful for the financial support from the Natural Science Foundation of Beijing, China (Grant No. 3132015) and the Fundamental Research Funds for the Central Universities (Grant No. 12ZX01).

## Authors



**Chengbing He, born in October, 1974, Sichuan, China**

**Current position, grades:** an associate professor in the School of energy, power and mechanical engineering at North China Electric Power University (NCEPU)

**University studies:** the Ph.D. degree in Thermal power Engineering from NCEPU

**Scientific interest:** on-line monitoring and analyse of coupled bending and torsional vibrations, fault diagnose of thermal equipment in power plant, and state maintenance of thermal equipment, etc.

**Publications:** 31



**Dongchao Chen, born in December, 1985, Jilin, China**

**Current position, grades:** the Ph.D. candidate in the School of energy, power and mechanical engineering at North China Electric Power University (NCEPU)

**University studies:** the bachelor's degree in Thermal and Power Engineering and the master's degree in Thermal Engineering from NCEPU

**Scientific interest:** On-line monitoring of coupled bending and torsional vibrations, fault diagnose of thermal equipment in power plant.

**Publications:** 2

# Experimental studies of differential heating for artificial upwelling

Ming Lv\*, Xu Yan, Xin Nie, Huachen Pan, Haiqiang Liu

*School of Mechanical Engineering, Hangzhou Dianzi University, 310018, Hangzhou, China*

*Received 1 June 2013, www.tsi.lv*

## Abstract

PIV experiments were carried out to study the mechanism and characteristics of a new artificial upwelling technology via differential heating named "Differential- Heating-Liquid-Upwelling" (DHLU). Results show that there is a small scaled area with high temperature around the modelling heat point source in DHLU system. The existence of this high temperature area is suggested as the power source of upwelling. Obvious upwelling flows upon the heating source were observed in DHLU system. The stream tube of upwelling for DHLU system is just like an upside-down cone. The max ascending velocity of horizontal layer increases firstly and decreases then as the height increases. A few of fluid masses with high ascend velocities were observed and the mechanism of the DHLU is revealed.

*Keywords:* artificial upwelling, differential heating, mechanism

## 1 Introduction

Upwelling is one of the most important conditions for rich marine fish stocks [1]. Upwelling can bring the nutrients which is rich in deep seas to the top and euphotic layer of the ocean. Therefore the phytoplankton which is at the bottom of most marine food chains can be fed. So the areas with natural upwelling are the most productive ocean fishing grounds in the world.

In recent years, only 75% of the world's commercial fish stocks are being fished at or above mean sustainable levels [1]. This situation is likely to get worse as the world's population grows. Artificial upwelling is considered as a promising technique for the enhancement of productivity in a sustainable way, as artificial upwelling can make up for the deficiencies of nature upwelling which has a restriction of time and space [1, 2].

Many studies were carried out for artificial upwelling technologies over the past few decades. McClimans et al. [3, 4] realized artificial upwelling with a freshwater source firstly, and then a combined method of fresh water and bubble curtain, with which significant amounts of nutrient-rich seawater could be lift to the light zone and provide an environment in which useful algae can survive at a large-scale fjord-experiment.

Kazuyuki et al. [5, 6] described an ocean nutrient enhancer named "TAKUMI" to upwelling deep ocean water with a pump which is powered by a diesel generator. The prototype of the machine was manufactured and set-up at the centre of Sagami Bay in Japan. However, the deep seawater flow rate is only about 1.2 m<sup>3</sup>/s.

Isaacs et al. [7] proposed to use wave energy to invert the density structure of the ocean and therefore the deep,

nutrient-rich water can be pumped into the euphotic layers. Liu et al. [8, 9] also described a similar wave-driven artificial upwelling device which consists of a buoy and a long pipe with a one-way valve. The main principle of the wave driving device is that the valve is only open on the down slope of a wave and close on the up slope. An estimated flow rate is about 0.45 to 0.95m<sup>3</sup>/s for a wave height of 1.90m and a wave period of 12s.

Tsubaki et al. [10] proposed an artificial upwelling technology based on the concept of "perpetual salt fountain" which is initiated by Stommel et al. In many areas of the tropical and subtropical ocean, warm salty water overlies colder fresher water. And that causes a famous vertical convective motion calling "salt finger". The salt fingers occur because of the difference in the diffusivities of heat and salt. When a pipe is inserted to connect deep sea and the surface and the pipe is filled with the low salinity deep sea water, the salinity of the water inside the pipe is lower than that outside. The upwelled deep sea water becomes almost the same temperature as the surrounding water. Hence buoyancy occurs in the pipe. The upwelling flow can be continues as long as the differences of the temperature and salinity exist. The flow rate with a single pipe was estimated as approximately 45m<sup>3</sup>/day [10]. The results would suggest that if such a perpetual salt fountain were to be viable for an ocean farming project, a large number of upwelling pipes would be necessary [11].

Liang et al. [12] proposed an air-lift pump for upwelling deep ocean water. For the air-lift pump the air is compressed into a vertical pipe, dipped in water. Bubbles ascend and the water level in the pipe rises due to the density decrease in the air-water mixture. Once the water level reaches the top of the pipe and the water flows out, the water flows continuously from the lower

\* *Corresponding author* e-mail: lvmingcn@163.com

end. For the air-lift pump the seawater flow rate ratio could be hundred times higher when compared to the air flow rate [12]. McClimans et al. [4] and Fan et al. [2] studied the air-lift pump experimentally and the effect of the air-lift upwelling method was confirmed.

In this paper, a new technology of artificial upwelling via differential heating is proposed. This renovation technology is based on the phenomena that the vertical temperature gradient of fluid, which is hot at bottom and cold at top, can cause the vertical density gradient inside the fluid, which will cause the inner vertical convection of the fluid for the reason of buoyancy. In this paper, PIV experiments were carried out to study the mechanism and characteristics of this new technology which we named "Differential- Heating-Liquid-Upwelling" (DHLU for short in follows).

## 2 Experimental setup

A schematic view of the experimental apparatus is shown in Figure 1. Heat source of the DHLU system was modelled by a self-made heating system. The heating system consisted of a heating stick, a quartz fibre tube and a temperature & power controller. The heating stick with a diameter of 10 mm and a length of 25 mm was placed inside the heat-insulated quartz fibre tube with an inner diameter of 30 mm and a length of 400 mm. And at the top, a length of 20 mm of the heating stick was kept outside of the quartz fibre to simulate a point heat source. The heating temperature and the power of heating stick can be controlled by the Temperature & Power Controller separately.

The environment of the upwelling was modelled by a glass tank of 700×400×550 mm. The tank was filled with

water. The water depth was kept at 420 mm. At the top of the tank where was a water cooler. The water cooler was immersed into water to keep the temperature of the top water layer at a constant level. The heating system was fixed vertically at the bottom centre of the glass tank, as Figure 1 shows. The height of the heating stick's head from the bottom of tank was fixed at 100 mm to avoid the influence of the tank bottom. Rubber seals were used to seal the heating system.

A group of nine K-type thermocouples was arranged with its testing points placed nearly equidistantly upon the heating system along the central axis of heating stick, as shown in Figure 1 and Table 1. The No.1 to No.8 thermocouples were used to measure the vertical temperature variation in steady upwelling with height along the central axis of heating stick. The No. 9 thermocouple was placed in air to measure the room temperature. Temperatures measured by thermocouples were collected by a data logger of HP34970A with 34901A Armature Multiplexer Module (Agilent Technologies, Inc.)

The fluid field of the upwelling was measured by the V3V-2D PIV system (produced by TSI, USA). The PIV system includes a laser (Nd:YAG laser, 350mJ, 15 Hz), three cameras (PowerView Plus 4MP, 2K × 2K pixels), a synchronizer and Insight4G software. This system can measure a three-dimensional flow field with a max area of 140×140×100 mm. Figure 1 shows the positions of the laser and cameras of PIV system in experiment.

In Experiments, the power of heating stick was set at 200W. And then the fluid field was measured by PIV system when a steady upwelling could be observed.

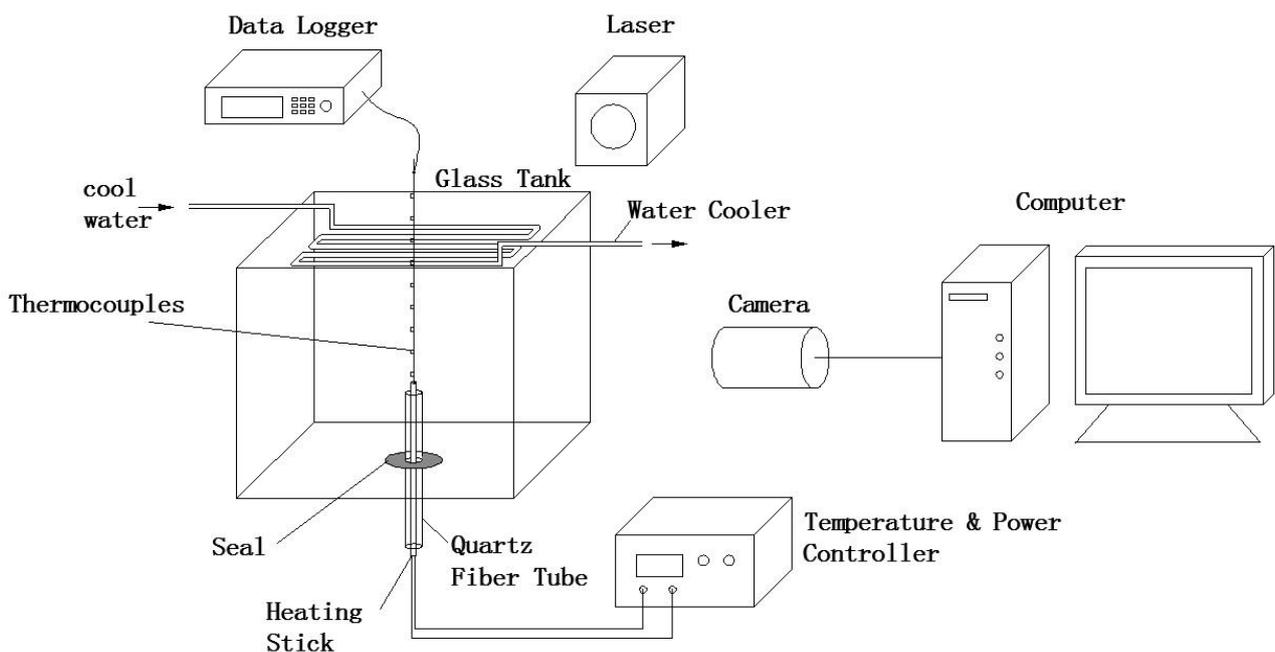


FIGURE 1 Schematic diagram of the experimental system

TABLE 1 Heights of the test points of thermocouples from the top of the heating stick

No.	Height (cm)
1	0
2	1
3	5
4	10
5	15
6	20
7	25
8	30
9	35

**3 Results and discussion**

Experiments were carried out on DHLU system. In steady case, the vertical temperature distribution in upwelling along the central axis of heating stick was shown in Figure 2. The temperature of heat source could be kept at about 61°C when the power of heating stick was set as 200 W in experiment. As the height increased from 0, the temperature dropped quickly firstly. And the temperature dropped to 26°C when the height rose only 5 cm from the top of heating stick. And then as the height increased again the drop of temperature became slowly. And finally the temperature of water reached an environmental water temperature of about 20°C. The room temperature measured in experiments was 17.5°C. Above all, we can see that there is a small scaled area with high temperature around the modelling heat point source in DHLU system. The existence of this high temperature area is suggested as the power source of upwelling.

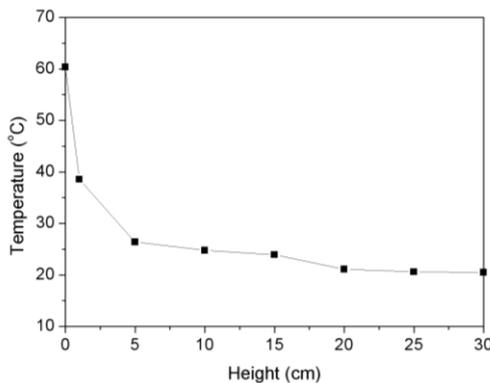
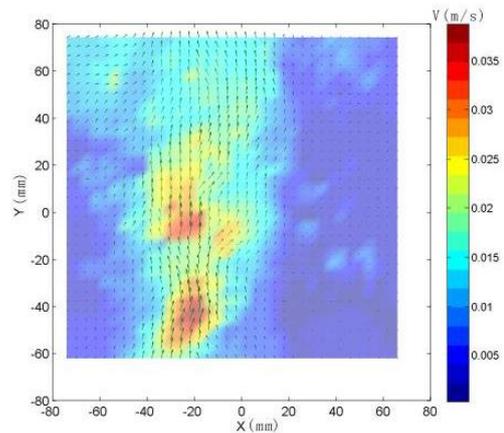


FIGURE 2 Variations of vertical temperature in steady upwelling vs. the height of test point from top of heating stick (P=200 W)

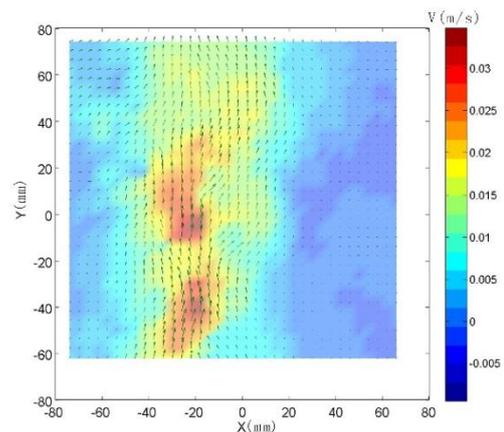
The fluid field of upwelling was measured by PIV in steady case via V3V technology. The measure area was set as X=140 mm, Y=140mm, Z=100mm. X and Z are horizontal coordinates. Y is vertical coordinate. The position of modelling heat point source was set at X=-20, Y=-70, Z=-650 mm.

Figure 3 shows the velocity field at a XY plane across the centre of heating stick in a steady case. Both the full velocity and the vertical velocity(Y-direction velocity) were figured out. It is shown that the temperature difference caused by the heating at bottom can cause obvious upwelling. There were obvious upwelling flows with up velocity upon the heating source. The stream

filament of upwelling appeared narrow at bottom and wide at top. More studies were carried out to analysed the three dimensional shape of the upwelling boundary with V3V technology. The 3D iso-surfaces of the velocity of 0.019 m/s was figured out, as Figure 4 shows. From the Figure we can see the explicit shape and distribution area of upwelling. The stream tube of upwelling for DHLU system is just like an upside-down cone.



(a) Velocity



(b) Vertical velocity

FIGURE 3 The velocity field at a XY plane across the centre of heating stick (Z= -650 mm)

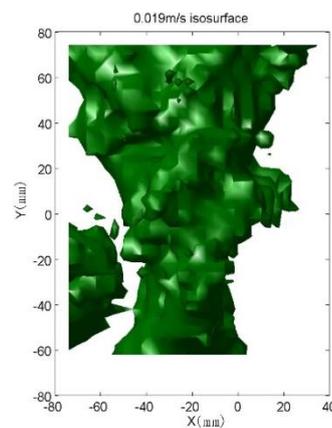


FIGURE 4 Iso-surfaces of velocity equals 0.019m/s in upwelling

Furthermore, several water masses with high speed were observed in upwelling. Speeds of those water masses are obviously higher than the fluid around. A string of those water masses appeared and floated up in steady case. This suggests that the mechanism of the DHLU may be as follows: The heating at bottom of the heat source produces a series of fluid masses with relative higher temperature and lower density. And those fluid masses ascend due to the density decrease and bring the fluid around ascend together. And a continuous steady upwelling is formed as the heating at bottom continuous.

We also analysed the vertical velocity field of horizontal layers at different heights from the modelling heat point sources, as shown in Figure 5 to Figure 7. Totally speaking, in steady case the max ascending velocity of horizontal layer increases firstly and decreases then as the height increases. The area of upwelling increases as the height increases.

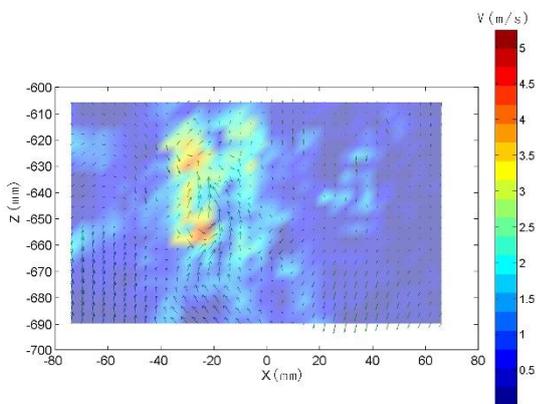


FIGURE 5 The vertical velocity field at a horizontal XZ plane with a height of 40 mm ( $Y = -30$  mm)

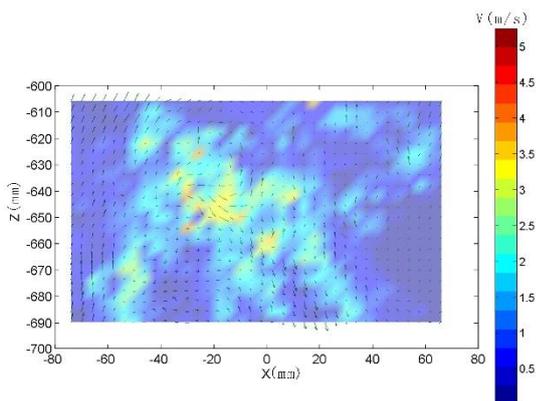


FIGURE 6 The vertical velocity field at a horizontal XZ plane with a height of 80 mm ( $Y = 10$  mm)

## References

- [1] Brian K 2003 Enhancing fish stocks with wave-powered artificial upwelling *Ocean Coast. Manage* **46** 901–15
- [2] Wei F, Jiawang C, Yiwen P, Haocai H, Chen T, Arthur C, Ying C 2013 Experimental study on the performance of an air-lift pump for artificial upwelling *Ocean Engineering* **59**(1) 47-57

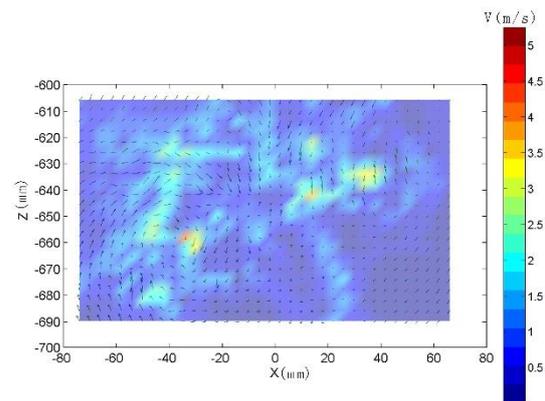


FIGURE 7 The vertical velocity field at a horizontal XZ plane with a height of 120 mm ( $Y = 50$  mm)

## 4 Conclusions

PIV experiments were carried out to study the mechanism and characteristics of a new artificial upwelling technology via differential heating named “Differential-Heating-Liquid-Upwelling” (DHLU).

Results show that there is a small scaled area with high temperature around the modelling heat point source in DHLU system. The existence of this high temperature area is suggested as the power source of upwelling. Obvious upwelling flows upon the heating source were observed in DHLU system. The stream tube of upwelling for DHLU system is just like an upside-down cone. The max ascending velocity of horizontal layer increases firstly and decreases then as the height increases. A few of fluid masses with high ascend velocity was observed and the mechanism of the DHLU is suggested as follows: The heating at bottom of the heat source produces a series of fluid masses with relative higher temperature and lower density. And those fluid masses ascend due to the density decrease and bring the fluid around ascend together. And a continuous steady upwelling is formed as the heating at bottom continuous.

## Acknowledgements

Authors wish to acknowledge the sponsor of the National Natural Science Foundation of China (No.51209062 and No. 51305113), the Science and Technology Project of Zhejiang Province (No. 2012R10003 and No. 2010R50003).

- [3] McClimans T A, Eidnes G, Aure J 2002 Controlled artificial upwelling in a fjord using a submerged fresh water discharge: computer and laboratory simulations *Hydrobiologia* **484**(1) 191–202
- [4] McClimans TA, Handa A, Fredheim A, Lien E, Reitan K I 2010 Controlled artificial upwelling in a fjord to stimulate non-toxic algae *Aquacultural Engineering* **42** 140-7

- [5] Kazuyuki O, Alan J M 2003 Real sea experiment of ocean nutrient enhancer "TAKUMI" upwelling deep ocean water *Oceans Proceedings* 881-5
- [6] Kazuyuki O, Hiroyuki N 1999 The deep ocean water upwelling machine using density current-Creation of fishing ground and absorption of CO<sub>2</sub> *Oceans '99 MTS/IEEE* 2 1019-24
- [7] Isaacs J D, Castel D, Wick GL, 1976 Utilization of the energy in ocean waves *Ocean Engineering* 3 175-87
- [8] Liu C C K, Jin Q 1995 Artificial upwelling in regular and random waves *Ocean Engineering* 22(4) 337-50
- [9] Liu C C K 1999 Research on artificial upwelling and mixing at the University of Hawaii at Manoa *IOA Newsletter* 10(4) 1-8
- [10] Tsubaki K, Maruyama S, Komiya A, Mitsugashira H 2007 Continuous measurement of an artificial upwelling of deep sea water induced by the perpetual salt fountain *Deep-Sea Research I* 54 75-84
- [11] Williamson N, Komiya A, Maruyama S, Behnia M, Armfield S W 2009 Nutrient transport from an artificial upwelling of deep sea water *J Oceanogr* 65 349-59
- [12] Liang N K, Peng H K 2005 A study of air-lift artificial upwelling *Ocean Engineering* 32 731-45

Authors	
	<p><b>Lv Ming, born on February 26, 1982, Hunan, China</b></p> <p><b>Current position:</b> Ph.D., lecturer of Department of Ocean Engineering in Hangzhou Dianzi University  <b>University studies:</b> Zhejiang University  <b>Scientific interest:</b> major in clean energy production and artificial upwelling  <b>Experience:</b> Zhejiang University 2004/9-2009/12 Energy and Environment Engineering PhD; research subjects engaged in clean energy production and artificial upwelling</p>
	<p><b>Yan Xu, born on March 8, 1989, Anhui, China</b></p> <p><b>Current position:</b> postgraduate of Department of Ocean Engineering in Hangzhou Dianzi University  <b>University studies:</b> Hangzhou Dianzi University  <b>Scientific interest:</b> major in artificial upwelling  <b>Experience:</b> Hangzhou Dianzi University 2012/9-current Ocean Engineering postgraduate; research subjects engaged in artificial upwelling</p>
	<p><b>Nie Xin, born on November 15, 1974, JiangXi China</b></p> <p><b>Current position:</b> Ph.D., Associate Professor of Department of Ocean Engineering in Hangzhou Dianzi University  <b>University studies:</b> Zhejiang University  <b>Scientific interest:</b> Fluid engineering and energy engineering  <b>Experience:</b> Zhejiang University 2000/9-2007/3 Thermophysics Engineering PhD; research subjects engaged in theory and method of Fluid engineering and energy engineering; Computational Fluid Dynamics (CFD) technologies and fuel energy clean utilization</p>
	<p><b>Pan Huachen, born on February 27, 1956, JiangSu China</b></p> <p><b>Current position:</b> Ph.D., Professor of Department of Ocean Engineering in Hangzhou Dianzi University  <b>University studies:</b> Nanjing University of Aeronautics and Astronautics  <b>Scientific interest:</b> Fluid engineering  <b>Experience:</b> Nanjing University of Aeronautics and Astronautics 1983/3-1986/3 Engine Inlet Engineering PhD; research subjects engaged in theory and method of Fluid engineering; Computational Fluid Dynamics (CFD) technologies</p>
	<p><b>Liu Haiqiang, born on April 19, 1980, JiangXi China</b></p> <p><b>Current position:</b> Ph.D., lecturer of Department of Ocean Engineering in Hangzhou Dianzi University  <b>University studies:</b> Zhejiang University  <b>Scientific interest:</b> Intelligent design and digital product design  <b>Experience:</b> Zhejiang University 2005/9-2010/9 Mechanical Manufacturing and Automation PhD; research subjects engaged in theory and method of Product Data Management; PLM methodology and research integrated techniques of CAX/PDM and focused on build the integrated product data model</p>

# Method of multi-feature fusion based on SVM and D-S evidence theory in Trojan detection

Shengli Liu<sup>\*</sup>, Xiang Gao, Pan Xu, Long Liu

State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, 450002, China

Received 1 March 2014, www.tsi.lv

## Abstract

According to the low accuracy and low stability of the single feature-based method for Trojan detection, a multi-feature fusion method based on SVM and DS evidence theory is proposed. First, three types of flow features such as session, upload data of session/download data of session, distribution of data packet size are extracted from the data stream. Then the SVM classification results of each single feature are used as evidences to construct the basic probability assigned (BPA). Finally, we use DS combination rule of evidence to achieve the decision fusion and give the final detection results by fusion results. The experimental results showed that the accuracy of multi-feature fusion method was 97.48% which has good performance on accuracy and stability compared with the single feature method in Trojan detection.

*Keywords:* Trojan detection, support vector machine, DS evidence theory, multi-feature fusion

## 1 Introduction

With the fast development of information technology, the rapid growth of data has become a serious challenge as well as a good opportunity in many industries [1]. In the era of big data, Trojan technology based on data stream communication has become the primary means of attackers to steal confidential information, which poses a serious threat to network security. Therefore, the research of Trojans based on the data stream communication brooks no delay.

At present, many scholars, at home or abroad, have made a lot of research and exploration about Trojan detection methods, mainly divided into two host-side and network-side [2, 3]. Among them, only the host-side software protection can't achieve effective detection for Trojan, and the Trojan detection method based on data flow analysis of network traffic gradually becomes a hot topic, attracting more and more attention. However, the existing work, mostly focus on the researches of a single feature[4-6], typically by establishing and maintaining a pre-defined characteristics database, using the database to match the characteristic information with the network data flow. If the match is successful, it gives an alarm. In addition, the comprehensive utilization of multi-feature also has made some achievements [7-9], but these studies are simply integrated multiple features without effective integration, resulting in the low rate of Trojan detection, and prone to false negatives and false positives, thus affecting the accuracy and validity of the detection system.

In order to further improve the accuracy of Trojan detection, on the basis of feature extraction, we propose a

combining multi-feature fusion Trojan detection method based on SVM and DS evidence theory. Using the DS evidence theory's advantage [10-11] of dealing with uncertainty information and the better classification capabilities of SVM in small sample, we combine multiple single features of Trojan detection information and get the final test results according to the decision rules. It's important to note that this paper is to study the Trojans of information stealing. After such Trojans are implanted into the target system, they will record or collect all kinds of important information of the target system, such as stealing all kinds of user names and Passwords, and send the information to the attacker through a particular way.

## 2 Feature extraction of Trojan

Due to the limit of firewall and NAT networks, the controlled terminal initiates a connection request to the control terminal in general Trojans. After that the connection is established, once the controlled terminal takes orders of the control terminal, they generally open a new session to execute, which will be ended when the session is finished. Figure 1 shows the timing sequence.

Based on the flow characteristics of the Trojans' communication, this paper detects Trojans by analysing the roles of both Trojans communication sides (controlled terminal and control terminal) playing respectively and the inconsistency of performance in the communication and the roles that they play. By analysing the communication theory of many Trojans and a great amount of communication features, we select three types of effective communication behavioural characteristics to

<sup>\*</sup>Corresponding author e-mail: liushengl2013@163.com

distinguish between normal network communication behaviours and Trojan communication behaviours.

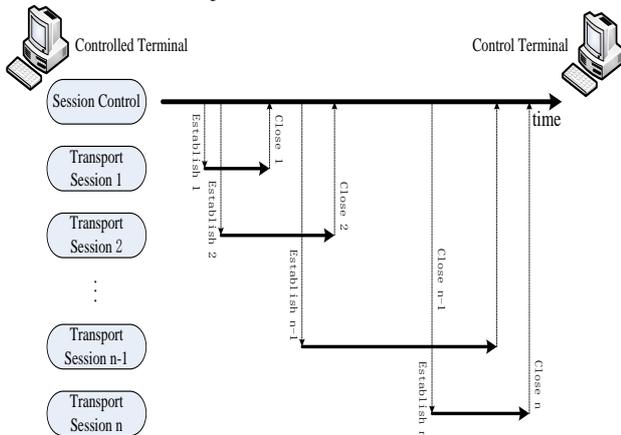


FIGURE 1 The timing sequence of Trojans communication session

2.1 SESSION FEATURE

There is significant difference between Trojans and normal HTTP C/S server. For Trojans, interactive commands, file resource searching and file transfer need artificial waiting and operating time. Therefore a longer time of communication is needed, while the normal HTTP C/S server performance to release the connection after completing the first request, so there is seldom a long connection. We use three parameters to describe the session features. Let  $T_1$  be the number of TCP sessions,  $T_2$  be control session duration and  $T_3$  be IP flow duration.

- 1) The number of TCP sessions  $T_1$ : when the IP connections launch the first TCP session, the value of  $T_1$  is one, before the end of the session, the value of  $T_1$  plus 1 when a TCP session is finished.
- 2) Control session duration  $T_2$ : the control session duration is the duration of the first TCP sessions.
- 3) IP flow duration  $T_3$ : the duration of the data communication between one pair of IP is from the first TCP session between the pair of IP to the end of all sessions.

TABLE 1 The characteristic of data flow

session feature			upload data volume/download data volume features	distribution feature of packet size	
$T_1$	$T_2$	$T_3$	$D$	$S_1$	$S_2$
7.0	184.25	190.45	6.83	0.64	0.75

3 SVM and DS evidence theory

3.1 SVM

SVM (Support-Vector Machines) [12] is an important product of statistical learning theory. SVM is able to make experience risk minimization in a fixed structural

2.2 SESSION UPLOAD DATA VOLUME / DOWNLOAD DATA VOLUME FEATURES

Because of that the data flow of Trojan communication is the upload data flow from the inside out; we extract characteristic D of upload data volume / download data volume in the session. When the controlled terminal initiates the communication, it is presented as the client of connecting C/S model requested resources. After that the connection is established, it will turn into the service side to provide resources, in which cases the amount of data sent is much larger than the amount of data received. Thus this ratio is greater than one for Trojans, while in normal HTTP C/S server, the upload data is far less than the amount of download data, the ratio is less than one.

2.3 THE DISTRIBUTION FEATURE OF PACKET SIZE

For the characteristic of packet size distribution in the communication process, we extract the following characteristics:

- 1) The number of receiving small packets / the number of small packets  $S_1$ , most of the packets received in the controlled terminal from the Trojans are the control commands, which are mainly small packets different from the big packets for sending information. So the value is generally greater than 0.5, and the normal HTTP C/S server is always the case that clients receive resource information from services, such as website information, most of which are big packets.
- 2) The number of big packets sent / the number of large packets  $S_2$ , most of the packets sent in the controlled terminal from the Trojans are the resource files and host information, which rarely receiving a large package, so the value is generally greater than 0.5. The normal HTTP C/S server is the case that the client requests the resource, and only sending the request to the server, receiving server's resources, so the number of big packets sent is low.

To sum up, we extract six features from three types, which are session feature, upload data volume / download data volume features and distribution feature of packet size. Grey pigeons Trojan sample is used as an example, intercepting data flow of a period of time in these samples. Table 1 shows every characteristic of data flow.

risk, because the problem will be converted into a constrained quadratic programming problem in SVM ultimately, so SVM is usually able to get a global optimal solution. SVM is a method that by defining the optimal linear hyper plane and summarizing the search of optimal linear hyper plane algorithm as solving a convex programming problem, and using Mercer nuclear

expansion theorem. The sample space is mapped into a dimensional or even infinite dimensional feature space by nonlinear mapping function T, and making it possible to solve regression and highly nonlinear classification problems in the sample space by using linear learning machine method. The classification function as

$$f(x) = \sum_{x_i \in S_V} a_i y_i k(x, x_i) + b. \tag{1}$$

In the formula,  $a_i$  is Lagrange multiplier,  $b$  is a given threshold according to the training samples,  $S_V$  is support vector set and  $k(x, x_i)$  is kernel function. Among them, only the kernel function satisfies Mercer conditions, then it corresponds to inner product of a certain space, and the nonlinear transformation can be changed by the inner product to linear transformation of high-dimensional space.

The process of solving Equation (1) is the training process of SVM. Therefore the results through the training and evaluation are as same as the classification function used in the actual detection.

### 3.2 DS EVIDENCE THEORY

The evidence theory established by Dempster and Shafer, which is called the D-S theory, is an important method of uncertainty reasoning. DS theory combines the trust function from two or more evidence bodies into a new trust function as a basis for decision making. The principle is as follows:

Let  $\Theta$  be recognition framework. We define function  $m: 2^\Theta \rightarrow [0,1]$ , if there is  $m: 2^\Theta \rightarrow [0,1]$  such that:  $M(\phi) = 0$  ( $\phi$  is empty set),  $\sum m(A) = 1$  ( $A \in 2^\Theta$ ), then we call  $m(A)$  as that basic probability assigning (BPA) on the framework, which means the precise degree of confidence in the proposition, which is also known as the mass function. If  $m(A) > 0$ , A is focal elements.

Let  $m_1, m_2, \dots, m_n$  be BPA of different evidences on recognition framework,  $m(A)$  can be expressed as:

$$m(A) = \sum_{A_1 \cap A_2 \cap \dots \cap A_i = A} (\prod_{1 \leq i \leq n} m_i(A_i)) / k. \tag{2}$$

K is the normalization constant, and

$$k = 1 - \sum_{A_1 \cap A_2 \cap \dots \cap A_i = \phi} (\prod_{1 \leq i \leq n} m_i(A_i)).$$

### 4 The decision level fusion Trojan detection algorithm

By analysing the characteristic information of Trojans in the detection, we can get that Trojan session feature, session upload data volume / download data volume feature and distribution feature of packet size are independent of each other, so you can take advantage of DS theory, which has the ability of composite the independent evidences to combine the identification information from the SVM of different characteristics. Finally, the target type is given by decision module (Trojans or normal HTTP C/S server). The Trojan detection algorithm model is shown in Figure 2.

The basic idea of detection algorithm is as follows:

1) Trojan Detection of Single Feature SVM. According to Trojan feature extraction method mentioned before, we extract three types of single-feature, Trojan session feature  $T_1 T_2 T_3$ , session upload data volume / download data volume feature  $D$  and distribution feature of packet size  $S_1 S_2$ . Then we detect three types of single-feature above with SVM classifier.

2) The Choice of SVM Training Algorithm Using SMO (Sequential Minimal Optimization) algorithm [13] to train the training set, it is because of that the influence from the choice of kernel function parameters is much bigger than from the kernel function itself on the classification results. So this paper selects Radial Basis function (RBF) as the kernel function.

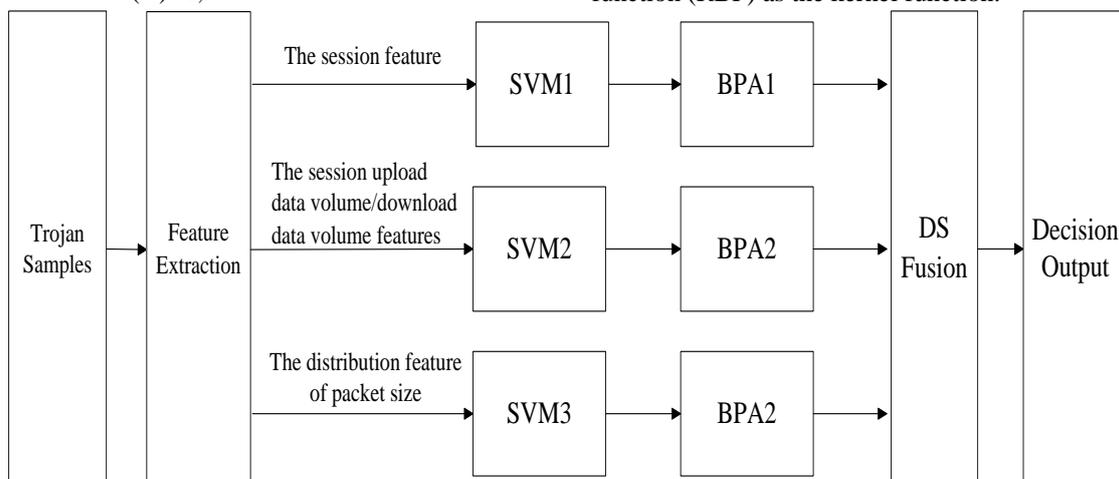


FIGURE 2 The Trojan detection algorithm model

In this method, given data sample  $(x_i, y_i)$ ,  $i = 1, 2, \dots, n$ , kernel function  $K(x_i, y_j)$  and adjustable parameters  $C$ , the necessary and sufficient conditions are Kuhn-Tucker (KKT) conditions that all training data samples should meet the following conditions:

$$a_i = 0 \Leftrightarrow y_i f(x_i) \geq 1, \tag{3}$$

$$0 < a_i < c \Leftrightarrow y_i f(x_i) = 1, \tag{4}$$

$$a_i = c \Leftrightarrow y_i f(x_i) \leq 1, \tag{5}$$

$f(x_i)$  is the output of data sample with number  $i$ . The diagram of KKT conditions is as follows

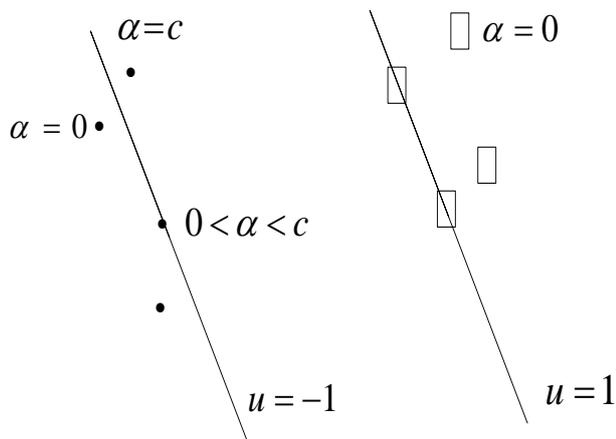


FIGURE 3 The schematic diagram of KKT conditions

The SMO algorithm breaks the problem down to the smallest size which is able to achieve, and each optimization only processes the optimization problem of two data samples. Its greatest advantage is that with analytical method for solving smallest optimization problems, avoiding the iterative algorithm and improving the overall computing speed. The algorithm does not require processing large matrix and no additional storage space is required. So this algorithm is suitable for processing the network data flow.

3) BPA Function Building. The standard output of SVM is  $\{1, -1\}$ , which is not a probabilistic output but a hard decision output, so we can't use it as BPA of evidence body. In order to construct the BPA of evidence theory, we build probabilistic modelling by combining many classification probabilistic results of two categories. First, assuming the given data of type 1, and for any  $x$ , we estimate the probability of matching class by sigmoid function [14], namely:

$$h_{ij} \approx p(y = i | y = i \text{ or } j, x). \tag{6}$$

For posterior probability  $p_i$ :

$$\min_p \left( \frac{1}{2} \sum_{i=1}^l \sum_{j:j \neq i} (h_{ij} p_i - h_{ij} p_j)^2 \right). \tag{7}$$

According to Equation (7), we can get the posterior probability after the training of the sample set. Then

testing the learning sample set, and you can get the accuracy of classification. So BPA function is defined as following:

$$m_j(A) = p_i q_i \tag{8}$$

4) The Decision Fusion and Judgment Rule. We can calculate the confidence of evidences by Equation (8), and the confidence of overall evidences by Equation (2).

Let  $A_j (j = 0, 1)$  be sample type (data flow of Trojan or normal data flow),  $A_w$  is the target category (data flow of Trojan or normal data flow). The classification decision obeys the following rules:

1)  $m(A_w) = \max\{m(A_j)\}$ , the target class is the class with the greatest confidence.

2)  $m(A_w) - m(A_j) > \varepsilon_1$  ( $\varepsilon_1 > 0$ ), the difference value of target class and other class should be greater than a certain threshold.

3)  $m(A_w) - m(\theta) > \varepsilon_2$  ( $\varepsilon_2 > 0$ ), the confidence of target class should be greater than the assign value of uncertain reliability.

4)  $m(\theta) < \varepsilon_3$  ( $\varepsilon_3 > 0$ ), the uncertain reliability of target class's evidence shouldn't be too large, and the assign value of uncertain reliability should be less than a certain threshold.

### 5 The prototype of detection system

In this paper, in order to verify the efficiency of Trojan communication detection method, a rapid Trojan detection system has been designed. The system consists of data acquisition module, feature extraction module, Trojan detection module and alarm response module. The structure is shown in Figure 4:

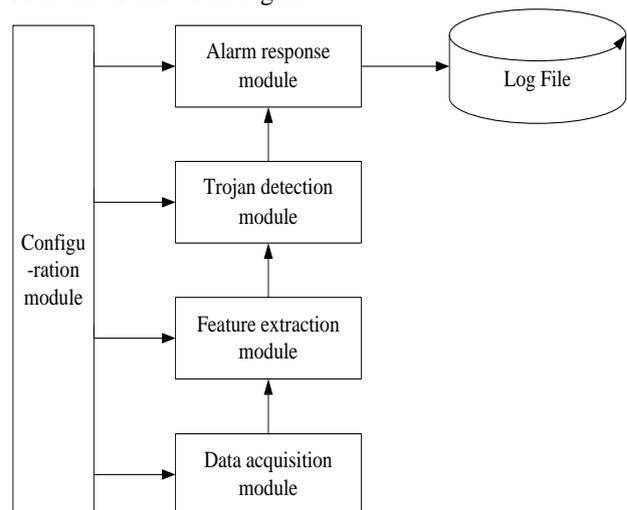


FIGURE 4 The structure of detection system

In this system, we use Winpcap to get data in data acquisition module. Feature extraction module is responsible for extracting the Trojans communication session, the session upload data volume / download data volume and packet size distribution. Trojan detection

module detects Trojan detection in real time based on Fusion Trojan detection algorithm of decision-making level, and alarm response module extract the appropriate information recorded in the log file. Configuration module is used to configure other modules.

**6 Trojan detection experiments**

**6.1 EXPERIMENT DESCRIPTION**

The experiment is conducted in a small laboratory environment. Shown in Figure 5, there are 20 hosts in the LAN access the internet and the host network bandwidth is 3Mbps. Among them, we implanted Trojans in host PC1, and build DNS in host PC5. The program containing the above-mentioned Trojan detection algorithm has been installed in a separate server host.

The operating environment:

- CPU: Xeon E5 2.3G
- Memory: 16G

In this paper, based on grey pigeons Trojan (Win32.Hack.Huigezi), we designed three types of Trojans with high concealment by increasing redundancy, dividing sending and discontinuous transmission to conceal their communication data flow features, which

are called Wake1.0, Wake2.0 and Wake3.0 respectively. Wake1.0 masks itself by increasing redundancy, Wake2.0 masks itself by dividing sending, and Wake3.0 masks itself by discontinuous transmission. The initial data sample consists of 410 normal network sessions and 145 Trojan sessions. Five groups of experiments were performed. Each of them randomly selected 90% of the initial data sample as the SVM training sample, and selected alternatively 10% of the data sample as the test sample. Then the test samples were divided into 10 test sets according to the type of data flow, and the data type of each test set was Trojan data flow or normal data flow.

First, we extracted six features from three types by using the method mentioned, which are session feature, upload data volume / download data volume feature and distribution feature of packet size, and normalized to them. Then we classified the data flow by multi-feature fusion method. In the method, we selected RBF as the kernel function of SVM model and used cross-validation method to select the error penalty parameter  $C$  and kernel parameter  $\sigma$  :  $C = 40$  ,  $\sigma = 2.43$  . We obtained judgment threshold from a lot of experiments:  $\varepsilon_1 = 0.6$  ,  $\varepsilon_2 = 0.7$  ,  $\theta = 0.1$  .

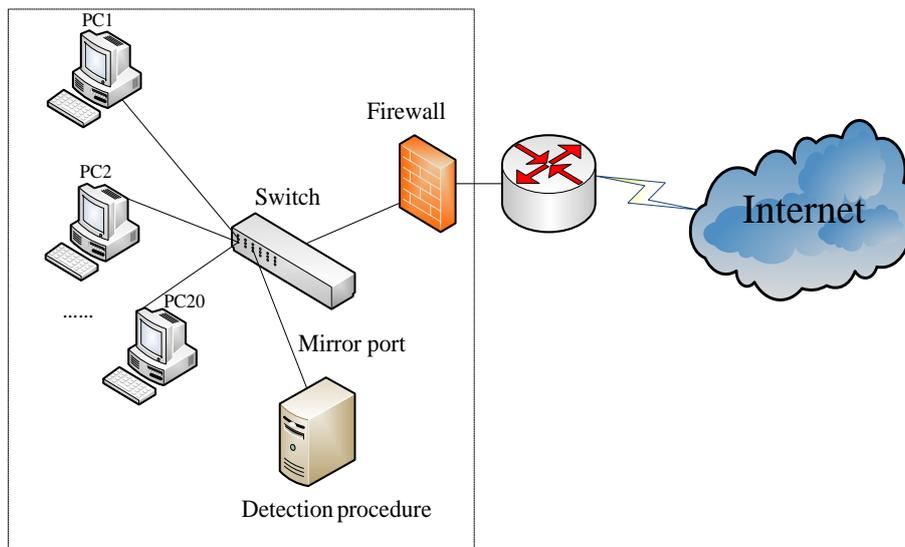


FIGURE 5 The experimental topological diagram

TABLE 2 The related records of three groups' experiments

The types of Samples	Single feature	Confidence function value			Detection results
		$m(A_0)$	$m(A_0)$	$m(\theta)$	
2-3 Wake1.0	$m_1$	0.5733	0.3125	0.1142	Indefinite
	$m_2$	0.6673	0.2134	0.1193	Indefinite
	$m_3$	0.6014	0.1756	0.2230	Indefinite
	$m$	0.8242	0.1405	0.0353	Wake1.0
4-6 Normal data flow	$m_1$	0.1012	0.7914	0.1074	Indefinite
	$m_2$	0.2345	0.5678	0.1977	Indefinite
	$m_3$	0.1947	0.6235	0.1818	Indefinite
	$m$	0.0754	0.8997	0.0249	Normal data flow
5-7 Wake3.0	$m_1$	0.5678	0.2565	0.1757	Indefinite
	$m_2$	0.7912	0.1345	0.0743	Indefinite
	$m_3$	0.6322	0.2379	0.1299	Indefinite
	$m$	0.8943	0.0924	0.0133	Wake3.0

6.2 ANALYSIS AND COMPARISON

1) We recorded the value of confidence function and recognition results of single-feature and fusion multi-feature in the experiments. Table 2 shows the related records of three groups' experiments randomly selected. In which,  $m_1$  is the session feature,  $m_2$  is upload data volume / download data volume,  $m_3$  is distribution feature of packet size and  $m$  is the fusion result. For example, No.2-3 means the data flow of the third test set in group 2.

Table 2 shows the experimental results:

a) Comparing the value of confidence function combined by session, upload data of session / download data of session, and distribution of data packet size with

the value of single-feature confidence function, the confidence of actual target increases a lot, which means the uncertainty of the detection is greatly reduced.

b) Three types of data samples which are unable to detect by multi-feature SVM (such as 2-3 and 5-7 in the table), can be accurately detected after the data fusion, so the multi-feature fusion method based on D-S evidence theory has good performance on accuracy and stability, which enhances the ability to detect Trojans.

2) Using the test set in the test samples as a unit, we recorded the detection rate and the multi-feature detection rate of five groups' experiments. The results are shown in Table 3. The detection rate means the percentage of correct classification in data samples, and  $m_1, m_2, m_3$  has the same meaning of Table 2.

TABLE 3 The experimental results

Experimental group	The detection rate of single feature			The fusion recognition rate of multiple features
	$m_1$	$m_2$	$m_3$	
1	60.25%	80.15%	65.34%	97.28%
2	87.23%	92.34%	89.23%	96.96%
3	70.34%	85.00%	80.00%	96.44%
4	70.46%	90.25%	78.05%	100%
5	62.25%	81.25%	68.75%	96.78%
Average	70.11%	85.80%	76.27%	97.48%

Table 3 shows the experimental results:

1) Comparing three types of single-feature detections, because of that the great difference between Trojan data flow and normal data flow in upload data volume and download data volume, the method based on upload data of session / download data of session performs better than the method based on session and distribution of data packet size. On the other hand, the test data which is complex in fact results in the inaccurate calculation to reduce detection rate. In summary, the single-feature detection has a high error rate, poor reliability and stability.

2) The average accuracy of multi-feature fusion detection reaches 97.48% and has less volatile. Comparing with the single-feature detection, the accuracy and stability has improved significantly. Because that the DS evidence theory is based on SVM's posterior probability and credits assigned of classification accuracy structure, it combines the different feature detection information with session, upload data of session / download data of session, and distribution of data packet size, and makes full use of multi-feature information to improve the accuracy and stability of Trojan detection.

Additional, we use BotHunter detection software [15] to test the samples above. BotHunter is a kind of driver-based IDS detection software which can detect the behaviors of network scanning and file downloading during the process of bot software infection. It matches with data flow feature and gives an alert when the

infected host is highly unusual. Some of the experimental results are shown in Table 4.

TABLE 4 The detection result of BotHunter

Experimental group	The types of Samples	The detection rate
2	Wake1.0	88.3%
3	Wake2.0Wake1.0	90.6%
5	Wake3.0	85.2%

We can get the conclusion that the detection rate of the method proposed in this paper is higher than the detection rate of BotHunter software, which has a better practical significance.

7 Conclusions

In order to improve the detection rate of Trojans, we propose a multi-feature fusion method based on SVM and DS evidence theory. First, we use three types of SVM classifiers respectively based on session, upload data of session / download data of session, distribution of data packet size for detecting Trojans. Then the SVM classification results of each single feature are used as evidences to construct the basic probability assigned (BPA). Finally, we make use of DS evidence theory to achieve the decision fusion and give final detection results by fusion results. The experimental results showed that the accuracy of multi-feature fusion method was 97.48% in the real network environment, which has good performance on accuracy and stability compared with the single feature method in Trojan detection.

**References**

[1] Wang Y Z, Jin X L, Cheng X Q 2013 Network Big Data Present and Future *Journal of Computers* **36** 1-15

[2] Liu T, Guan X, Zheng Q, Lu K, Song Y, Zhang W 2009 Prototype Demonstration Trojan Detection and Defense System *Proc of the IEEE Consumer Communications and Networking Conference* 1-2

[3] Zhang L, White G B 2007 An Approach to Detect Executable Content for Anomaly Based Network Intrusion Detection. *Proc of the IEEE Parallel and Distributed Processing Symposium* 1-8

[4] Dusi M, Croti M, Gringoli F, Salgarelli L 2009 Tunnel Hunter: Detecting application-layer tunnels with statistical Finger printing *Computer Network: The International Journal of Computer and Telecommunications Networking Archive* **53** 81-97

[5] Perdiscia R, Leea W, Feamster N 2010 Behavioral Clustering of HTTP-Based Malware and Signature Generation Using Malicious Network Traces *Proc of the 7th USENIX conference on Networked systems design and implementation* 391-404

[6] Crotti M, Dusi M, Gringoli F, Salgarelli L 2007 Detecting Http Tunnels with Statistical Mechanisms *Proc of the IEEE International Conference on Communications* 6162-8

[7] Roesch M 1999 Snort Lightweight Intrusion Detection for Networks *Proc of the LISA 99 System Administration Conference* 229-38

[8] Borders K, Prakash A 2004 Web Tap Detecting Covert Web Traffic *Proc of the 11th ACM Conf Computer and Communications Security* 110-20

[9] Pack D 2002 Detecting HTTP Tunnelling Activities *Proc of the IEEE Workshop on Information Assurance and Security* 1-8

[10] Xh X, Zl L 2005 A method for situation assessment based on D-S evidence theory *Electronics Optics & Control* **12** 36-7

[11] Li Y, Cai Y Z, Yin R P 2008 Support Vector Machine Ensemble Based on Evidence Theory for Multi-Class Classification *Journal of Computer Research and Development* **45** 571-8

[12] Bai P, Zhang X B, Zhang B 2008 Support Vector Machine and Its Application in Mixed Gas Infrared Spectrum Analysis *Xi'an Xidian university press*

[13] Platt J C 1998 Fast training of support vector machines using sequential minimal optimization *Advances in Kernel Methods Support Vector Machines Cambridge MA MIT Press*

[14] Platt J C 1999 Probabilistic output for support vector machine and comparisons to regularized likelihood methods *Advances in Large Margin Classifiers Cambridge MA MIT Press*

[15] Gu Guofei, Porras P, Yegneswaran V, Fong M 2007 Bot Hunter Detection malware infection through IDS-driven dialog correlation *Proc of the 16th USENIX Security Symposium on USENIX Security Symposium* 1-16

Authors	
	<p><b>Sheng-li Liu, born in 1973</b></p> <p><b>Current position, grades:</b> professor of State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou.  <b>Scientific interest:</b> applied mathematics and information security.  <b>Publications:</b> more than 18 papers published in various journals.  <b>Experience:</b> teaching experience of 15 years.</p>
	<p><b>Xiang Gao, born in 1984</b></p> <p><b>Current position, grades:</b> doctoral student of State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou.  <b>Scientific interest:</b> information security, artificial intelligence.  <b>Publications:</b> more than 10 papers published in various journals.  <b>Experience:</b> researching experience of 7 years.</p>
	<p><b>Pan Xu, born in 1988</b></p> <p><b>Current position, grades:</b> master student of State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou.  <b>Scientific interest:</b> information security and data mining.  <b>Publications:</b> more than 3 papers published in various journals.  <b>Experience:</b> researching experience of 3 years.</p>
	<p><b>Long Liu, born in 1983</b></p> <p><b>Current position, grades:</b> doctoral student of State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou.  <b>Scientific interest:</b> applied mathematics and artificial intelligence.  <b>Publications:</b> more than 8 papers published in various journals.  <b>Experience:</b> researching experience of 7 years.</p>

# UAV trajectory optimization generation based on Pythagorean hodograph curve

Yi Zhang, Xiuxia Yang\*, Weiwei Zhou

Department of Control Engineering, Naval Aeronautical and Astronautical University, Yantai, China

Received 1 March 2014, www.tsi.lv

## Abstract

The study of autonomous planning of UAVs (Unmanned Aerial Vehicles) flyable on-line path to adapt unstructured environment and improve manoeuvring warfare capability has an important practical significance. A path planning algorithm on-line of UAV based on Pythagorean Hodograph (PH) curve is put forward, which can consider the kinematics and dynamic constraints. The effect of the key parameters on the trajectory generation are analysed, and the appropriate value range that satisfy the constraints are given. To overcome the blindness iteration, the method of improved estimation of distribution algorithm (EDA) is used to trajectory optimization. In the estimation of distribution algorithm, the probability selection mechanism of global elite individual based on interval selection is raised, which improve the speed and precision of the path generation. When the obstacle is detected by the UAV, only if the position and the direction of the interrupt point is given, the trajectory can be replanned online. Simulation results for UAV trajectory optimization generation prove the validity and practicability of the algorithm.

*Keywords:* unmanned aerial vehicle, PH curve, estimation of distribution algorithm, trajectory generation, backstepping

## 1 Introduction

Unmanned Aerial vehicles (UAVs) hold good promise for autonomously carrying out various operations in unstructured environment. However, autonomous trajectory planning is still among the most difficult and important problems for the UAVs should allow more flexibility and thus more complex applications, such as avoid collision with moving obstacles. The literature that deals with this problem is rapidly growing.

Some planning algorithms, such as rapidly-exploring random trees [1], genetic algorithm [2] and A\* algorithm [3], divide the object region or explore large areas, and find the collision avoidance path points, then the path points are smoothed to flyable path, where vehicles on a team may be regarded as obstacles, making cooperation almost an inevitable approach for the problem.

Some scholars put forward the method of applying curve to plan the whole path directly, such as Shanmugavel [4] put forward that using Dubins curve to plan multiple UAVs' paths. Dubin path is composed by circular arcs and straight line, and from the perspective of the theory of differential geometry, in order to get a smooth flight trajectory, the first two derivative of the path is existed at least, that is, the curvature of the curve is continuous, but Dubin path is  $C_1$  type curve, and is not the  $C_2$  type curve. The reference [5] gives the method of using the Clothoids curve to get path that the curvature is continuous, the curvature linearly change along the curve, but the path length is not easy to generate closed solution. Jolly [6] applies the Bezier curve to the path planning of

mobile robot; the reference [7] uses the seven order Bezier curve to plan the UAV's flight path, considering constraints of system performance such as the curvature. At present, the Pythagorean Hodograph (PH) curve is widely used in aircraft path planning online [8-10], the planning path is  $C_2$  type path that has continuous curvature, and has a lot of advantages compared to other curve. The overall curvature is small. The length, curvature and bending energy of the curve can be calculated in closed form. At the same time, the length and curvature of PH trajectory is easy to coordinate.

The references [8-10] give the method of generating PH curve path, and consider the kinematics constraint and safety constraint. When the path doesn't meet the requirements of curvature constraints, increase the tangent vectors' length of the start and the end control points of the PH curve by iteration, but this article doesn't study the optimization parameter' selection for the coordination of the curvature and the path length, and the iteration is time consuming.

In this paper, the effect of the key parameters on the trajectory generation are analysed, and the appropriate value range that satisfy the constraints are given. To overcome the blindness iteration, the method of improved estimation of distribution algorithm (EDA) is used to trajectory optimization. The concept of Estimation of Distribution Algorithm (EDA) is first proposed in 1996 and develops rapidly in about 2000 [11], which has become the cutting-edge research in the field of evolutionary computation. EDA presents a new evolutionary pattern. In traditional genetic algorithm,

\* *Corresponding author* e-mail: yangxiuxia@126.com

with the population representing a group of candidate solutions for optimization problems, each individual of the population has its own adaptive value, and then do the genetic manipulation of simulated natural evolution, such as choose, crossover and mutation, repeating the manipulation until the problem is solved. However, in EDA, there is no traditional genetic operation such as crossover and mutation, but the learning and sampling of the probabilistic model. EDA describe the distribution of the candidate solutions in space through a probability model, uses statistical learning tools from the perspective of the macro group to establish a probabilistic model to describe the distribution of solutions, and then take a random sampling of the probabilistic model to generate new populations, repeat it and achieve the optimization of the population. Traditional evolutionary algorithm achieves the population evolution based on the genetic manipulation for the every individual in the population (crossover and mutation, etc.) and establishes the mathematical model from the "micro" level. But EDAs directly describes the evolution trend of the population through the establishment of a mathematical model based on the whole population from the "macro" level in biological evolution. To satisfy the value range of the PH curve and select the optimum parameters, the probability selection mechanism of global elite individual based on interval selection is raised, which improve the speed and precision of the path generation.

**2 PH Trajectory Optimization Generation**

**2.1 PROBLEM PRESENTATION**

Let the starting pose of the  $i^{th}$  UAV is  $pose_{si}(x_{si}, y_{si}, \theta_{si})$  and the final pose of the same UAV is  $pose_{fi}(x_{fi}, y_{fi}, \theta_{fi})$ , then the path of the UAV is defined as the parametric curve  $r_i(t) = [x(t), y(t)]$  with curvature  $\kappa_i$  connecting the initial and final pose:

$$Pose_{si}(x_{si}, y_{si}, \theta_{si}) \xrightarrow{r_i(t)} Pose_{fi}(x_{fi}, y_{fi}, \theta_{fi}).$$

In this case Pythagorean Hodograph curve in parametric form subjected to the following constraints:

- 1)  $|\kappa_i(t)| \leq \kappa_{max} 10$ , where  $\kappa_{max}$  is the maximum curvature attainable by the UAVs.
- 2) The vehicles must keep safe distance to avoid inter collision while tracing their corresponding trajectories. This can be written in mathematical form as follow:  $R_{si} \cap R_{sj} = \varnothing$ .

That is, the intersection of the safety circles of radius  $R_{si}$ ,  $R_{sj}$  corresponding to  $i$ -th and  $j$ -th UAVs must be empty.

- 3) The vehicles must keep safe distance to avoid collisions with known obstacles while tracing their corresponding trajectories. This can be written in mathematical form as follow:

$d(obstacle, UAV) \geq R_{obstacle} + R_s$ , where  $d(obstacle, UAV)$  is the distance between the centre of the circle enclosing the obstacle and the centre of the safety circle of the UAV.  $R_{obstacle}$  is the radius of the circle enclosing the square obstacle.

**2.2 USING PH TRAJECTORY GENERATION AND ITS ANALYSIS**

Reference [10] show that the lowest order of the PH path that has a point of inflection is the fifth, called the quintic PH. The presence of an inflection point allows the path to have more flexibility so that the path can easily be manipulated. Hence, the quintic PH curve is used for path planning, in which six control points are needed.

Supposing the path is  $s(t) = \int_{t_1}^{t_2} |r'(t)| dt = \int_{t_1}^{t_2} \sqrt{x'(t)^2 + y'(t)^2} dt$ .

Selecting appropriate  $\sigma(t)$  to satisfy  $\sigma^2(t) = x'^2(t) + y'^2(t)$ .

Then the length of the path is

$$s(t) = \int_0^{t_1} |\sigma(t)| dt \tag{1}$$

We can produce PH curve by selecting appropriate  $u(t), v(t), w(t)$  to construct  $\dot{x}(t), \dot{y}(t)$  and satisfy Equation (1). Therefore

$$x'(t) = w(t)[u^2(t) - v^2(t)], y'(t) = 2w(t)u(t)v(t),$$

$$\sqrt{x'^2(t) + y'^2(t)} = \sqrt{w(t)\{u^2(t) + v^2(t)\}}^2$$

$$\sigma(t) = w(t)[u^2(t) + v^2(t)]$$

where  $w(t) = 1$ . The Bernstein form of  $u(t), v(t)$  is

$$u(t) = \sum_{k=0}^2 u_k \binom{2}{k} t^k (1-t)^{2-k}, t \in [0, 1],$$

$$u(t) = u_0(1-t)^2 + 2u_1(1-t)t + u_2t^2 10.$$

$$v(t) = \sum_{k=0}^2 v_k \binom{2}{k} t^k (1-t)^{2-k}, t \in [0, 1],$$

$$v(t) = v_0(1-t)^2 + 2v_1(1-t)t + v_2t^2.$$

Considering the numeric stability, PH curve can be give by Bézier curve. The curve is

$$r(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \sum_{k=0}^5 P_k \binom{5}{k} (1-t)^{5-k} t^k, \tag{2}$$

where  $P_k = (x_k, y_k), k=0,1,2,3,4,5$  are control points, which have the following relation:

$$P_1 = P_0 + \frac{1}{5}(u_0^2 - v_0^2, 2u_0v_0). \tag{3}$$

$$P_2 = P_1 + \frac{1}{5}(u_0u_1 - v_0v_1, u_0v_1 + u_1v_0). \tag{4}$$

$$P_3 = P_2 + \frac{1}{5}(u_1^2 - v_1^2, 2u_1v_1) + \frac{1}{15}(u_0u_2 - v_0v_2, u_0v_2 + u_2v_0). \tag{5}$$

$$P_4 = P_3 + \frac{1}{5}(u_1u_2 - v_1v_2, u_1v_2 + u_2v_1). \tag{6}$$

$$P_5 = P_4 + \frac{1}{5}(u_2^2 - v_2^2, 2u_2v_2). \tag{7}$$

Giving the initial and final tangent vector length  $(\Delta x_s, \Delta y_s), (\Delta x_f, \Delta y_f), P_1, P_4$  can be determinate.

$$r(0) = P_0 = (x_s, y_s), \tag{8}$$

$$r'(0) = 5(P_1 - P_0) = [\Delta x_s, \Delta y_s], \tag{9}$$

$$r(1) = P_5 = (x_f, y_f), \tag{10}$$

$$r'(1) = 5(P_5 - P_4) = [\Delta x_f, \Delta y_f], \tag{11}$$

$$P_1 = P_0 + \frac{1}{5}[\Delta x_s, \Delta y_s], \tag{12}$$

$$P_4 = P_5 + \frac{1}{5}[\Delta x_f, \Delta y_f]. \tag{13}$$

According to reference [10]:

$$\begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \pm \begin{bmatrix} \frac{\sqrt{\varepsilon_s + \Delta x_s}}{2} \\ \text{sign}(5\Delta y_s) \sqrt{\frac{\varepsilon_s - \Delta x_s}{2}} \end{bmatrix}, \tag{14}$$

$$\begin{bmatrix} u_2 \\ v_2 \end{bmatrix} = \pm \sqrt{\frac{5}{2}} \begin{bmatrix} \sqrt{\varepsilon_f + \Delta x_f} \\ \text{sign}(\Delta y_f) \sqrt{\varepsilon_f - \Delta x_f} \end{bmatrix}, \tag{15}$$

$$\begin{bmatrix} u_1 \\ v_1 \end{bmatrix} = -\frac{3}{4} \begin{bmatrix} u_0 + u_2 \\ v_0 + v_2 \end{bmatrix} \pm \sqrt{\frac{1}{2}} \begin{bmatrix} c + a \\ \text{sign}(b) \sqrt{c - a} \end{bmatrix}, \tag{16}$$

where  $\varepsilon_s = \sqrt{\Delta x_s^2 + \Delta y_s^2}, \varepsilon_f = \sqrt{\Delta x_f^2 + \Delta y_f^2},$

$$a = \frac{9}{16}(u_0^2 - v_0^2 + u_2^2 - v_2^2) + \frac{5}{8}(u_0u_2 - v_0v_2) + \frac{15}{2}(x_4 - x_1),$$

$$b = \frac{9}{8}(u_0v_0 - u_2v_2) + \frac{5}{8}(u_0v_2 - u_2v_0) + \frac{15}{2}(y_4 - y_1),$$

$$c = \sqrt{a^2 + b^2}$$

Therefore  $P_2, P_3$  can be determinate.

From the generation of the PH trajectory, it can be seen that if the start and the final pose are known, the tangent length  $\varepsilon_s, \varepsilon_f$  of the start and the end points can determine the UAV trajectory. The trajectory has continuous curvature. From equation (3) to (16) it can be seen that after the pose of the initial control point  $pose_s(x_s, y_s, \theta_s)$  and the final control point  $pose_f(x_f, y_f, \theta_f)$  and the tangent length  $\varepsilon_s, \varepsilon_f$  are given, four paths will be generated. Select the path with minimum energy, and justify whether the path is satisfy the curvature constrain. If not, justify the tangent vector length  $\varepsilon_s, \varepsilon_f$ . The bending energy is calculated by the following expression:

$$E = \int \kappa^2(t) \sigma(t) dt \tag{17}$$

The curvature of the UAV path is:

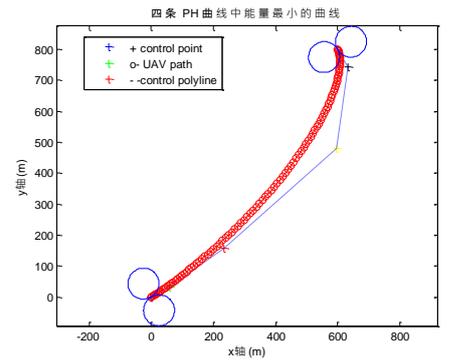
$$\kappa(t) = \frac{|\dot{r}(t) \times \ddot{r}(t)|}{|\dot{r}(t)|^3} = \frac{2[u(t)\dot{v}(t) - v(t)\dot{u}(t)]}{w(t)[u(t)^2 + v(t)^2]^2}. \tag{18}$$

In references [8-10], the parameter  $\varepsilon_s, \varepsilon_f$  are selected by iteration, which is time consuming and the optimize parameters cannot be get. On the following, through analysing the relationship between tangent vector length  $\varepsilon_s, \varepsilon_f$  and the path performance under the different initial and the final position and the tangent line direction, the parameters value range are given.

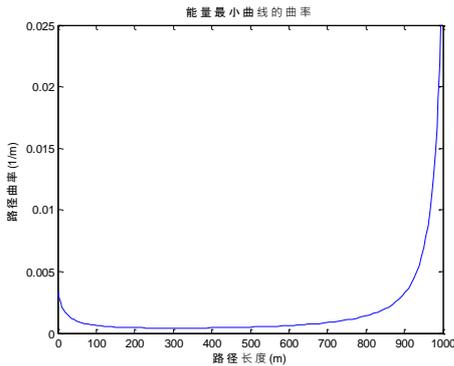
Suppose the initial and the final control points are:  $(P_s, \theta_s) = (0, 0), (P_f, \theta_s) = (600, 800)$ . The distance between the two points is:  $|P_0P_5| = 1000$ . Suppose the maximum curvature of the planning trajectory is 0.02.

TABLE 1 The effect of different tangent vector length and initial/final angle to the curvature

tangent vector length	$\frac{1}{15}d$	$\frac{1}{10}d$	$\frac{1}{5}d$	$\frac{1}{3}d$	$\frac{1}{2}d$	$\frac{2}{3}d$
initial/final angle (°)						
0/120	0.0235	0.0118	0.0034	0.0038	0.0076	0.0157
30/120	0.0249	0.0128	0.0042	0.0056	0.0129	0.0273
60/120	0.0270	0.0143	0.0055	0.0074	0.0144	0.0215
90/120	0.0297	0.0161	0.0067	0.0079	0.0118	0.0141
120/120	0.0324	0.0180	0.0076	0.0075	0.0092	0.0099
150/120	0.0408	0.0221	0.0111	0.0113	0.0138	0.0154



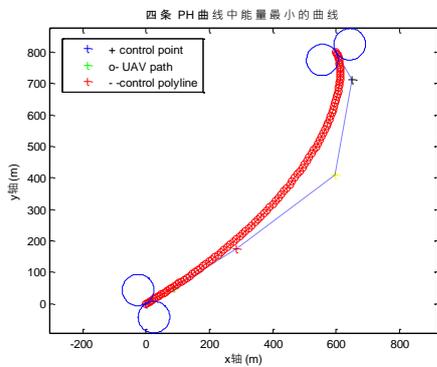
(a) The generated PH path



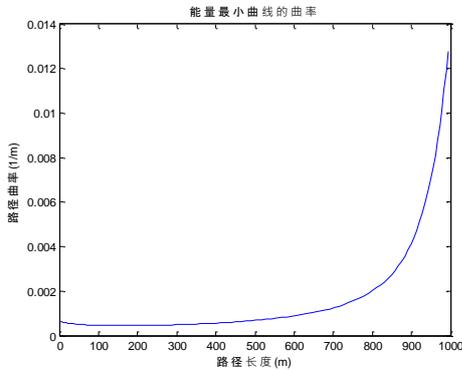
(b) The curvature of the PH path

FIGURE 1 PH path and its curvature when  $\theta_{si} = \pi/6, \theta_{fi} = 2\pi/3,$

$$\varepsilon_0 = \varepsilon_1 = \frac{1}{15}d$$



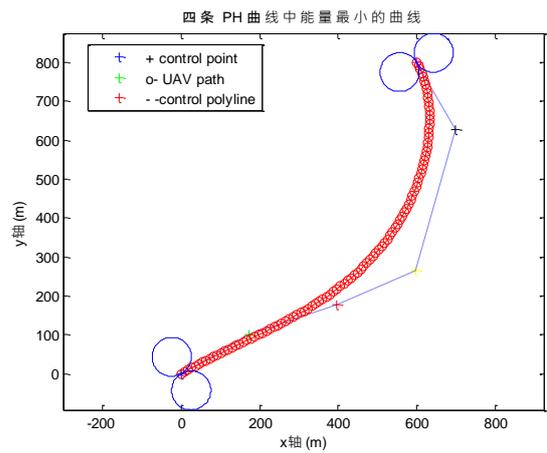
(a) The generated PH path



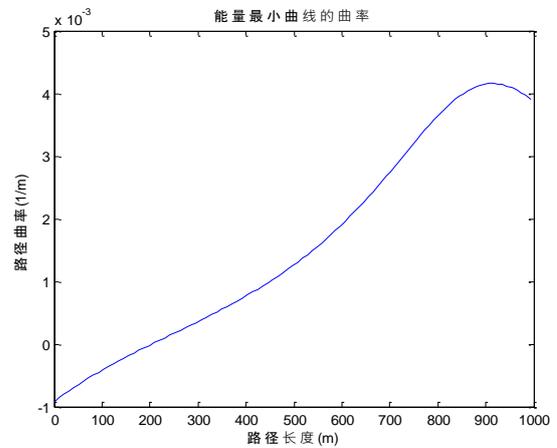
(b) The curvature of the PH path

FIGURE 2 PH path and its curvature when  $\theta_{si} = \pi/6, \theta_{fi} = 2\pi/3,$

$$\varepsilon_0 = \varepsilon_1 = \frac{1}{10}d$$



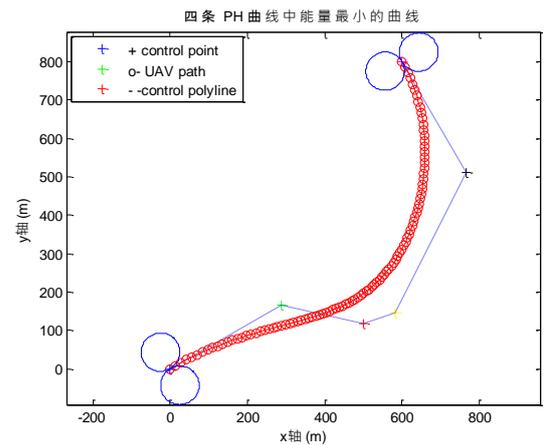
(a) The generated PH path



(b) The curvature of the PH path

FIGURE 3 PH path and its curvature when  $\theta_{si} = \pi/6, \theta_{fi} = 2\pi/3,$

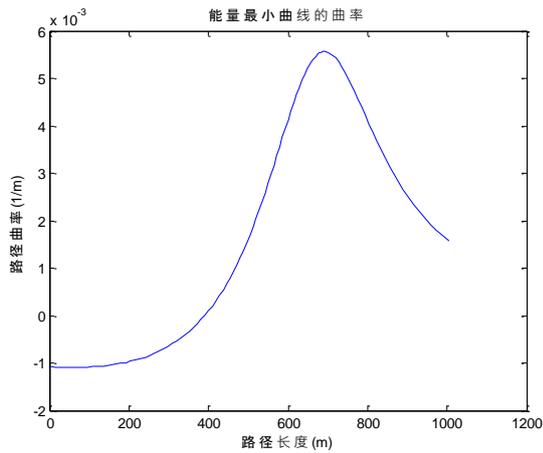
$$\varepsilon_0 = \varepsilon_1 = \frac{1}{5}d$$



(a) The generated PH path

FIGURE 4 PH path and its curvature when  $\theta_{si} = \pi/6, \theta_{fi} = 2\pi/3,$

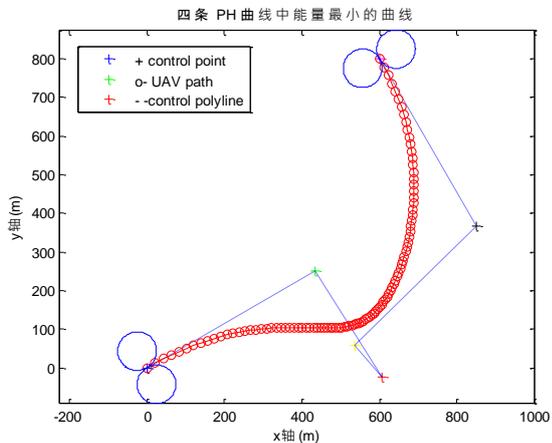
$$\varepsilon_0 = \varepsilon_1 = \frac{1}{3}d$$



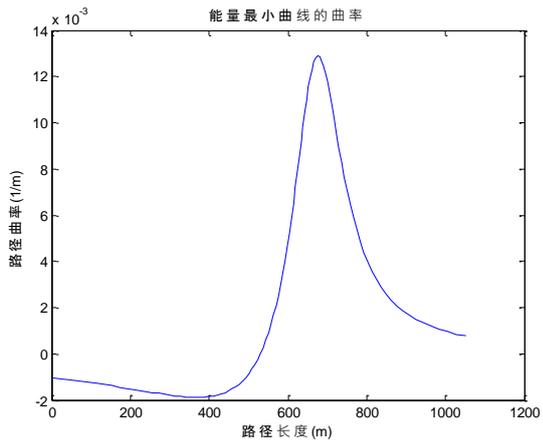
(b) The curvature of the PH path

FIGURE 4 PH path and its curvature when  $\theta_{si} = \pi/6, \theta_{fi} = 2\pi/3,$

$$\varepsilon_0 = \varepsilon_1 = \frac{1}{3}d$$



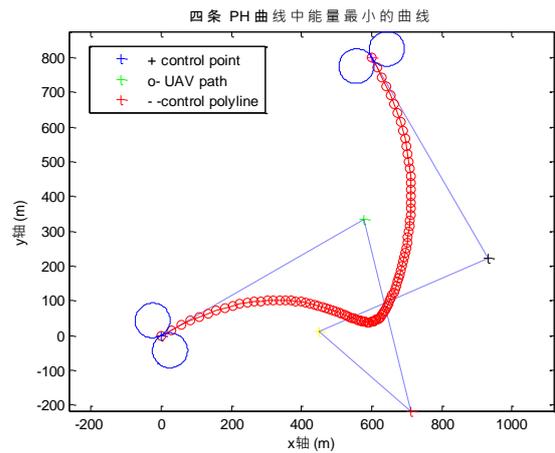
(a) The generated PH path



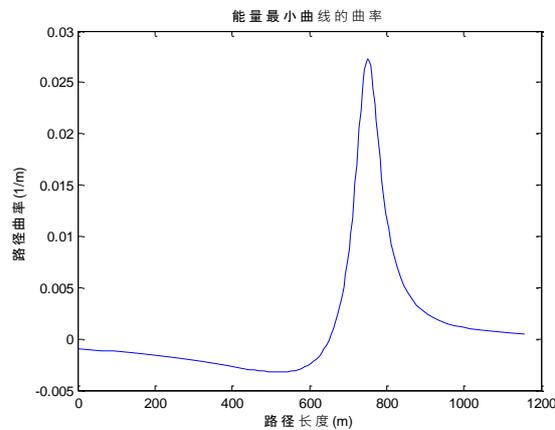
(b) The curvature of the PH path

FIGURE 5 PH path and its curvature when  $\theta_{si} = \pi/6, \theta_{fi} = 2\pi/3,$

$$\varepsilon_0 = \varepsilon_1 = \frac{1}{2}d$$



(a) The generated PH path



(b) The curvature of the PH path

FIGURE 6 PH path and its curvature when  $\theta_{si} = \pi/6, \theta_{fi} = 2\pi/3,$

$$\varepsilon_0 = \varepsilon_1 = \frac{2}{3}d$$

For convenience, set the tangent vector length  $\varepsilon_0, \varepsilon_1$  is proportional to the distance  $d$  of the initial and the final points where  $d = |P_0P_5|$ . Similarly, to get the affection of the initial and the final heading angle to the PH path curvature, set the heading angle at some step, and the simulation results are shown in table 1.

Suppose the initial/final angle is  $30/120^\circ$ , from Figures 1-5, the minimum bending energy PH curve with different tangent vector length and the corresponding curvature are shown.

It can be seen that when the tangent vector length is short, the curvature of the PH curve is very big and not easy to satisfy the constrains, which is shown in Figure 1

( $\varepsilon_0 = \varepsilon_1 = \frac{1}{15}d$ ). With the increase of the tangent vector

length, such as  $\varepsilon_0 = \varepsilon_1 = \frac{1}{5}d$  (in Figure 3), the curvature

is small, which can satisfy the demand of the performance. When increase the tangent vector length

continuously, such as  $\varepsilon_0 = \varepsilon_1 = \frac{1}{3}d$  (in Figure 4), the

peak appear in the curvature curve, accordingly, the bending energy and the length increase. From the above analysis, it shows that the tangent vector length  $\varepsilon_s, \varepsilon_f$  should be selected appropriately in some interval. At the same time, with the different initial and final angle  $\varepsilon_s, \varepsilon_f$  may be selected differently. From above it show that  $\varepsilon_s, \varepsilon_f$  can be selected in the interval  $\varepsilon \in \left[ \frac{1}{15}d, \frac{2}{3}d \right]$ .

Under the different combination of the initial and the final tangent vector length, the PH path maximum curvature change with the initial heading angle increase at the step  $\pi/6$  is shown in Figure 7, where the final angle is  $2\pi/3$ . The initialization condition of the best PH curve can be found from Figure 7. At the same time, the heading angle that satisfy the curvature constrain can also be found from the different initial and the final tangent vector length.

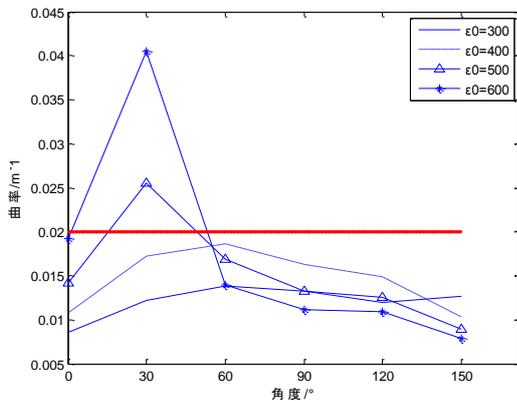


FIGURE 7 PH path maximum curvature change with the initial heading angle increase

### 2.3 PATH OPTIMIZATION GENERATION BASED ON ESTIMATION OF DISTRIBUTION ALGORITHM

In order to make full use of the performance of PH curve, and overcome the defects of blind iteration in the present literature, the method for generating and optimizing flight path based on distribution estimation algorithm is put forward.

EDAs use probabilistic modelling of promising solutions to estimate a distribution over the search space, which is then used to produce the next generation by sampling the search space according to the estimated distribution. An algorithmic framework of most EDAs can be described as follows:

#### Framework of EDA

```

Pop InitializePopulation()
while Stopping criteria are not satisfied do
  Popsel = Select(Pop)/*Selection*/
  Prob = Estimate(Popsel) /*Estimation*/
  Pop = Sample(Prob)/*Sampling*/
endwhile
    
```

An EDA starts with a solution population Pop and a solution distribution model Prob. The main loop consists of three principal stages. The first stage is to select the best individuals (according to some fitness criteria) from the population. These individuals are used in the second stage in which the solution distribution model Prob is updated or recreated. The third stage consists of sampling the updated solution distribution model to generate new solutions offspring.

Estimation of distribution algorithm is suitable especially for the situation of this article, which needs to find the feasible range of  $\varepsilon_s, \varepsilon_f$  first, which meets the dynamic performance constraints and then optimize the parameters. On the following, the probability selection mechanism of global elite individual based on interval selection is introduced to the EDA to adapt the interval optimization of the PH trajectory planning. The improved EDA is applied to optimize parameters of the flight path, in order to complete the manoeuvre in high dynamic environment with the threats.

#### 1) Performance Index

In order to get the shorter path and has strong manoeuvrability, considering both the length  $s$  and curvature  $\mathcal{K}$ , and the curvature of path reflects by bending energy  $E$ , so the performance index is defined as:

$$J = \min(\lambda_1 \times s + \lambda_2 \times E), \tag{19}$$

where the bending energy  $E$  is calculated by the equation (17), the length of path is calculated by the Equation (1).

#### 2) Gene coding

The real values  $\varepsilon_s$  and  $\varepsilon_f$  are applied to coding the individual gene in estimation of distribution algorithm. Suppose the distance between initial point to final point is  $|p_s p_f|$ , the range of tangent vector length of the starting

and ending points are set to  $\varepsilon_s \in \left[ \frac{1}{15}|p_s p_f|, \frac{2}{3}|p_s p_f| \right]$ ,

$\varepsilon_f \in \left[ \frac{1}{15}|p_s p_f|, \frac{2}{3}|p_s p_f| \right]$ , which can satisfy the

constraints of system dynamic performance, and doesn't make the length of path and bending energy too large.

Using this heuristic information, the number of iterations can be reduced when generate and optimize the trajectory.

#### 3) Individuals Chaotic Initialization

In order to improve the population diversity, initialize the individuals using chaotic variables.

The Logistic mapping chaotic model is induced:

$$r_{d+1} = \mu r_d(1-r_d), \tag{20}$$

where  $r_d \in (0,1)$ ,  $\mu$  is the control parameter. When  $\mu \in (3.56, 4.0)$ , equation (20) is in a chaotic state. Take  $\mu = 3.6$ , the coverage of chaotic sequence can meet the requirements of the population diversity.

Mapping the chaotic variables to the range of the decision variables  $(x_{j\_min}, x_{j\_max})$ , the  $j$ -th component of the  $i$ -th individual  $X_{ij}$  can be get:

$$X_{ij} = x_{j\_min} + (x_{j\_max} - x_{j\_min})r_{d+1}, \tag{21}$$

where  $j = 1, 2, \dots, M$ ,  $M$  is the number of the variables,  $i = 1, 2, \dots, N$ ,  $N$  is the number of the individuals.

4) Establishment of the Probability Estimation Model and Generation of the New Population

Divide the range  $[a_j, b_j]$  of all variables  $x_j$  into equal span  $[a_{jk}, b_{jk}]$ , where  $k$  is the  $k$ th search interval,  $k = \{1, 2, \dots, n\}$ ,  $n$  is the interval number of division. Select the elite individuals at the probability  $p_e$ . The next generation individuals' genes are generated according to the probability of the elite individual in each interval  $[a_{jk}, b_{jk}]$ . That is, suppose the probability of the elite individual in the  $k$ th search interval of the  $j$ th dimension search variable is  $p_{jk}$  (shown in Figure 8)

$$p_{jk} = \frac{\text{The number of elite individual contained in Kth interval of locus } j}{N}. \tag{22}$$

When new individuals are generated, create random number  $p^j$  in  $(0, 1)$  for each gene. If  $\sum_{k=1}^s p_{jk} < p^j \leq \sum_{k=1}^{s+1} p_{jk}$  ( $s = \{1, 2, \dots, n-1\}$ ), select the values within the  $(s+1)$ th interval  $[a_{js}, b_{js}]$  to generate the new genes. If  $p^j \leq p_{j1}$ , select the values within interval  $[a_{j1}, b_{j1}]$  to generate the new genes. Thus, the next generation population will be generated at a rather large probability in the interval with more excellent individual.

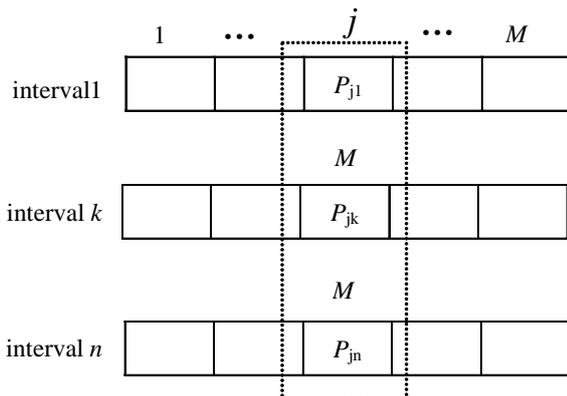


FIGURE 8 Probability value of the elite individual locus in every search interval

5) Individual Diversity Adjustment Based on Interval Information Entropy

a) Diversity calculation of each locus

The interval information entropy theory is used to define the population diversity. The  $j$ -th locus information entropy  $H_j(N)$  is defined as

$$H_j(N) = -\sum_{k=1}^n p_{jk} \log p_{jk} \tag{23}$$

$H_j(N)$  is the overall information measure of the  $j$ th locus. From the definition of information entropy, the greater of  $H_j(N)$  the better of the locus diversity. Population information is abundant in the early evolutionary, which may achieve the maximum entropy  $\log|n|$ . With the increase of the evolution generation  $t$  and the improvement of the fitness, the population diversity will reduce gradually. When the individuals in the population achieve consistent, the diversity is zero.

b) Population diversity calculation and adjustment.

The threshold for diversity of each generation is defined as:

$$\varphi(t) = A \cdot (\log|n|) \cdot e^{-\frac{Bt}{G}}, \tag{24}$$

where  $A \in (0,1)$  is regulating coefficient,  $B > 0$  is accelerating factor,  $t$  is the current evolution generation,  $G$  is the total evolution generation.

Justify the population diversity in the process of evolution. If the diversity is below to the threshold given by equation (23), the chaos substitution needs to be adopted. The detail operations are:

i)  $M$  individuals are selected randomly from the population with probability  $P_\alpha$ , where  $P_\alpha \in (0,1)$ .

ii) The chaos substitution were adopted. Low-diversity gene loci  $j$  of  $X_i^t$  was chosen, a new gene was generated according to the Equation (20) after normalization, then chaos substitution was carried out with the probability  $p_c$ .

6) Vaccination Based on the System Information

In order to make full use of the system information, the immune evolutionary mechanism is added. Extract the vaccine at first, and according to the system information and the heuristic information, such as if initial boundary curvature does not meet the requirement of the curvature's constraints, the value of  $\varepsilon_s$  should be increased, and inoculate it with the probability  $P_u$ . Reserve the best individual of each generation.

7) Gauss Perturbation Optimization

In order to find the optimal individual, Gauss perturbation is adopted for the neighbourhood optimization of the best individual. The variable perturbation is carried out as follows:

$$x'_{b,j} = x_{b,j} + N(0, \sigma), \tag{25}$$

where  $x_{b,j}$  is the  $j$ th variable of the optimal individual,  $x'_{b,j}$  is the individual values of variables after perturbation,  $N(0, \sigma)$  is the amount of Gaussian distribution,  $\sigma$  is the variance,  $N(0, \sigma)$  may be replaced

by other random distribution. If the fitness of the individual  $x_{b,j}$  exceeds  $x_{b,j}$ , the perturbation is carried out. Retain the original optimal individual.

#### 8) Performance analysis of the algorithm

a) The improved algorithm is convergent by probability 1, and the convergence speed is accelerated.

For the optimal reservation mechanism in the algorithm, although the other operators are added, the convergence of the algorithm is not affected. Based on the certification of literature [12], the convergence probability of this algorithm is 1. The global elite individual probability selection based on EDA can effectively improve the convergence speed. After the vaccination operation, the system information can be fully used to speed up the evolution towards the optimal solution.

b) The global optimal solution is appropriate to get.

The loci interval information entropy is used to judge the population diversity and it can more accurately reflect the information of the individual, and it can obtain the individual that satisfies the evolution rule by the adjustment of the threshold in the process of evolution, which is more meet the requirements of the evolution than using the fixed density individual choice. Chaos initialization and Gauss perturbation optimization make the search space increase and the probability of local optimum decrease.

#### 9) Step of the Algorithm

a) Determine the range of the parameters and decoding them with real value. Divide the interval, using the chaotic initialization and perform individual vaccination with probability  $p_u$ .

b) Calculate the interval probability estimation model of elite individuals, and generate the new population. Calculate locus information entropy of each individual and adjust the low diversity genes using chaotic replace operation.

c) Using immune mechanics, perform individual vaccination with probability  $p_i$ .

d) For the individuals satisfying the dynamic constrains, calculate the fitness  $J$  and preserve the best individual  $X_b$ . Optimize the best individual by Gauss perturbation for the neighbourhood. If the best fitness keeps constant in  $g$  generations or arrives the max generation, then output the best individual and stop, else return to step b).

### 2.4 COLLISION AVOIDANCE PH TRAJECTORIES GENERATION

When UAV detects the unknown moving threat or obstacle and it need to be avoided, the interrupt point of the trajectory should be given. The UAV will replan the trajectory according to the velocity and the position of the obstacles. After the UAV velocity turn angle and the interrupt point coordinate are determined, the new

trajectory can be generated between the obstacle detected point and the interrupt point. Then take the interrupt point as the initial point and the object or the next interrupt point as the final point, the next trajectory can also be generated. To keep the continuous of the trajectory, on the detect point and interrupt point, the direction of the tangent line keep unchanged. Through adjusting  $\varepsilon_s, \varepsilon_f$  of each trajectory, the flyable trajectory that satisfy the performance constrains can be given.

### 3 Simulation Results

1) The optimization PH trajectory generation based on EDA

Supposing UAV sets up from the initial point  $P_s$  (0, 0) to the object point  $P_f$  (200, 500). The initial

angle  $\theta_s$  is  $\frac{\pi}{6}$  and the final angle  $\theta_f$  is  $\frac{5}{4}\pi$ . The

maximum curvature constrain is  $k_{\max} = \pm 0.02$ .

Estimation of distribution algorithm is used to generate the trajectory. After normalizing the curvature and the trajectory length, the weight coefficient is set to  $\lambda_1 = 0.008, \lambda_2 = 40$ . Determining the parameter range of every gene, the elite individuals selection probability

$p_e = \frac{1}{3}$ , the diversity select threshold coefficient are set to  $A=0.8, B=0.8$ . The low diversity individual probability  $P_a=0.6$ . The chaotic replacement probability  $p_c = 0.5$ . The inoculation probability  $P_i = 0.5$ . The initialization individuals gene inoculation probability  $p_u = 0.6$ . The Gauss perturbation variance  $\sigma = 0.1$ .

Set the population individuals number  $N=10$ , the evolution generation  $G=10$ . The tangent vector length is

selected as  $\varepsilon_s \in \left[ \frac{1}{15} |p_s, p_f|, \frac{2}{3} |p_s, p_f| \right]$ ,

$\varepsilon_f \in \left[ \frac{1}{15} |p_s, p_f|, \frac{2}{3} |p_s, p_f| \right]$ . The optimization trajectory is

generated and the corresponding optimized parameters  $\varepsilon_s = 97.3$ ,  $\varepsilon_f = 98.2$ , the path length  $s = 548.63$ , the maximum curvature  $\bar{k}_{\max} = 0.0111$ .

Set the evolution generation  $G=5$ , the simulation results are shown from Figure 9 to 11. The optimized parameter  $\varepsilon_s = 88.3$ ,  $\varepsilon_f = 99.8$ ,  $s = 546.6$ ,  $\bar{k}_{\max} = 0.013$ . Compared with the simulation results of  $G=10$ , the near optimal trajectory is already found. For the online planning, the optimal trajectory is not needed, while the iteration times is reduced largely, which show that the flyable trajectory can be generated in a short time.

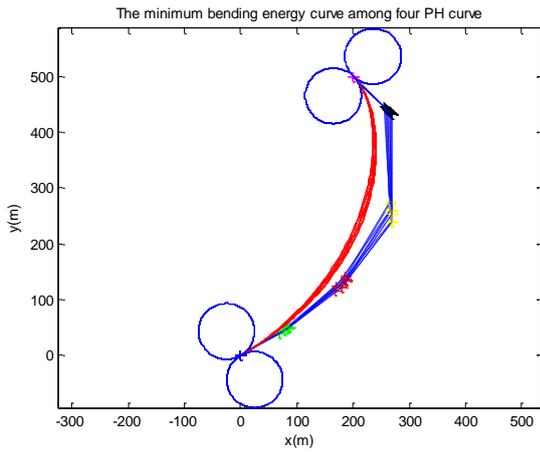


FIGURE 9 N Trajectories of the Last Generation Individuals

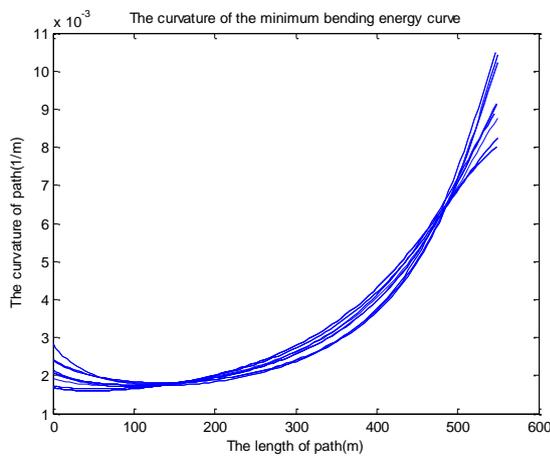


FIGURE 10 Curvature Change with the Length of the Trajectories in Figure 9

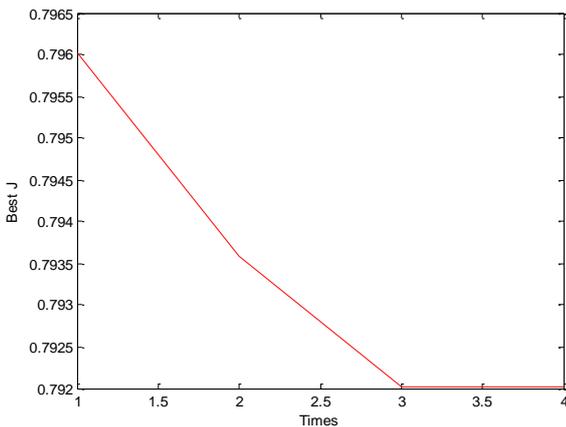


FIGURE 11 The Optimum Individual Fitness Evolution Curve

2) Collision Avoidance Trajectories Generation Online

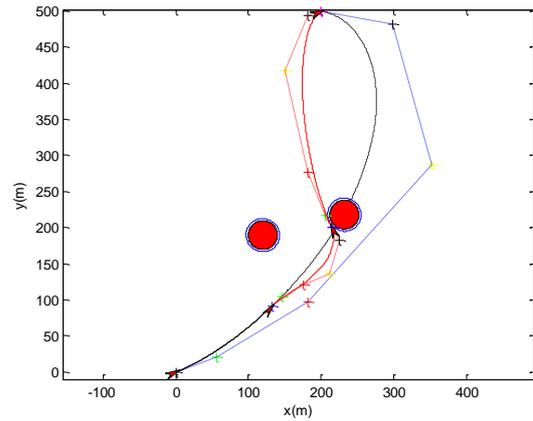


FIGURE 12 UAV planning Trajectory Based on PH Curve

It is assumed that the starting point  $P_s$  and the final point  $P_f$  are  $(0, 0)$  and  $(200, 500)$ . The heading angle  $\theta_s, \theta_f$  are  $\frac{\pi}{9}$  and  $\frac{17}{18}\pi$ . Suppose the detect distance  $S_t$  is 100. The average UAV speed  $V_u$  is 50. On the location  $(132, 90)$  of the PH path, the obstacle is detected. The radius of the barrier's circle is 20, the movement velocity is  $V_o = 40$ , the line-of-sight angle between UAV and obstacle is  $\theta_o = 6.6^\circ$ .

It can be judged whether UAV need to avoid obstacle according to the obstacle avoidance principle of velocity. The coordinates of the insertion point are  $(215, 200)$  and the angle that UAV needs to counter clockwise turn is  $\alpha = 76.4^\circ$  by calculation at the insertion point.

Using PH curve, UAV collision avoidance trajectory is generated, which is shown in Figure 12. The UAV successfully realize the obstacle avoidance online.

4 Conclusion

Pythagorean Hodograph (PH) curve is used to plan the UAV trajectory, which is flyable and satisfy the kinematics and dynamics constrains. The effect of the key parameters on the trajectory generation are analysed, and the appropriate value range that satisfy the constrains are given. EDA is used to optimize the trajectory overcome the blindness iteration. The proposed probability selection mechanism of global elite individual based on interval selection is appropriate to the PH trajectory interval optimization, which improve the speed and precision of the path generation. The method of PH optimization generation can be generalized to three-dimension trajectory generation.

Acknowledgments

This research is supported by Aeronautical Science Foundation of China under Grant No20135584010.

References

[1] Mangal Kothari, Ian Postlethwaite 2013 A Probabilistically Robust PathPlanning Algorithm for UAVs Using Rapidly-Exploring Random Trees *Journal of Intelligent and Robotic Systems* 71(2) 231

[2] Xuan Zou, Bin Ge, Peng Sun 2012 Improved Genetic Algorithm for Dynamic Path Planning *International Journal of Information and Computer Science* 5 28-31

[3] Li Xia, Wei Ruixuan, Wang Zhike 2011 Three-dimension path planning for UAV using improved A\* algorithm in complicated threat environment *High Technology Letters* 1 13-8

[4] Madhavan Shanmugave, Antonios Tsourdos, Brian White, Rafa Zbikowski 2009 Co-operative Path Planning of Multiple UAVs Using Dubins Paths With Clothoid Arcs *Control Engineering Practice* doi:10.1016.

[5] Dai R, John E Cochran Jr, 2009 Path Planning for Multiple Unmanned Aerial Vehicles by Parameterized Corno-Spiral *American Control Conference Hyatt Regency Riverfront St. Louis MO USA* June

[6] K G Jolly, R Sreerama Kumar, R Vijayakumar 2009 A Bezier Curve Based Path Planning in a Multi-Agent Robot Soccer System Without Violating the Acceleration Limits *Robotics and Autonomous Systems* 57 23-33

[7] Armando Alves Neto, Douglas G Macharet, Mario F M Campos 2013 Feasible path planning for fixed-wing UAVs using seventh order Bézier curves *Journal of Braz Comput Soc* 19 193-203

[8] M Shanmugavel, A Tsourdos, R Zbikowski, B A White, C A Rabath, N Lechevin 2006 A solution to simultaneous arrival of multiple UAVs using Pythagorean Hodograph curves *Proceedings of the 2006 American Control Conference Minneapolis Minnesota USA* June 2006 2813-8

[9] M Shanmugavel, A Tsourdos, R Zbikowski, B A White 2007 3D path planning for multiple UAVs using Pythagorean Hodograph curves *AIAA Guidance, Navigation and Control Conference and Exhibit 20 - 23 August 2007 Hilton Head South Carolina: AIAA* 2007-6455

[10] M A Shah 2011 Cooperative path planning and cooperative perception for UAVs swarm *Cranfield University PhD Thesis*

[11] Larrinaga P, Lozano J A 2002 Estimation of Distribution Algorithms *A New Tool for Evolutionary Computation. Boston: Kluwer Academic Publishers*

[12] Yi Hong, Qingsheng Ren, Jin Zeng, Yuchou Chang 2005 Convergence of estimation of distribution algorithms in optimization of additively noisy fitness functions *17th IEEE International Conference on Tools with Artificial Intelligence* 223-7

Authors	
	<p><b>Yang Xiuxia, born in 1975, Laizhou, Shandong, China</b></p> <p><b>Current position, grades:</b> Vice professor at the department of control engineering of naval aeronautical and astronautical university</p> <p><b>University studies:</b> Ph.D. degree in electrical engineering from naval university of engineering, Wuhan, China in 2005</p> <p><b>Scientific interest:</b> nonlinear control theory with applications to robots, aircraft and other mechanical systems</p>
	<p><b>Zhang Yi, born in 1971, Rongcheng, Shandong, China</b></p> <p><b>Current position, grades:</b> Vice professor at the department of control engineering of naval aeronautical and astronautical university</p> <p><b>University studies:</b> master degree in control theory and application from naval aeronautical and astronautical university, Yantai, China in 2001</p> <p><b>Scientific interest:</b> nonlinear control theory with applications to robots, aircraft and other mechanical systems</p>
	<p><b>Zhou Weiwei, born in 1991, Nanjing, Jiangsu, China</b></p> <p><b>Current position, grades:</b> master degree student in control theory and application at naval aeronautical and astronautical university</p> <p><b>University studies:</b> bachelor degree in test and control engineering from naval aeronautical and astronautical university, Yantai, China in 2014</p> <p><b>Scientific interest:</b> nonlinear control theory with applications to robots, aircraft and other mechanical systems</p>

# Reliable UDP over the air transfer in digital radio system

**Juan Zhang\*, Hesong Jiang, Hong Jiang, Chunmei Chen**

*The Open Fund of Robot Technology Used for Special Environment Key Laboratory of Sichuan Province, School of Information Engineering, Southwest University of Science and Technology, Mianyang, China*

*Received 5 May 2014, www.tsi.lv*

---

## Abstract

Reliability is very important in digital radio point-to-point transmission system, especially for bulk data transfer in narrow band channel. Currently most applications are based on raw UDP service, which does not guarantee the reliability, and existing reliable UDP transfer protocols do not satisfy the performance expectations. The article presents R2UDP(Reduced Reliable UDP) over the air transfer suitable to radio system, SBACK (selective bundled ACK) and smart probe improves the transfer efficiency and saves the bandwidth, and also minimizes the impact of bulk data transfer to other traffic on the shared channel.

*Keywords:* R<sup>2</sup>UDP, bundled ACK, smart probe, digital radio, narrow band channel

---

## 1 Introduction

In digital radio system, some data application needs to transfer bulk data over the air in low speed underlying channel. This application does not occur often to warrant having a dedicated reversion channel. And it has the lowest priority and shares channel resource with other traffic loading. As a result, it will take a long duration for whole transfer and occupy the bandwidth for a long time, not only because the channel speed is low and packet loss probability is high, but also the transfer is interrupted often by other higher priority traffic occurred on the shared channel.

Currently methods of the data transmission mainly has TCP (transmission control protocol), SCTP (Stream Control Transmission Protocol), and the UDP (user datagram protocol). TCP and SCTP protocols are connection oriented, ensure the reliability of the data, but the processing is complex, efficiency is not high, occupy more resources, unable to support the massive concurrent connections; UDP protocol adopted for non-connection transmission, fast speed, high efficiency, and can support massive concurrent connections, but there are many shortcomings of poor reliability [1-2]. Through the comparison, the UDP protocol is a more appropriate choice to further improve the speed of data transmission.

To study and improve the reliability of UDP protocol has become a hot issue in current. The literature [3, 4] proposed the RUDP protocol for the environment of a large number of communication terminal frequently sending small size message to the dedicated server. The protocol is very similar with the TCP timeout mechanism. Although compared with TCP, the protocol is much simplified, the kinds of delay is not suitable for fast data transmission of wireless narrow-band system.

The literature [5-7] proposed the RUDP that sends a message into waiting queue and continues to send the next packet, do not receive confirmation of the message and then send a message. Compared with the literature [3-4], RUDP saves a lot of waiting time. However, it is still to confirm each message and spend a lot of resources. Lack of RUDP based on [5-7], the literature [8-13] put forward the concept of batch confirmation or timing validation BA-RUDP (Bulk Ack-Reliable UDP). The BA-RUDP algorithm is better than the previous RUDP. It has saved a lot of time and resources, but it still exists some problems: the sending end of the sending pointer is only one. Namely, the sending end each received confirmation; the sending pointer need temporarily stop the sending operation if the sending end has the need to retransmit a packet. This process is still a stop waiting process, spend considerable time.

This paper presents R2UDP (Reduced Reliable UDP) protocol for the specific application of digital wireless narrow-band system of point-to-point file transfer. The protocol creatively adds selective BACK (Bundled ACK) and smart probe, improves the transmission efficiency and save the bandwidth, but also reduces the influence of mass data on the shared channel transmission.

## 2 Problem statement

In digital radio system, some data application needs to transfer bulk data over the air in low speed underlying channel. This application does not occur often to warrant having a dedicated reversion channel. And it has the lowest priority and shares channel resource with other traffic loading.

As a result, it will take a long duration for whole transfer and occupy the bandwidth for a long time, not only because the channel speed is low and packet loss

---

\*Corresponding author e-mail: zhangjuanr@swust.edu.cn

probability is high, but also the transfer is interrupted often by other higher priority traffic occurred on the shared channel.

For example, as shown in Figure 1, the PCR OTAP (Over The Air Programming) application deployed in

remote PC server and subscriber exchanges large size configuration data through the DMR (Digital Mobile Radio) radio system, such as IPSC (IP Site Connect), single site. The system context diagram is illustrated below.

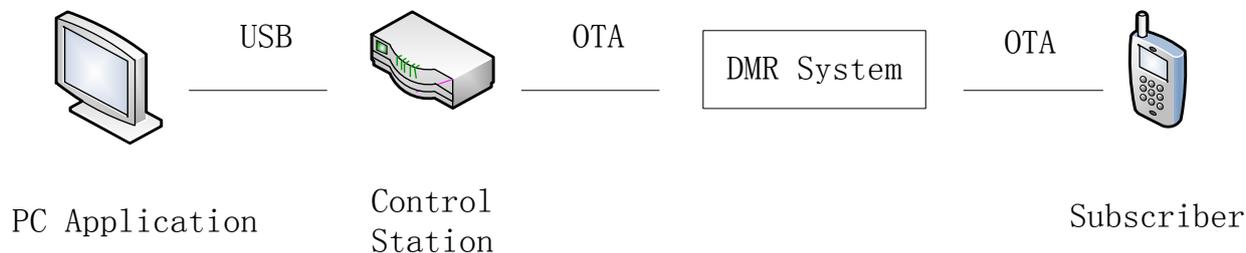


FIGURE 1 System Context Diagram

The PC application and the subscriber application are two end applications for bulk data transfer in the existing radio data application infrastructure. The control station is a L3 device to route the data between the RAN and the wired USB connection. It is transparent to application and does not expect any specific change. The channel in the RAN is low speed with less than 9600 bps data rate, it suffers from poor link quality and also shares with high priority traffic, e.g. voice call. The channel of USB connection is high speed and has good link quality.

In the radio system above, there is a critical parameter of ‘time to transfer’ between the two ends, and the bottle neck is the performance of the OTA link in the RAN. The ‘time to transfer’ over the air is impacted by many factors, besides the inherent wireless channel characteristic of high error rate due to RF impairment, the round trip delay varies in different system, and the transmission channel will become unavailable when the subscriber is engaged in voice call, or it is out of RF range, or radio user switches the channel, etc.. All of them worsen the transfer success rate.

In case of bulk data transfer, we are pursuing shorter transfer time with less bandwidth consumption, because we need to minimize the impact of long time channel occupation by this lowest priority data application. If the duration is too long, it will not meet the requirement of the application itself, and also it will impact other traffic transfer occurred on the channel.

To achieve this goal, in upper L7 application layer, the raw bulk data is compressed and fragmented, and the transfer is designed as resilience. It means that the application only needs to transfer those unsuccessful fragments in continuous sessions and does not transfer the data from the scratch, in this way the bandwidth resource is saved and the whole transfer time is shorter. Here the fragment size is balanced by the underlying channel speed and feature interaction, if it is too large, it will delay other higher priority traffic access on this channel, and this will be unacceptable.

Although the L7 handling gets benefits to some extent, the performance is still not desired. The main problem is the efficiency of reliable transfer for each fragment. So an effective transfer mechanism is expected

to transport the data over the air faster and consume the bandwidth resource as less as possible.

Current known transfer mechanisms can be found in transport layer protocols, e.g. TCP, UDP and some standard file transfer protocols, e.g. TFTP. Since the UDP does not provide reliable transfer service, it is upper layer application’s responsibility to retransmit the failed packet to guarantee the reliability. The UDP based transfer is the direction of over the air transfer in radio system because of its simplicity, lightweight, and some proprietary utility of UDP/IP header compression. The existing UDP based reliable transfer scheme is not desirable because the retransmission strategy is low efficient or aggressively clogs channel, or there is redundancy as it is heavy weight model based which is not suitable to the radio system. In consequence, the channel utilization and system throughput is decreased, and the user interface is hard to be utilized in special application scenario.

Accordingly, there is a need for a reliable UDP transfer, which is lightweight and efficiently transports the bulk data over the air in radio system.

### 3 R<sup>2</sup>UDP mechanism analysis

This paper provides an efficient reliable UDP transfer over the air. It is a balance of efficiency and complexity in radio system. R<sup>2</sup>UDP (Reduced Reliable UDP) is referred in following chapters. The security is not covered as it is supposed to be application layer’s responsibility. It has the following goals.

- 1) It will guarantee the success rate of each data UDP packet transfer, but not allow the big size data packet is loaded on the channel too fast to clog the system.
- 2) It will reduce the reliability overhead as much as possible to decrease the overall transfer time and save the bandwidth, e.g. less cost on acknowledgement traffic and handshake.
- 3) It will be lightweight and simple in radio system model.
- 4) It will provide simple and general low level user interface, like socket interface.

3.1 GENERALS

R<sup>2</sup>UDP is a ‘thin’ layer based on UDP and completely compatible with UDP packet format to embrace existing proprietary UDP/IP header compression. It attempts to provide only those services necessary, in order to be efficient in operation and small in size. It is efficient because essentially there is small overhead and retransmission intelligence.

R<sup>2</sup>UDP is deployed as 2 ends: ‘Fat’ sender and ‘thin’ receiver. One device could be both sender and receiver for full duplex reliable channel. The sender initiates bulk data transmission with UDP packets. It is the key functional part to manage the pace of retransmission. The

receiver receives the bulk data and returns ACK to indicate which packets are arrived successfully.

In the example below, the PC device is the data sender that sends bulk data over the air to the receiver of subscriber. As shown in Figure 2.

As shown in Figure 3, there are 2 symmetric data paths provided for full duplex reliable channel as below. Each path consists of bidirectional UDP packets (data and ACK) and is independent on each other. The following chapters only describe one path.

As shown in Figure4, R<sup>2</sup>UDP encapsulates the payload data and send it with raw UDP service. The transfer unit of R<sup>2</sup>UDP is called as segments as follows.

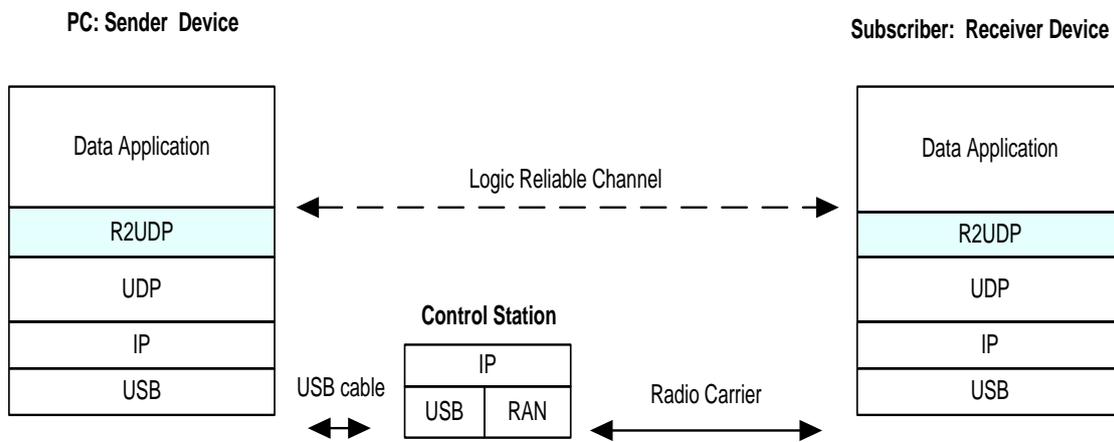


FIGURE 2 Protocol Stack Overview

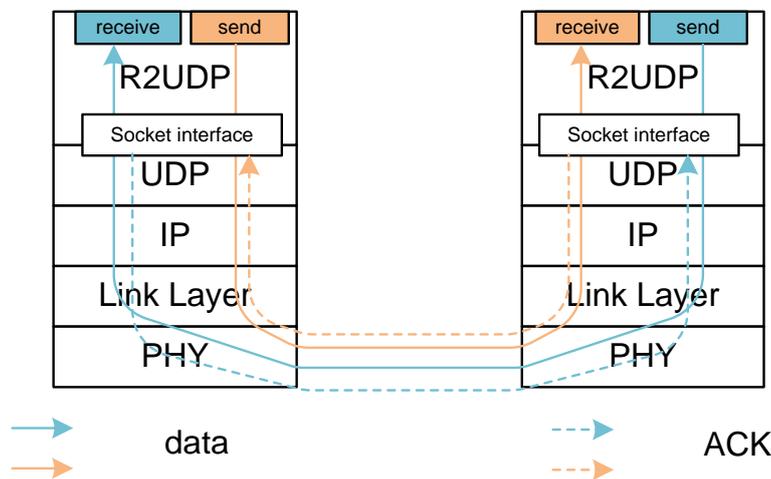


FIGURE 3 Logical Reliable Channels

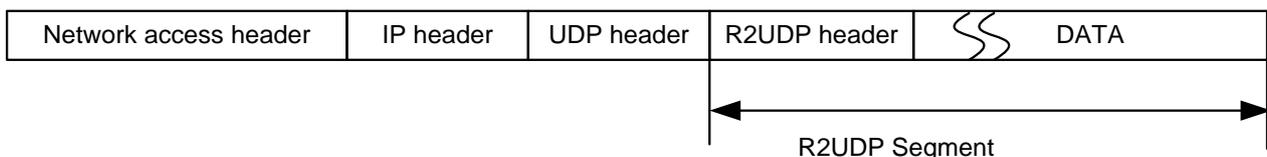


FIGURE 4 Segments in Frame

R<sup>2</sup>UDP provides efficient reliable transfer through main mechanisms here.

- 1) Manage sending buffer with several windows and data blocks. The data block in next window is only transferred when the current window transferring completes. It essentially provides more patience to avoid channel clogging.
- 2) Bundle acknowledgement for cumulative ACK and selective retransmission. It essentially reduces the traffic and increase the data throughput.
- 3) Active probe by the sender to pull in the last delivery information. In this way, the sender is capable of pacing the flow with more patience instead of put bulk data aggressively.
- 4) Reduced header format and handshake overhead.

### 3.2 RELIABLE COMMUNICATION

#### 3.2.1 Window and Block Number

As shown in Figure 5, a buffer from user will be split into several windows, and each window contains several blocks, which share one ACK. Each window engages one

window number to differentiate it with others, and each block engages one block number to identify it among blocks in the window. The initial window number and block number are 0 after data path reset which happens prior to the beginning of first transfer. The window number is increased by 1 each time a window buffer is sent successfully.

Sender is responsible for filling in window and blocks number, and make sure that a window must be successfully received by receiver before a new window starts.

Window number and block number are filled in R<sup>2</sup>UDP header, which forms a data segment with data block.

The window size can be modified during transfer according to the total data length. If the remained data is less than previous window size, e.g. there is only 4 data blocks, while the previous window size is 5, the sender updates the last window size as 4. E.g. transferring 2K bytes buffer from user, window size = 3, block size = 256 bytes. The window n+1 is only transferred upon completion of the window n.

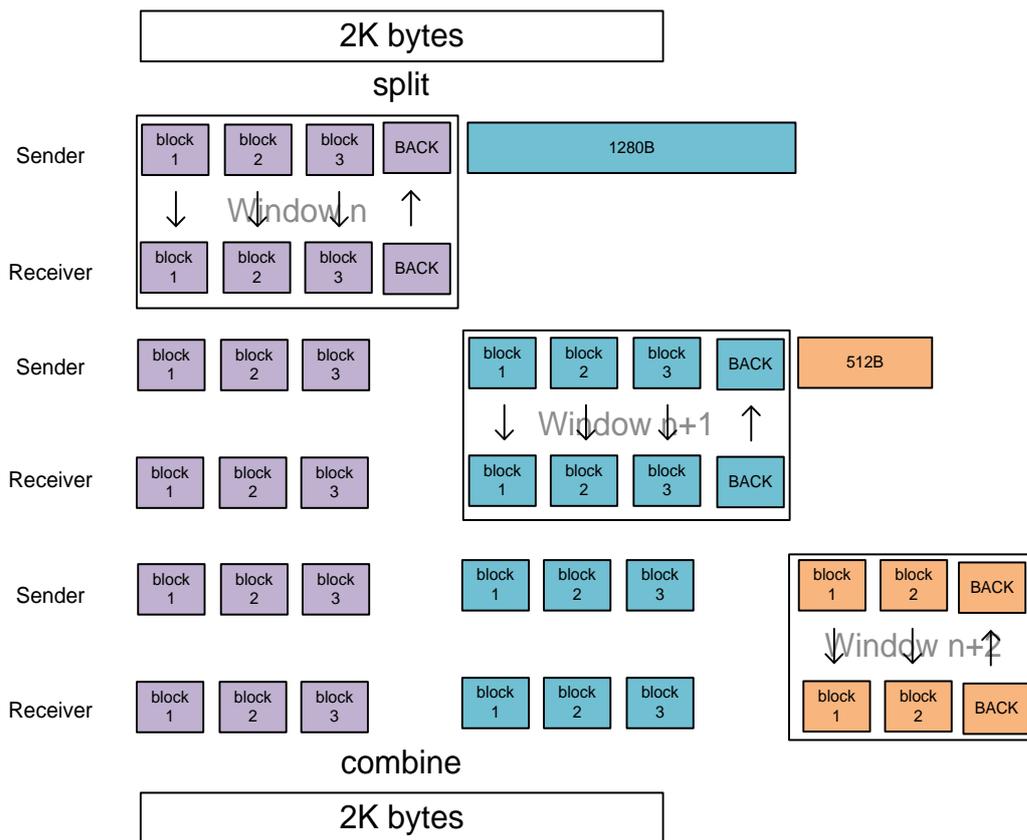


FIGURE 5 Window Management

#### 3.2.2 Bundled Acknowledgement

R<sup>2</sup>UDP assumes it has only an unreliable datagram service to deliver segments. To guarantee delivery of segments in this environment, R<sup>2</sup>UDP engages mechanism of ACK and selective retransmission.

As shown in Figure 6, here the ACK is a bundled acknowledgement for data segments belong to the same window. Receiver will update the received data block number in cumulative BACK when missing segment arrives. The ACK segment will not be acknowledged.

3.2.4 Selective Retransmission

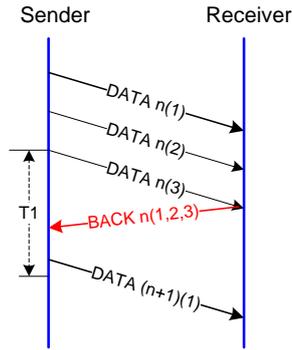


FIGURE 6 Window size = 3, normal case

3.2.3 Active Probe

To detect missing segments, the sender utilizes a retransmission timer for each window transmitted. As shown in Figure 7, the timer could be set according to the single segment transfer time in the network and the amount of segments. When an acknowledgement for a window is received, the timer for that window is cancelled. As shown in Figure 8, if the timer for a window expires before an acknowledgement is received, a PROBE segment is transmitted. Receiver will reply a BACK to identify the window number and received data blocks.

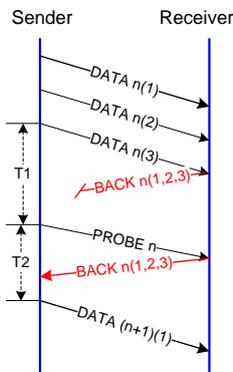


FIGURE 7 Window size = 3, BACK loss

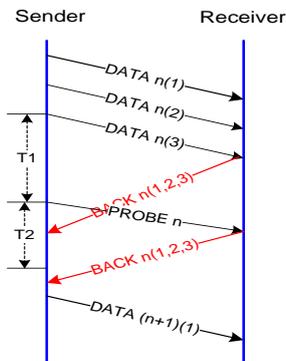


FIGURE 8 Window size = 3, BACK delay

As shown in Figure 9, sender sends out data blocks of a window. Receiver will reply a BACK once all blocks in the window are received. Otherwise if not all blocks are received, receiver will wait for the blocks not arrived, or reply a BACK for blocks arrived on receiving a PROBE. Sender retransmits only those missing data blocks and waits for the positive BACK.

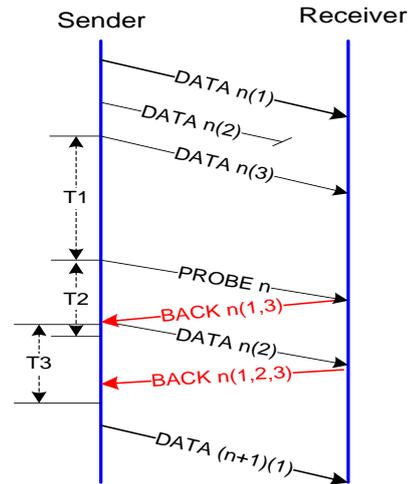


FIGURE 9 Window size = 3, 1 data segment loss

3.2.5 Reset Path

As shown in Figure 10, the data path is to reset before the first transfer. Sender is responsible to send a RST upon reset. On receiving a RST segment, receiver will discard incomplete window, and then reply a RACK to indicate its availability. Receiver will update its window number on receiving a DATA segment.

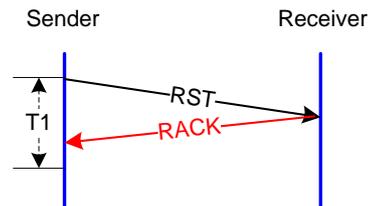


FIGURE 10 Reset a data path

4 Simulation results

We analysed the mechanism and performance of the R2UDP protocol through the Motorola radio to test and NS-2 network simulation.

4.1 TESTING ON VOICE INTERFERENCE

The testing conditions are as follows: the radio is Motorola Mag One A8, the channel bandwidth is

12.5KHZ, and the use of transferring the file is bandwidth is 9600bps, the file is 10M between the radios. Comparison of the protocol is TCP, TFTP (the bottom for the UDP protocol), and R2UDP. The total test time is 10 minutes and 5 seconds of the phone every 30 seconds in Table 1. The total test time is 10 minutes and 5 seconds of the phone every 20 seconds in Table 2.

TABLE 1 Interference 30s

Protocol	Data
TCP	4.054M
TFTP	4.163M
R <sup>2</sup> UDP	4.769M

TABLE 2 Interference 20s

Protocol	Data
TCP	3.292M
TFTP	3.856M
R <sup>2</sup> UDP	4.293M

From Table 1 and Table 2 shows that, even in the frequent case of a voice interrupted, R2UDP showed better transmission performance. TCP and TFTP transmission performance dramatically decrease with frequent access interference.

4.2 SIMULATION ANALYSIS ON NO VOICEINTERFERENCE

Through the NS-2 simulation comparative analysis of TCP, UDP, R2UDP, the specific simulation environment is shown in table 3. Given the error rate of Irate=4%, different bandwidth situation, compare the average throughput as shown in Figure 11. Given the bandwidth of 9.6K, different error rate comparison of throughput is shown in Figure 12.

As shown in Figure 11, compared to Vegas and Reno, R2UDP showed good performance in a higher rate of error (4%), a narrow bandwidth (less than 12K). As shown in Figure 12, compared to Vegas and Reno, R2UDP showed good performance with the increased error rates in a narrowband system (9.6K), while Reno and Vegas respectively in the error rate is 10% and 45% in the case of a throughput of 0.

TABLE 3 Parameter settings

Type	Value
Link delay	10ms
Packet Size	256Byte
Window size	3

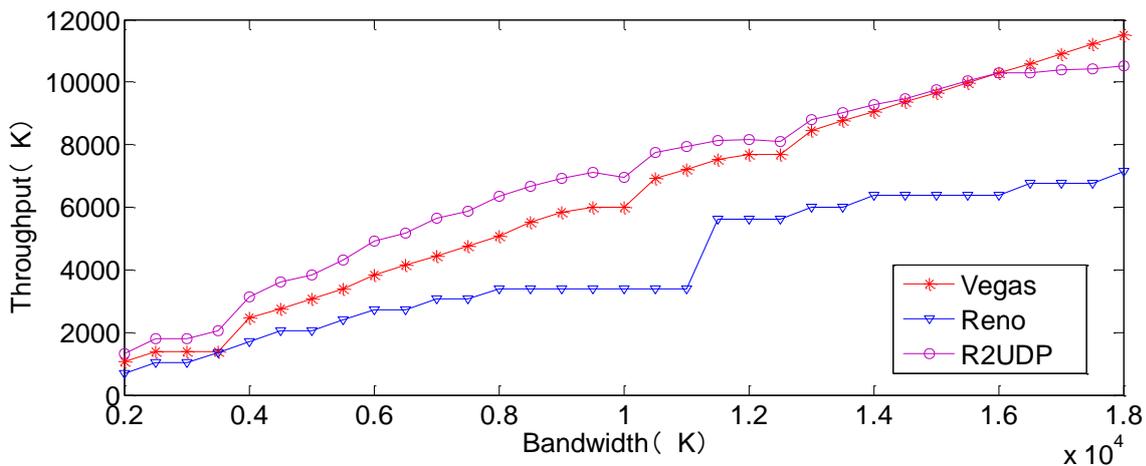


FIGURE 11 Throughput at 4% error rate

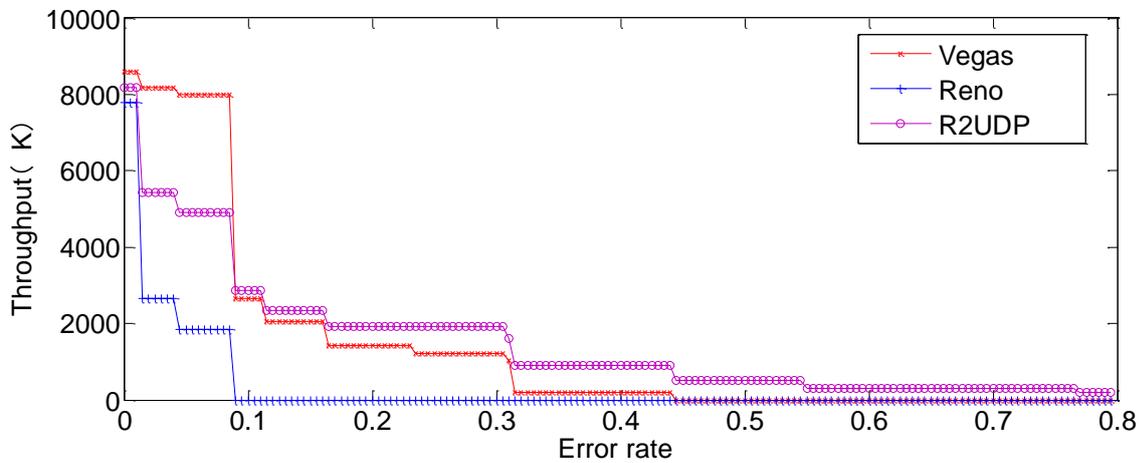


FIGURE 12 Throughput at 9.6K bandwidth

**5 Conclusions**

This paper designs a simple and reliable R<sup>2</sup>UDP protocol in order to adapt to the digital narrowband radio systems. This protocol adds Selective BACK and Smart Probe Frame. According to the measured data of Motorola Digital Radio experiments indicate that: in the case of frequent voice interrupted, R<sup>2</sup>UDP showed better transmission performance; and through the analysis of NS2 network simulation proves that, even in a higher rate of error, the narrow bandwidth cases still showed a better performance.

**Acknowledgments**

This work is supported by the National Natural Science Foundation of China (NO.61072138, No. 61379005), and Southwest University of Science and Technology (12zx7127).

**Appendix** Examples of operation

**A1 No Segment Loss and Delay**

In this case, the receiver receives all segments of the window and replies a BACK in time. It is the simplest case.

Time	Sender	Receiver
1	DATA  2   n   0   Payload   -->	
2	DATA  2   n   1   Payload   -->	
3	DATA  2   n   2   Payload   -->	
4		<--  BACK  2   n   0  00000111
5	DATA  2   n+1   0   Payload   -->(next window starts)	

**A2 Data Segments Loss**

In this case, the receiver replies a BACK indicate which segments are received, and then the sender retransmits those lost segments only.

Time	Sender	Receiver
1	DATA  2   n   0   Payload   -->	
2	DATA  2   n   1   Payload   -->	

3	DATA  2   n   2   Payload   --> Lost
4	Wait because not all segments of window 1 are received
5	Time out
6	PROB  2  1  -->
7	<--  BACK  2   n   0  00000111
8	DATA  2   n   2   Payload   -->
9	<--  BACK  2   n   0  00000111
10	DATA  2   n+1   0   Payload   -->(next window starts)

**A3 BACK Segments Loss**

In this case, the receiver replies a BACK indicate, which segments are received, and then the sender retransmits those lost segments only.

Time	Sender	Receiver
1	DATA  2   n   0   Payload   -->	
2	DATA  2   n   1   Payload   -->	
3	DATA  2   n   2   Payload   -->	
5	Lost <--  BACK  2   n   0  00000111	
6	Time out	
7	PROB  2  0  -->	
8	<--  BACK  2   n   0  00000111	
9	DATA  2   n+1   0   Payload   --> (next window starts)	

**A4 Communication Over Long Delay Path**

In this case, the sender sends a PROBE segment to query the result of last transfer when timeout.

Time	Sender	Receiver
------	--------	----------

```

+-----+-----+-----+-----+
1 |DATA| 2 | n | 0 | Payload | -->
+-----+-----+-----+-----+
+-----+-----+-----+-----+
2 |DATA| 2 | n | 1 | Payload | -->
+-----+-----+-----+-----+
+-----+-----+-----+-----+
3 |DATA| 2 | n | 2 | Payload | -->
+-----+-----+-----+-----+
+-----+-----+-----+-----+
4 [Long Delay] <-- |BACK| 2 | n | 0
|00000111|
+-----+-----+-----+-----+
5 Time out
+-----+-----+
6 |PROB| 2 |1| -->
+-----+-----+
+-----+-----+-----+-----+
7 <-- |BACK| 2 | n | 0 |00000111|
+-----+-----+-----+-----+
+-----+-----+-----+-----+
8 |DATA| 2 | n+1 | 0 | Payload | --> (next window
starts)
+-----+-----+-----+-----+
The BACK at '7' will be ignored by sender because it
duplicates with '4'.

```

**A5 Communication over Long Delay Path with Lost DATA Segments**

In this case, the sender sends a PROBE segment to query the result of last transfer when timeout. After the BACK is received, it retransmits those lost segments according to the BACK.

Time	Sender	Receiver
1	DATA  2   n   0   Payload   -->	
2	DATA  2   n   1   Payload   -->	
3	DATA  2   n   2   Payload   -->	Lost
4	Wait because not all segments are received	
5	Time out	
6	PROB  2  0  -->	
7	<--  BACK  2   n   0  00000111	
8	DATA  2   n   2   Payload   -->	
9	[Long Delay] <--  BACK  2   n   0  00000111	

```

+-----+-----+-----+-----+
10 Time out
+-----+-----+
11 |PROB| 2 |0| -->
+-----+-----+
+-----+-----+-----+-----+
12 <-- |BACK| 2 | n | 0 |00000111|
+-----+-----+-----+-----+
+-----+-----+-----+-----+
13 |DATA| 2 | n+1 | 0 | Payload | --> (next window
starts)
+-----+-----+-----+-----+
The BACK at '12' will be ignored by sender because it
duplicates with '9'.

```

**A6 Communication over Long Delay Path with Lost BACK Segments**

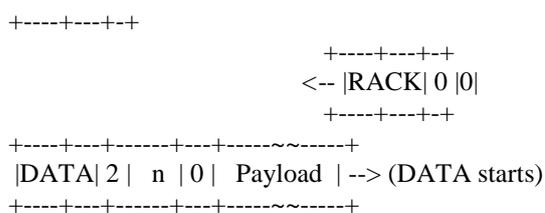
In this case, the sender sends a PROBE segment to query the result of last transfer when timeout. After the BACK is received, it retransmits those lost segments according to the BACK.

Time	Sender	Receiver
1	DATA  2   n   0   Payload   -->	
2	DATA  2   n   1   Payload   -->	
3	DATA  2   n   2   Payload   -->	
4	Lost <--  BACK  2   n   0  00000111	
5	Time out	
6	PROB  2  1  -->	
7	[Long Delay] <--  BACK  2   n   0  00000111	
8	Time out	
9	PROB  2  1  -->	
10	<--  BACK  2   n   0  00000111	
11	DATA  2   n+1   0   Payload   --> (next window starts)	

The BACK at '10' will be ignored by sender because it duplicates with '7'.

**A7 Sender Reset**

In this case, when the sender is reset, it sends a RST to reset the data path. Then the next window number will reset to 1.



**References**

[1] Partridge C, Pink S 1993 *IEEE/ACM Transactions on Networking* 1(4) 429 - 40

[2] Zhang Yong-qiang, Gao Hong-bin 2010 Design and implementation of RUDP protocol for multiple mobile agent communication *International Conference on Computer Application and System Modeling (ICCSAM)*, 22-24 Oct. (2010) 8 614-8

[3] Wang Jigang, Gu Guochang, Xie Shibo, Xu Lifeng 2006 Reliable and Efficient Data Transfer Protocol Based on UDP in Cluster System *Proceedings of the First International Multi-Symposiums on Computer and Computational Sciences* 1 518-24

[4] Doan Thanh Tran, Eunmi Choi 2007 A reliable UDP for ubiquitous communication environments *Proceedings of the 2007 Annual Conference on International Conference on Computer Engineering and Applications* 1-6

[5] Saini T K, Kumar S, Dhaka M K 2014 Analysis of routing protocols using UDP traffic under dynamic network topology. *2014 IEEE International Advance Computing Conference (IACC)*, 21-22 Feb 2014 160-165

[6] Fang Jiaoli, Liu Ming 2011 Design and Implementation of Embedded RUDP *The 2<sup>nd</sup> International Conference on Networking and Distributed Computing (ICNDC)*, 21-24 Sep 2011 7- 9

[7] Alrabaee S 2014 Concordia Inst. Using model checking for Trivial File Transfer Protocol validation *2014 International Conference on Communications and Networking ComNet 19-22 March 2014*, 1-7

[8] Ono F, Takizawa K, Tsuji H, Lin Shan, Kagawa T 2014 Measurement of TCP and UDP performance over UAS relay networks *2014 International Conference on Unmanned Aircraft Systems (ICUAS)* 27-30 May(2014) 389-94

[9] Manabe J, Funasaka J, Ishida K 2012 An improved UDP protocol for video transmission over Internet-to-wireless networks *The 3<sup>rd</sup> International Conference Networking and Computing (ICNC)* 5-7 Dec 2012 162 – 8

[10] Lam P P -K, Liew S 2004 CUDP-Liter: an improved UDP protocol for real-time multimedia applications over wireless links *International Symposium on Wireless Communication Systems*, 20-22 Sep 2004 314-8

[11] Finamore A, Mellia M, Meo M, Rossi D 2010 *IEEE/ACM Transactions on Networking* 18(5) 1505 –15

Authors	
	<p><b>Juan Zhang, born in 1982, Mianyang, China</b></p> <p><b>Current position, grades:</b> associate professor of School of Information Engineering, Southwest University of Science and Technology, PhD</p> <p><b>University study:</b> Doctor of Engineering in Signal and Information Processing from University of Chinese Academy of Sciences (2012). Now she is.</p> <p><b>Research interests:</b> cognitive radio and intelligent learning.</p>
	<p><b>Hesong Jiang, born in 1982, Mianyang, China</b></p> <p><b>Current position, grades:</b> lecturer of School of Information Engineering, Southwest University of Science and Technology</p> <p><b>University study:</b> Master of Engineering in Signal and Information Processing from University of Chinese Academy of Sciences (2009)</p> <p><b>Research interests:</b> cognitive radio and intelligent learning.</p>
	<p><b>Hong Jiang, born in 1969, Mianyang, China</b></p> <p><b>Current position, grades:</b> The full professor of School of Information Engineering, Southwest University of Science and Technology</p> <p><b>University study:</b> Doctor of Engineering in Communications professional from University of Electronic School of Information Engineering (2004)</p> <p><b>Research interests:</b> cognitive radio and intelligent learning.</p>
	<p><b>Chenmei Chen, born in 1977, Mianyang, China</b></p> <p><b>Current position, grades:</b> associate professor of School of Information Engineering, Southwest University of Science and Technology</p> <p><b>University study:</b> She received her Master of Engineering in Signal and Information Processing from University of Chinese Academy of Sciences (2008)</p> <p><b>Research interests:</b> cognitive radio and intelligent learning</p>

# A genetic algorithm for the vehicle routing optimization problem of logistics park distribution

Wenqiang Chen\*

*School of Economics and Management, Chang'an University, the middle of South 2nd ring, Xi'an, Shannxi, China*

*Received 26 October 2013, www.tsi.lv*

---

## Abstract

The Vehicle Routing Problem of Logistics park distribution (VRPLPD) is an extension of the vehicle routing problem, which deals with simultaneous distribution of goods to customers. With the increasing importance of logistics activities, it is of great theoretical and practical significance to determine efficient and effective vehicle routes for simultaneous delivery activities. The study attempts to propose a genetic algorithm approach to tackle this problem. Numerical example is presented with parameter settings in order to demonstrate the applicability and feasibility of the proposed approach. The simulation is carried out in Simulink package of MATLAB. It is shown that Genetic Algorithms are highly effective in optimizing vehicle routing problem.

*Keywords:* Vehicle routing problem (VRP), Logistics park distribution, Genetic algorithms

---

## 1 Introduction

The Vehicle Routing Problem (VRP) can be described as the problem of designing optimal delivery routes from one or several depots to a number of geographically scattered customers within some constraints [1]. The VRP, the heart of distribution management, plays a pivotal role in the fields of physical distribution and logistics [2], which serves as an effective way to ascertain the optimal set of routes within specific constraints [3-4].

The VRP can be regarded as the travelling salesman problem (TSP). In order to solve the TSP, customers are partitioned into vehicles to minimize the required number of vehicles without violating the capacity constraint. For each vehicle, the VRP seeks to find out the lowest-cost (usually the shortest-distance) driving path, which is the same as what the TSP requires. The Vehicle Routing Problem with Time Window (VRPTW) is an extension of the VRP. In the VRPTW, customers have predefined time windows and a vehicle serves a customer within the time window [5]. The Vehicle Routing Problem of Logistics park distribution (VRPLPD) is typical of the VRPTW. In Logistics park distribution, the time window is a rigid constraint. If a Logistics park distribution vehicle arrives at the downstream distributors' location earlier, it must wait until the beginning of the time window; if a Logistics park distribution vehicle arrives at the downstream distributors' location later than the end of the time window, the solution is not acceptable [5]. Due to the complexity of the challenging problem, the VRPLPD is of great theoretical and practical significance in the fields of physical distribution and logistics research [5-6].

The VRPLPD is multi-objective, with minimization of the total travel distance as the most common one. The

common way to achieve the objective is to minimize the total distance with a vehicle. This paper proposes a Genetic Algorithm (GA) to accomplish the objective. The remainder of this paper is organized as follows. Section 2 conducts an extensive and in-depth literature review. Section 3 defines the notation and problem formulation. The proposed approach is elaborated in Section 4, and numerical example in Section 5. Section 6 draws conclusion and provides future research directions. By looking into the trade-off between these solutions, scholars and practitioners in the sphere of vehicle routing can acquire more information and make more informed decisions.

## 2 Literature review

The literatures on vehicle routing problems are extensive and have been contributed by numerous scholars [7]. Hadjar et al (2009) adopted a pricing approach to solve the problem of multiple depot vehicle scheduling with time windows. They developed a dynamic time window reduction technique, which was used at every node of the price tree to tighten the time windows [8]. Chen Hsueh et al (2009) solved vehicle routing with time windows for perishable food products by utilizing a heuristic method [9]. Zachariadis et al (2009a) proposed a heuristic models methodology for the CVRP with two-dimensional loading constraints, the paper attempted to find the minimum cost routes that a vehicle started and terminated at a central depot [10]. Fleszar et al (2009) proposed a variable neighbourhood search heuristic for open vehicle problem. Proposed solution was based on reversing segments of sub-routes and exchanging segments between routes [11]. Li et al (2009) proposed a

---

\*Corresponding author e-mail: cwqiang007@126.com

Lagrangian relaxation based-heuristic for the real-time vehicle rerouting problems with time windows. In real-time vehicle, rerouting problems there are service disruption because of vehicle breakdowns. Therefore, some vehicles must be rerouted [12]. Based on ant colony optimization, Fuellerer et al (2009) developed an effective heuristic algorithm for the two-dimensional loading vehicle routing problem. There is a combination of two problems: loading of the freight into the vehicles and routing the vehicles successfully [13].

Genetic algorithms (GAs) were put forward by John Holland in the 1960s and were further developed by Holland, his students and colleagues at the University of Michigan in the 1960s and the 1970s [14-15]. Genetic Algorithms have been extensively studied, experimented and applied in many fields in engineering world. GAs not only provide an alternative method to solve problem but also consistently outperform other traditional methods in most of the problems link [16]. Genetic algorithms have shown great advantages in solving the combinatorial optimization problem, including certain types of vehicle routing problem [17]. Numerous researchers have conducted studies to solve VRP using GAs. Baker et al (2003) considered the application of GAs to solve the VRP, in which customers of unknown demands were supplied from a single depot [18]. Wang et al (2009) primarily focused on solving a capacitated vehicle routing problem by applying a novel hybrid genetic algorithm [19]. Chan et al (2004) considered the problem of scheduling a single production plant in order to satisfy delivery time constraints. They proposed two approaches, an exact method suitable only for very simple cases, and a GA for instances of more realistic size. The paper did not address a realistic application scenario, as it considered only a single depot, ignored limited resources for transportation [20]. In another recent research work, Feng et al. (2004) focused on scheduling for a single depot, equipped with a fleet of vehicles with identical capacity and a fixed (customer and depot independent) loading/unloading times [21].

**3 Notation and problem formulation**

The vehicle routing problem of logistics park distribution (VRPLPD) can be viewed as a vehicle routing optimization problem of single distribution centre within time windows. This problem can be described as: the goods are delivered with more than one vehicle to multiple customers from a logistics park, each vehicle departures from park and returns to the park after the completion of delivery. Each customer's location and demand is changeless, the time to send the goods is set within a certain window, and each vehicle load is fixed.

A feasible solution to the VRPLPD must satisfy the following constraints:

- 1) Each customer must be served by one vehicle exactly one time along the designated route;
- 2) The route of each vehicle must start from and end at

the depot;

- 3) The total demand of the customers served by each vehicle shouldn't exceed the maximum capacity;
- 4) A vehicle must arrive at customer no later than the end of the time window;
- 5) The service shouldn't start before the beginning of the time window;
- 6) The length of each route is no more than maximum mileage of a vehicle.

The parameters of Vehicle Routing Problem of Logistics Park Distribution (VRPLPD) are defined as follows:

- 1) The VRPLPD involves two types of objects: locations and vehicles. A special location 0 represents the depot. The remaining N locations correspond to N customers. The centre of Logistics park distribution has k vehicles. The maximum capacity of vehicle k is  $Q_k$ , the maximum mileage is  $D_k$ , The fixed cost of vehicle k is  $C_k$ , the average cost is  $M_k$  ( $k=1 \sim K$ ), the dispatching cost for all vehicles is Z;
- 2) Each customer i has a fixed demand  $A_i$  and a time window  $[E_i, L_i]$ ;
- 3) The distance from customers i to j is  $d_{ij}$ , the distance from logistics centre to each customer is  $d_{0i}$ ;
- 4) Arrival time of vehicle at customer i is  $t_i$ , the time of vehicle arrival at customer i from the customer j is  $t_{ij}$ , unloading per ton cargo needs time  $t_i$ . The serve time is S ;

$$y_{ik} = \begin{cases} 1, & \text{if vehicle } k \text{ delivers goods} \\ & \text{to Customer } i \\ 0, & \text{else} \end{cases}$$

$$X_{ijk} = \begin{cases} 1, & \text{if vehicle } k \text{ arrives loaction} \\ & \text{of customer } j \text{ from customer } i \\ 0, & \text{else} \end{cases}$$

The model, which is shown as below:

$$\min Z = M_k \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^K d_{ij} X_{ijk} + \sum_{k=1}^K c_k, \tag{1}$$

$$\min S = \sum_{i=1}^N \sum_{j=1}^N \{ [t_j - (t_i + t_{ij})] X_{ijk} \}.$$

Subject to

$$\begin{aligned} \sum_{i=1}^N A_i y_{ik} &\leq Q_k, \\ \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^K d_{ij} X_{ijk} &\leq D_k, \\ \sum_{k=1}^M y_{ik} &= \begin{cases} 1, & i = 1 \sim N \\ 0, & i = 0 \end{cases} \\ E_i &\leq t_i \leq L_i. \end{aligned} \tag{2}$$

**4 Genetic Algorithms**

Genetic Algorithms (GAs) are adaptive heuristic search algorithm premised on the evolutionary ideas of natural

selection and genetic. The basic concept of GAs is designed to simulate processes in natural system, which is necessary for evolution, specifically for those principles, which are firstly laid down by Charles Darwin of survival of the fittest. As such, they represent an intelligent exploitation of a random search within a defined search space to solve a problem [22].

An initial population of individuals (chromosomes) evolves through generations until reaching the satisfactory quality criteria, and a maximum number of iterations or time limits are reached. New individuals (children) are generated from individuals forming the current generation (parents) by means of genetic operators (crossover and mutation). GAs simulate the survival of the fittest among individuals over consecutive generation for solving a problem. Each generation consists of a population of character strings that are analogous to the chromosome that we see in our DNA. Each individual represents a point and a possible solution in a search space. The individuals in the population are then made to go through a process of evolution [23-24].

Algorithmically, the basic genetic algorithm is outlined as below:

**Step I [Start]** Generate random population of

chromosomes, that is, suitable solutions for the problem.

**Step II [Fitness]** Evaluate the fitness of each chromosome in the population.

**Step III [New population]** Create a new population by repeating following steps until the new population is completed:

- a) [Selection] Select two parent chromosomes from a population according to their fitness.
- b) [Crossover] With a crossover probability, cross over the parents to form new offspring, that is, children. If no crossover is performed, offspring is the exact copy of parents.
- c) [Mutation] With a mutation probability, mutate new offspring at each locus.
- d) [Accepting] Place new offspring in the new population.

**Step IV [Replace]** Use new generated population for a further run of the algorithm.

**Step V [Test]** If the end condition is satisfied, stop, and return the best solution in current population.

**Step VI [Loop]** Go to step II.

The framework of the proposed genetic algorithm is shown in Figure 1.

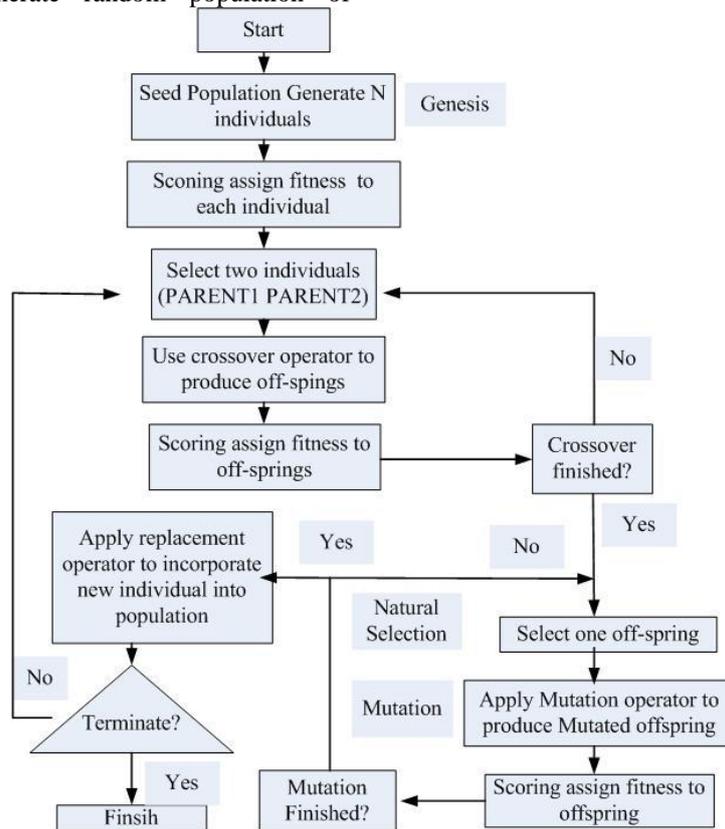


FIGURE 1 The framework of the proposed genetic algorithm

**5 Numerical example**

In this numerical example, there are 39 customer nodes, the customer nodes coordinates are shown in table 1. The depot 0 coordinates are (41, 43). A vehicle capacity is 30 units, each customer has the same demands.

Initially, a pilot study is conducted to determine the appropriate population size and number of generations for GA. The proposed approach is applied with combinations of population sizes  $M = \{50,100,150\}$ , and number of generations  $T = \{2000, 3000, 4000\}$ . The other parameters used in GA are crossover rate 0.75 and mutation rate 0.1.

TABLE 1 Customer's coordinates (unit: km)

<b>Customer ID</b>	1	2	3	4	5	6	7	8	9	10	11	12	13	14
<b>X</b>	25	8	50	41	62	18	19	38	52	42	29	9	24	51
<b>Y</b>	52	62	37	30	36	40	54	17	44	13	44	48	43	32
<b>Customer ID</b>	15	16	17	18	19	20	21	22	23	24	25	26	27	28
<b>X</b>	11	19	31	29	57	36	44	24	33	40	51	59	34	61
<b>y</b>	40	60	58	64	28	64	51	27	16	41	59	44	22	64
<b>Customer ID</b>	29	30	31	32	33	34	35	36	37	38	39			
<b>X</b>	61	31	45	54	18	59	24	15	13	7	60			
<b>y</b>	18	25	14	24	22	52	13	15	27	27	58			

Using Matlab to solve the problem, results of several iterations are summarized in Table 2. For each combination, the best objective function value (column I), the average of the best values obtained through several

runs (column II), the worst of the best values obtained through several runs (column III), the average objective function value of the population (column IV) are given.

TABLE 2 Computational results for the parameter settings

<b>Popsize</b>	<b>Number of generations</b>	<b>I</b>	<b>II</b>	<b>III</b>	<b>IV</b>
50	2000	405	435	489	503
100	2000	348	426	456	489
150	2000	326	450	437	514
50	3000	356	435	478	524
100	3000	378	406	436	536
150	3000	406	408	490	513
50	4000	379	456	431	543
100	4000	396	425	434	510
150	4000	356	411	453	514

In this example, the best objective function values and the corresponding distribution path through the evolution process are given in Figs.2 and Figs.3 respectively.

From these results, it can be concluded that population size 150 combined with number of generations 2096 results in good solutions considering the best objective function values. Total travelled distance by the vehicle is 326. The vehicle served 39 customer nodes through:

- 0 → 4 → 8 → 14 → 19 → 32 → 29 → 31 → 13
- 8 → 33 → 27 → 31 → 35 → 40 → 33 → 36
- 38 → 15 → 6 → 22 → 11 → 14 → 1 → 7 →
- 12 → 2 → 16 → 18 → 17 → 20 → 25 → 28 →
- 39 → 34 → 26 → 5 → 3 → 9 → 21 → 24 → 0

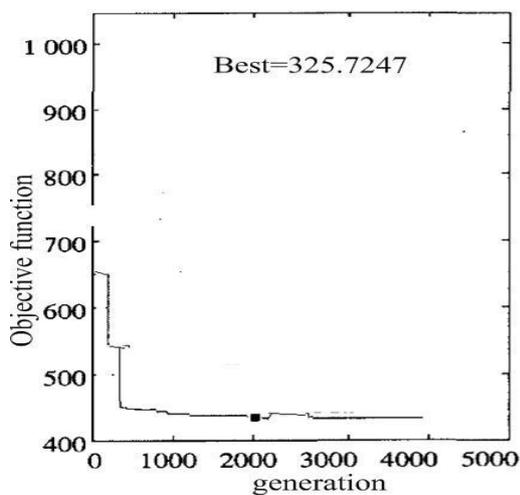


FIGURE 2 Best objective function value through the evolution process

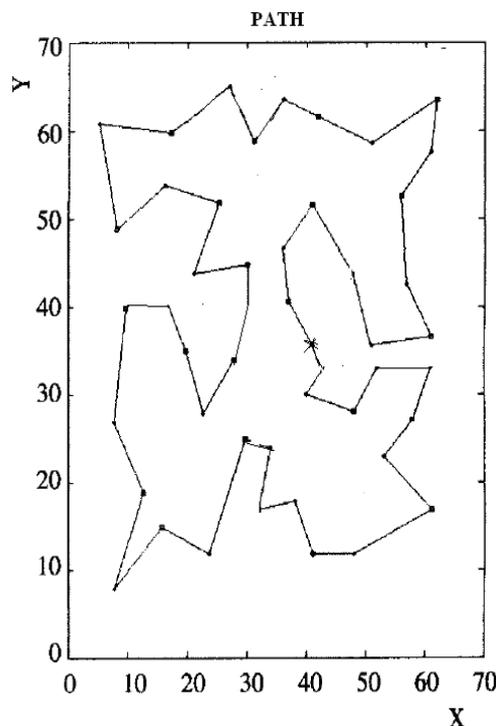


FIGURE 3 The optimal distribution path

**6 Conclusion**

The VRP is a well known combinatorial optimization problem. The Vehicle Routing Problem of Logistics Park Distribution (VRPLPD) is an extension of the VRP and considers simultaneous distribution of goods to

customers. The VRPLPD has been getting more and more attention due to the increasing importance of logistics activities. In the same vein, VRPLPD has a combinatorial nature. Due to GAs' effectiveness in solving complex optimization problems, an improved GA approach for solving VRPLPD is proposed in this study.

This study sheds light on the VRP field by providing an efficient and effective GA approach that produces highly feasible routes for VRPLPD. Details of the proposed approach are presented in the previous sections after introducing VRPLPD and its mathematical formulation. In the following section, a numerical example is provided.

According to the results of numerical example, the

proposed GA approach is proved to be effective and superior to other existent methods. Also, genetic algorithms can be applicable to other combinatorial optimization problems, but for different optimization problem, need to adopt different ways of encoding and different operation of genetic operators. Therefore, blending with other algorithms is highly advisable with the aim of gaining the most favourable solutions.

### Acknowledgements

This research was funded by the Fundamental Research Funds for the Central Universities (No. 2013G6234069).

### References

- [1] Pisinger D, Ropke S 2007 *Computers and Operations Research* **34**(8) 2403–35
- [2] Laporte G 1992 *European Journal of Operational Research* **59**(3) 345–58
- [3] Bräysy O, Gendreau M 2005 *Transport Sci* **39**(1) 119–39
- [4] Rabinovich E, Windle R, Dresner M, Corsi T 1999 *International Journal of Physical Distribution and Logistics Management* **29**(6) 353–74
- [5] Tsung Che Chiang, Wei Huai Hsu 2014 *Computers & Operations Research* **45**(5) 4525–37
- [6] Tazan A S, Gen Mitsuo 2012 **62**(3) 755–61
- [7] Ligil H B, Kara S S, Alcan P, Ozkan B, Caglar E G 2011 *Expert Systems with Applications* **38**(10), 12730–38
- [8] Hadjar A, Soumis F 2009 *Computers & Operations Research* **36**(7) 2160–72
- [9] Chen H, Hsueh C, Chang M 2009 *Computers & Operations Research* **36**(7) 2311–19
- [10] Zachariadis E E, Tarantilis C D, Kiranoudis C T 2009 *European Journal of Operational Research* **195**(3) 729–43
- [11] Fleszar K, Osman I H, Hindi K S 2009 *European Journal of Operational Research* **195**(3) 803–09
- [12] Li J, Mirchandani P B, Borenstein D 2009 *European Journal of Operational Research* **194**(3) 711–27
- [13] Fuellerer G, Doerner K F, Hartl R F, Lori M 2009 *Computers & Operations Research* **36**(3), 655–73
- [14] Holland J H 1992 *Adaptation in Natural and Artificial Systems* Cambridge MIT Press
- [15] Holland J H 1986 *Escaping brittleness: The possibilities of general-purpose learning algorithms applied to parallel rule-based systems* Boca Raton Florida: CRC Press
- [16] Pisinger D, Ropke S 2007 *Computers and Operations Research* **34**(8) 2403–35
- [17] Neri F, Cotta C 2012 *Swarm and Evolutionary Computation* **21**–14
- [18] Wang C H, Lu J Z 2009 *Expert Systems with Applications* **36**(2) 2921–36
- [19] Baker B M, Ayeche M A 2003 *Computers & Operations Research* **30**(5) 787–800
- [20] Chan F T S, Chung S H, Wadhwa S 2004 *International Journal of Production Research* **42** (1) 1–19
- [21] Feng C W, Cheng T M, Wu H T 2004 *Automation in Construction* **13** (3) 327–40
- [22] Schmitt L M, Nehaniv C L, Fujii R H 1998 *Theoretical Computer Science* **200**(1–2) 101–34
- [23] *Genetic Algorithms in Engineering and Computer Science* 1995 New Jersey: John Wiley & Sons Ltd
- [24] Mitchell M 1998 *An Introduction to Genetic Algorithms* Cambridge, Massachusetts London, England: MIT Press

### Author



**Chen WenQiang**, born on July 7, 1981, Xian, Shannxi, P.R. China

**Current position, grades:** The Lecturer of School of Economics and Management, Chang' an University, China.

**University studies:** B.Sc. in Transportation Planning and Management from Chang' an University of Economics and Management, in China, M.Sc. from Chang' an University in China.

**Scientific interest:** Transportation planning, logistics management

**Publications:** more than 10 papers published in various journals.

**Experience:** teaching experience of 3 years, has completed one scientific research projects

# A novel method for identifying system modal parameters using stabilization diagram

Wen-sheng Xiao<sup>1</sup>, Zhong-yan Liu<sup>1\*</sup>, Jian Liu<sup>1</sup>, Han-chuan Wu<sup>2</sup>

<sup>1</sup>College of Machinery and Electronic Engineering, China University of Petroleum, Qingdao, Shandong, China, 266555

<sup>2</sup>SJ Petroleum Machinery Co., Jingzhou, Hubei, China, 434024

Received 1 March 2014, www.tsi.lv

## Abstract

Modal parameters of a structure are important for system identification. In order to identify modal parameters of a structure more accurately, this paper proposes a parameter identification method combined with stabilization diagram. Stochastic subspace identification (SSI) is a recently developed method for identifying a linear system. Combining SSI and the proposed method can easily confirm system order. However, the proposed method has difficulty in distinguishing spurious modals. Therefore, the proposed method must be revised to ensure that the spurious modal can be detected and the SSI can be used to improve identification accuracy. Finally, a simulation is conducted on a fracturing pump truck, when the damping ratio increases from 10% to 40%, those spurious modals disappear. The results indicate that this method performs precise identification.

*Keywords:* Parameter Identification, Modal Analysis, Stochastic Subspace Identification, Stabilization Diagram, Fracturing Pump Truck

## 1 Introduction

Determining modal parameters has become important for system identification in the past decades. Results of experimental modal analysis (EMA) are used in practice; an overview of the EMA method can be found in [1–3]. In some cases, performing the vibration test on large structures is difficult or impossible because some excitations, such as wind or traffic, cannot be measured. In addition, using artificial excitation, such as hammer or drop weight, is impractical or, in some cases, expensive.

Therefore, output-only stochastic system identification methods have been developed. In these methods, ambient forces cannot be ignored and should be regarded as stochastic quantities with some unknown parameters. Stochastic subspace identification (SSI) is one of the methods for identifying system parameters. SSI has two implementation procedures: covariance-drive (SSI-cov) and data-drive (SSI-data) implementation [4]. Given that these methods need only the outputs of the structure for measurement, artificial excitation is unnecessary.

Estimating the modal parameters of the structures according to the measured data involves three steps: data collection, system identification, and determination of modal parameters [5–7]. This paper focuses on data collection. Thus, system identification should be treated as an important problem and is defined as construction of the system model according to the measured data. The SSI method is used in the time domain because of its convenience [8]. The modal parameters can be

determined according to a free vibration analysis of the identified system model.

In this paper, the stabilization diagram can be used to determine the system order, which is an important step for system identification. Other studies [9–11] confirmed the stabilization diagram method based on singular value decomposition. However, the obtained results are insufficient because the stabilization diagram method is a comparatively new method of determining system order. The stabilization diagram method can be used to distinguish real modals and modals in cases with excess noise. The stabilization diagram can delete certain system poles that meet the condition, but cannot be treated as real poles because they may belong to noise modals and not to the system. Thus, these poles can be distinguished and deleted by using the stabilization diagram.

This paper is organized as follows: Section 2 shows how the vibration structure can be modeled according to stochastic state-space models and modal analysis. Section 3 discusses the subspace identification method used for system identification. The use of a stabilization diagram to determine the system order is discussed in Section 4. Section 5 shows the application of this method to a real structure.

## 2 Stochastic state-space model for vibrating structures

For a linear dynamical system model, the following system of ordinary differential equations can be obtained:

$$M \frac{d^2 \mathbf{u}(t)}{dt^2} + C_2 \frac{d\mathbf{u}(t)}{dt} + K\mathbf{u}(t) = \mathbf{B}_2 \mathbf{f}(t), \quad (1)$$

\*Corresponding author e-mail: liuzhy@upc.edu.cn

where  $M$  represents the mass matrices,  $C_2$  represents the stiffness matrices,  $K$  represents the damping matrices,  $f(t)$  and  $u(t)$  represent the nodal forces and nodal displacement, respectively,  $B_2$  is the selection matrix, and  $t$  is the time. This equation can be converted into the state-space model as follows:

$$\frac{dx(t)}{dt} = A_c x(t) + B_c f(t), \tag{2}$$

where  $x(t) = \begin{bmatrix} u(t) \\ \frac{du(t)}{dt} \end{bmatrix}$ ,  $A_c = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C_2 \end{bmatrix}$ ,

and

$$B_c = \begin{bmatrix} 0 \\ M^{-1} \end{bmatrix} B_2, \tag{3}$$

where  $x(t)$  is the state of the structure. The quantities of interest can be grouped in an output vector ( $t$ ) as follows:

$$\begin{aligned} y(t) &= C_a \frac{du(t)}{dt} + C_v \frac{du(t)}{dt} + C_d u(t) \\ &= [C_d - C_a M^{-1} K C_v - C_a M^{-1} C_2] x(t) + C M^{-1} B_2 f(t), \tag{4} \\ &= Cx(t) + Df(t) \end{aligned}$$

when they are the linear combination of nodal displacements, velocities, or accelerations.

In these equations,  $C_a$ ,  $C_v$ , and  $C_d$  are the selection matrices. The discrete-time state-space model can be obtained after discretization in time:

$$x_{k+1} = Ax_k + Bf_k, \tag{5}$$

$$y_k = Cx_k + Df_k. \tag{6}$$

From the relationship above, the system matrices on continuous-time and the discrete-time can be obtained as follows:  $A = e^{A_c(\Delta t)}$  and

$$\begin{aligned} A &= e^{A_c(\Delta t)} B = \int_{k\Delta t}^{(k+1)\Delta t} e^{A_c((k+1)\Delta t - \tau)} d\tau B_c \\ &= (A - I)A_c^{-1} B, \tag{7} \end{aligned}$$

where  $\Delta t$  is the discrete-time step.

When system matrices  $A$ ,  $B$ ,  $C$ , and  $D$  are known, the outputs  $y_k$  are measured. However, the inputs cannot be known; thus,  $f_k$  remains unknown. In the state-space equation, the measurement noise on the measured outputs should not be neglected.

The state-space equation can be written as follows:

$$x_{k+1} = Ax_k + w_k, \tag{8}$$

$$y_k = Cx_k + v_k, \tag{9}$$

$$w_k = Bf_k \text{ and } v_k = Df_k + n_{y,k}, \tag{10}$$

where  $N_{y,k}$  can be considered the measurement noise. The stochastic terms  $w_k$  and  $v_k$  are unknown in the above equation. However, these variables are assumed to have a white noise nature and an expected value of zero. The covariance matrices can then be defined as follows:

$$E \left( \begin{bmatrix} w_p \\ v_p \end{bmatrix} \begin{bmatrix} w_p^T & v_p^T \end{bmatrix} \right) = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \cdot \delta(p-q). \tag{11}$$

The states and the output can be separated into a purely stochastic part as follows:

$$\begin{aligned} x_k &= x_k^d + x_k^s, \quad x_{k+1}^d = Ax_k^d + Bf_k, \quad x_{k+1}^s = Ax_k^s + w_k, \\ y_k &= y_k^d + y_k^s, \quad y_k^d = Cx_k^d + Df_k, \quad y_k^s = Cx_k^s + v_k. \end{aligned} \tag{12}$$

The state cannot be calculated exactly because of the stochastic terms. However,  $x_{k+1}^d$  can be estimated when the output vector  $y_k$  can be measured. The Kalman filter offers a method of determining the optimal linear estimate because of the unbiased and minimum variance of the estimator.

### 3 Reference-based deterministic-stochastic subspace identification

#### 3.1 IDENTIFICATION OF SYSTEM MATRICES

In some vibration tests, the sensors are less adequate than the test spots in the structures. Hence, several steps may be needed to complete the tests. Several test spots are selected as reference spots to unify every test step. Sensors in the reference spots are stabilized and sustained. In the state-space equation, the system matrices  $A$ ,  $B$ ,  $C$ ,  $D$ ,  $Q$ ,  $R$  and  $S$  are all unknown. The outputs can be grouped into the following block Hankel matrix:

$$Y_{0|2i-1} = \frac{1}{\sqrt{j}} \begin{bmatrix} y_0^{ref} & y_1^{ref} & y_2^{ref} & \dots & y_{j-1}^{ref} \\ y_1^{ref} & y_2^{ref} & y_3^{ref} & \dots & y_j^{ref} \\ \dots & \dots & \dots & \dots & \dots \\ y_{i-1}^{ref} & y_i^{ref} & y_{i+1}^{ref} & \dots & y_{i+j-2}^{ref} \\ y_i & y_{i+1} & y_{i+2} & \dots & y_{i+j-1} \\ y_{i+1} & y_{i+2} & y_{i+3} & \dots & y_{i+j} \\ \dots & \dots & \dots & \dots & \dots \\ y_{2i-1} & y_{2i} & y_{2i+1} & \dots & y_{2i+j-2} \end{bmatrix} = \begin{bmatrix} Y_p^{ref} \\ Y_f \end{bmatrix}. \tag{13}$$

The inputs can also be grouped into the following block Hankel matrix:

$$F_{0|2i-1} = \frac{1}{\sqrt{j}} \begin{bmatrix} f_0 & f_1 & f_2 & \dots & f_{j-1} \\ f_1 & f_2 & f_3 & \dots & f_j \\ \dots & \dots & \dots & \dots & \dots \\ f_{i-1} & f_i & f_{i+1} & \dots & f_{i+j-2} \\ f_i & f_{i+1} & f_{i+2} & \dots & f_{i+j-1} \\ f_{i+1} & f_{i+2} & f_{i+3} & \dots & f_{i+j} \\ \dots & \dots & \dots & \dots & \dots \\ f_{2i-1} & f_{2i} & f_{2i+1} & \dots & f_{2i+j-2} \end{bmatrix} = \begin{bmatrix} F_P \\ F_f \end{bmatrix}. \quad (14)$$

According to the subspace identification theorem, the following equation can be obtained:

$$a.s. \lim_{j \rightarrow \infty} \mathcal{G}_i = \lim_{j \rightarrow \infty} \Gamma_i \hat{X}_i, \quad (15)$$

where  $\mathcal{G}_i$  is the oblique projection of the row space of  $Y_f$  onto the joint row space of  $F_P$  and  $Y_P^{ref}$  in the direction of the row space of  $F_f$ ,

$$\mathcal{G}_i = Y_f / F_f \begin{bmatrix} F_P \\ Y_P^{ref} \end{bmatrix}, \quad (16)$$

where  $\Gamma_i$  is the extended observability matrix:

$$\Gamma_i = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{i-1} \end{bmatrix}, \quad (17)$$

where  $\hat{X}_i$  is the sequence of reference-based Kalman filter states:  $\hat{X}_i = [\hat{x}_i \ \hat{x}_{i+1} \ \dots \ \hat{x}_{i+j-1}]$ .

The theorem states that the rank of  $\mathcal{G}_i$  is equal to the system order  $n$ . The matrix  $\Gamma_i$  can be calculated according to the following singular value decomposition:

$$W_1 \mathcal{G}_i W_2 = [U_1 \ U_2] \begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U_1 S_1 V_1^T, \quad (18)$$

where  $rank(P_i^{ref}) = n, U_1 \in R^{li \times n}, S_1 \in R^{n \times n}, V_1 \in R^{j \times n}$ .

The state sequence of the Kalman filter can be obtained as follows:

$$\begin{aligned} O_i &= U_1 S_1^{1/2}, O_{i-1} = O_i(1:l(i-1),:) \\ \hat{X}_i &= O_i^+ P_i^{ref}, \hat{X}_{i+1} = O_{i-1}^+ P_{i-1}^{ref} \end{aligned} \quad (19)$$

The stochastic state-space model equations can be calculated as follows:

$$\begin{pmatrix} \hat{X}_{i+1} \\ Y_{ii} \end{pmatrix} = \begin{pmatrix} A \\ C \end{pmatrix} (\hat{X}_i) + \begin{pmatrix} w_i \\ v_i \end{pmatrix}. \quad (20)$$

The output sequence is represented as follows:

$$Y_{ii} = \begin{pmatrix} R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & R_{33} \end{pmatrix}. \quad (21)$$

The system and output matrices of the structures have the least squares solution:

$$\begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} \hat{X}_{i+1} \\ Y_{ii} \end{pmatrix} \hat{X}_i^+. \quad (22)$$

The noise sequence is given by

$$\begin{pmatrix} w_i \\ v_i \end{pmatrix} = \begin{pmatrix} \hat{X}_{i+1} \\ Y_{ii} \end{pmatrix} - \begin{pmatrix} A \\ C \end{pmatrix} (\hat{X}_i). \quad (23)$$

### 3.2 DETERMINATION OF MODAL PARAMETERS

The modal parameters of the system (eigenfrequencies, damping ratios, and mode shapes) can be obtained from the identified system description  $(A, C)$ . An eigenvalue decomposition of  $A$  obtains the diagonal matrix  $\Lambda$  of discrete-time system poles  $\lambda_i$  and corresponding right eigenvectors  $\Psi_i$ :

$$A = \Psi \Lambda \Psi^{-1}, A \Psi_i = \lambda_i \Psi_i. \quad (24)$$

The continuous-time system poles  $\lambda_{ci}$  can be calculated by the discrete-time system poles  $\lambda_i$  as follows:

$$\lambda_{ci} = \frac{\ln(\lambda_i)}{\Delta t}. \quad (25)$$

The undamped eigenequencies  $f_i$  and damping ratios  $\xi_i$  can be calculated from the continuous-time system poles  $\lambda_{ci}$  by

$$f_i = \frac{|\lambda_{ci}|}{2\pi} \text{ and } \xi_i = -100 \frac{\lambda_{ci}^R}{|\lambda_{ci}|}. \quad (26)$$

The experimental mode shape  $\Phi_i$  can be calculated as follows:

$$\Phi = C \Psi, \Phi_i = C \Psi_i. \quad (27)$$

### 4 Stabilization diagram

As the true system order is often unknown, a practical method is to calculate the model orders  $n$ . The true

system poles can be detected by comparing the modal parameters for different model orders. Thus, weakly excited system poles can be detected. This procedure is called stabilization diagram and is one of the novel methods for distinguishing system modals. The basic concept of the stabilization diagram is shown in Figure 1.

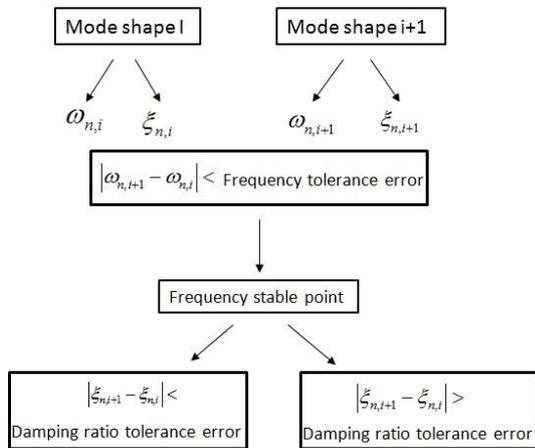


FIGURE 1 Theory of the stabilization diagram

Certain matrices, such as frequency, damping, and mode shape matrices, should be established to obtain accurate results. According to the frequency and damping matrices, every model order frequency and damping can be confirmed because they are both the average of each matrix.

Every point should then be judged according to whether they are stable or not based on the following:

- (1) The deviation between frequency and average frequency.
- (2) The deviation between damping and average damping.

In practice, when the assumed deviation of the damping ratio is under 10%, numerous mode shapes have similar frequencies. Thus, many false mode shapes are eliminated when the deviation of the damping ratio increases.

5 Sample analyses

One of the applications of the SSI modal analysis method is the fracturing pump truck, which has become increasingly important with the development of shale gas. Shale gas has a crucial role in fracturing work. Thus, learning the vibration characteristics of fracturing pump truck and determining abnormal vibrations is important.

5.1 EXPERIMENT SETUP

An experiment on truck vibration characteristics is carried out to simulate fracturing pump truck vibration characteristics under the support boundary condition, as shown in Figure 2. This experiment consists of 14 reference channels, which can collect all acceleration data.

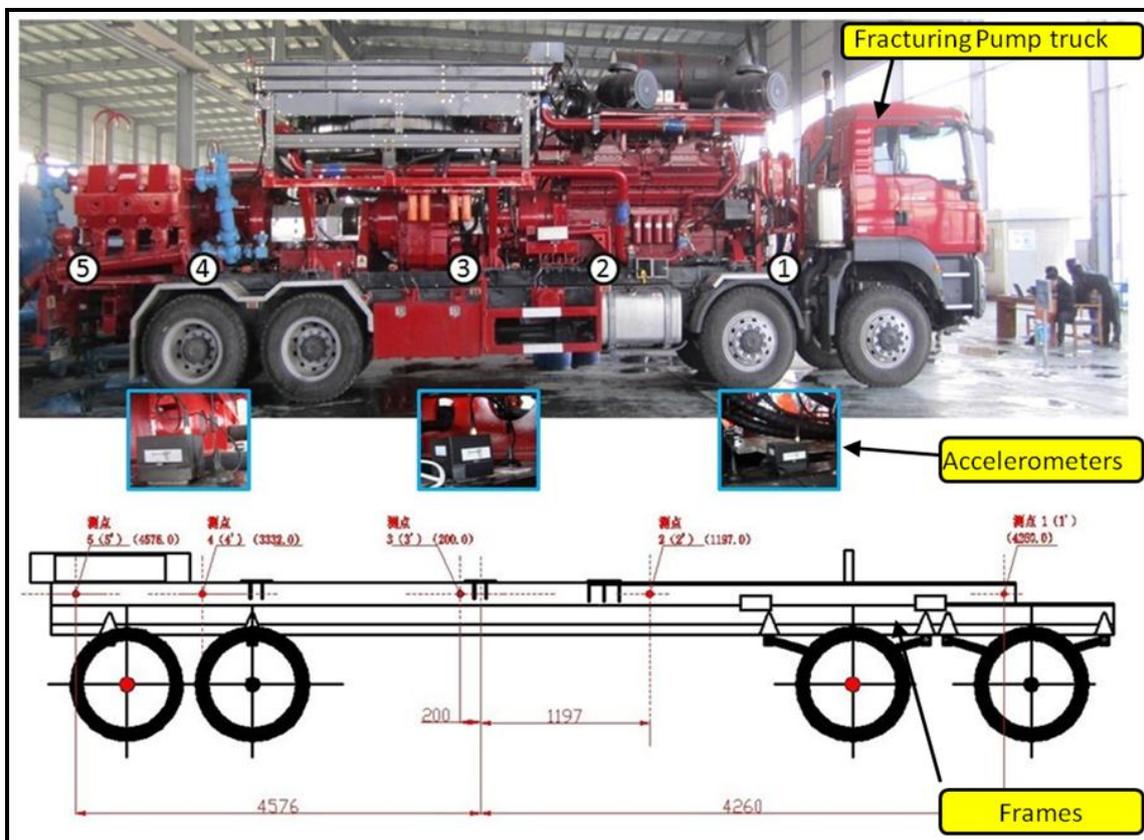


FIGURE 2 Experiment system: fracturing pump truck

5.2 SIMULATION OUTPUT

Simulations are performed to illustrate the function of the stabilization diagram. In these simulations,  $f_k$  is white noise, and  $v_k$  is a white noise vector. The only assumption of SSI is the infinite amount of measurement data. The stabilization diagram for this simulation is shown in Figure 3. Certain mathematical poles can be removed based on the following criteria: difference in

two consecutive eigenfrequencies  $df_i < 1\%$ ; difference in two consecutive damping ratios  $d\xi_i < 5\%$ ; and the highest modal transfer norms  $N_n = 3$ . The modal transfer norm  $n_i$  is the contribution of each mode to the total positive power spectral density. Given that the system and measurement noise terms are white noises, the contribution of the spurious modes is low enough that the modes are equal to an infinite number of samples.

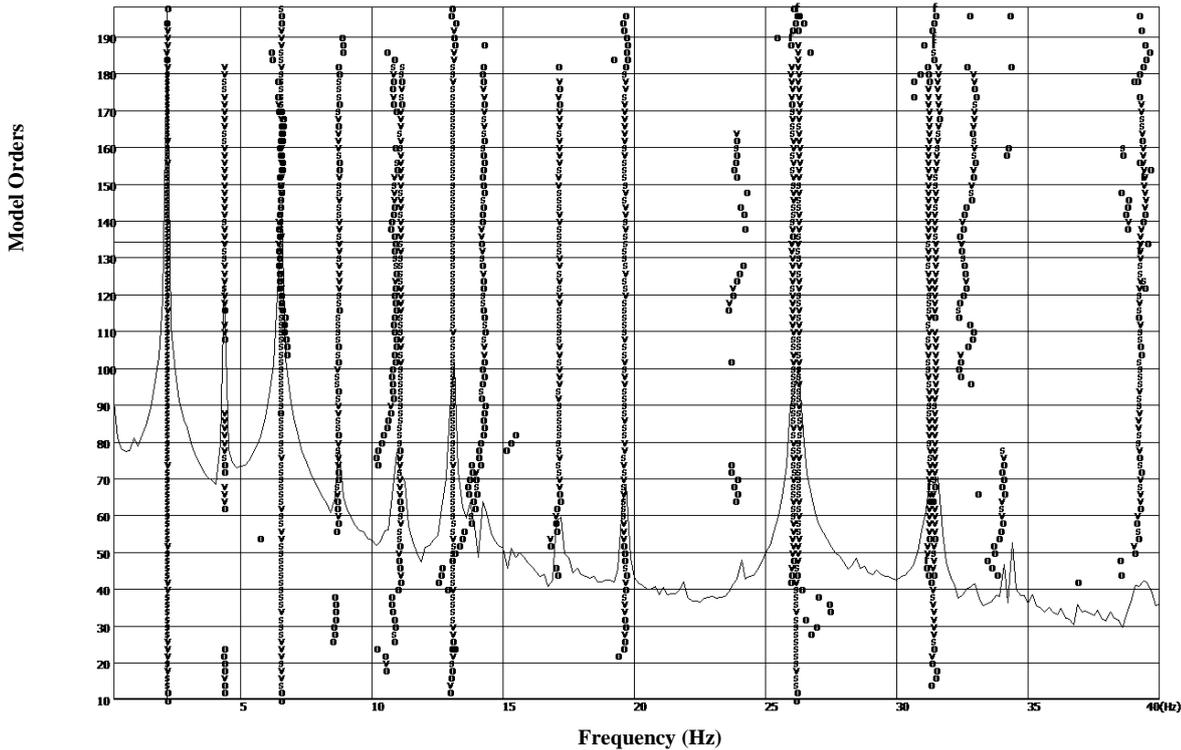


FIGURE 3 Stabilization diagram obtained by applying SSI. Stabilization criteria: 2% for frequencies, 10% for damping ratios, 2% for mode shape correlations, and  $df_i < 1\%$ ,  $d\xi_i < 5\%$ ,  $N_n = 3$

Spurious modes are removed as shown in Figure 4. The simulation shows that the stabilization criteria are similar to those of the first simulation, except for the damping ratio deviation. In this simulation, the mode shape shows the operational deflection shapes. The spurious modes that pass the stabilization criteria can be easily detected based on the nature of their mode shapes. The SSI method can distinguish the system mode shape accurately, as shown in Table 1. When the damping ratio is 10%, spurious modals occur and the frequency of these modals and mode shapes are similar, except for the obvious difference in the damping ratio. When the damping ratio is 40%, these spurious modals can be eliminated.

TABLE 1 Dynamic parameters of the fracturing pump truck frame

Number	Frequency/Hz		Measurement
	Damping criteria:10%	Damping criteria:40%	Damping ratio/%
1	2.186	2.186	0.24
2	4.369		0.46
3	4.378	4.378	0.14
4	6.449		0.34
5	6.515		0.27
6	6.539	6.539	0.18
7	8.711	8.711	0.61
8	8.840		0.92
9	10.492		1.15
10	13.037	13.037	0.21
11	15.505		
12	15.535	15.535	0.73
13	15.552		

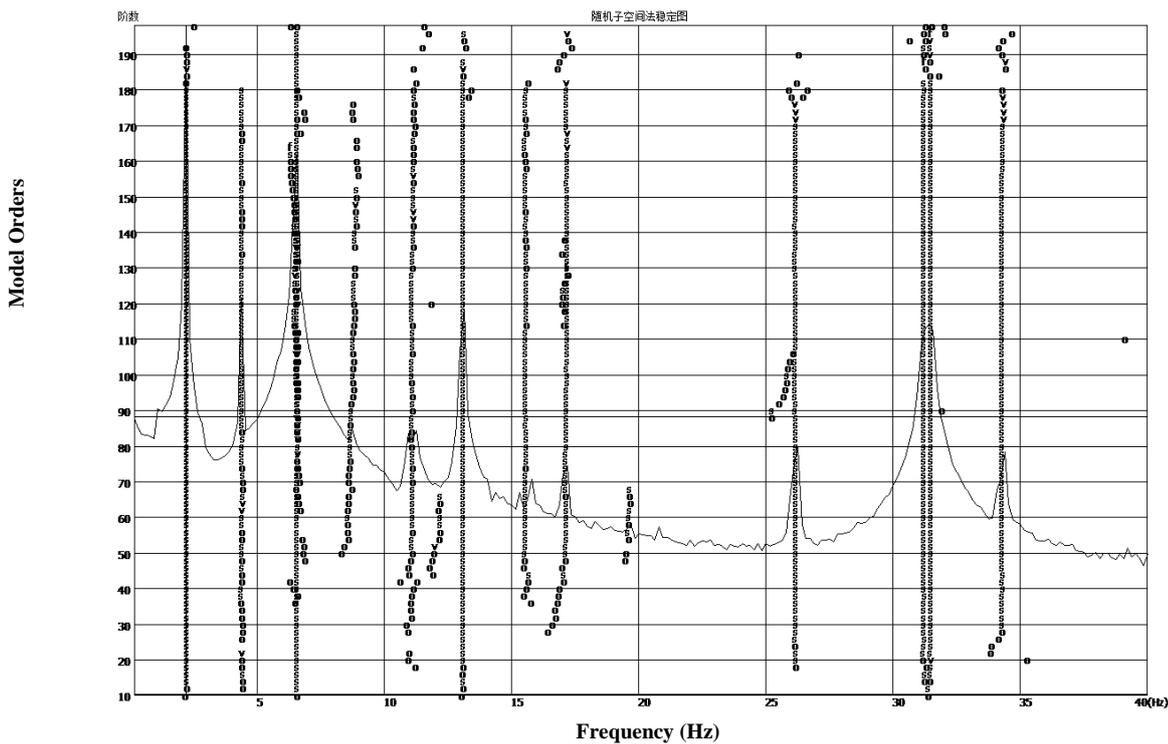


FIGURE 4 Stabilization diagram obtained by applying SSI. Stabilization criteria: 2% for frequencies, 40% for damping ratios, 2% for mode shape correlations, and  $df_i < 1\%$ ,  $d\xi_i < 5\%$ ,  $N_n = 3$

## 5 Conclusions

This paper presents a modal parameter identification method that combines SSI and stabilization diagram. The proposed method is used to evaluate a fracturing pump truck system, and it obtains ideal results. A simulation of the fracturing pump truck shows that the damping ratio can affect the accuracy of the results. In the stabilization diagram, most points can meet the demand for frequency and stability of mode shapes. Therefore, this method can effectively identify system parameters.

## References

- [1] Ewins D 2000 *Modal Testing, seconded* Baldock: Research Studies Press Chapter 8
- [2] Heylen W, Lammens S 1997 *Modal Analysis Theory and Testing* Leuven: Research Studies Press Chapter 6
- [3] Maia N, Silva J 1997 *Theoretical and Experimental Modal Analysis* Taunton: Research Studies Press
- [4] Peeters B, de Roeck G 1999 *Mechanical Systems and Signal Processing* **13** (6) 855–78
- [5] Cauberghe B 2004 *Applied frequency-domain system identification in the field of experimental and operational modal analysis* Ph.D. Thesis of Vrije Universiteit Brussel Belgium
- [6] Peeters B 2000 *System identification and damage detection in civil engineering* Ph.D. Thesis of Leuven Katholieke Universiteit Leuven Belgium
- [7] Bathe K-J 1996 *Finite Element Procedures seconded* Englewood Cliffs NJ: Prentice-Hall
- [8] Pintelon R, Schoukens J 2001 *System Identification* New York: IEEE Press

## Acknowledgment

This work was supported by the "National Science and Technology Major Project in 12th-5-year China: The Development and Application of 3000HP Combination Fracturing Unit, Project Number: (2011zx05048-01)" and "Postgraduate Innovation Engineering of China University of Petroleum".

- [9] Brincker R, Zhang L 2000 Modal identification from ambient responses using frequency domain decomposition *Proceedings of IMAC* **22**(18) 625-30
- [10] van Overschee P 1996 *Subspace Identification for Linear Systems: Theory-Implementation- Applications* Dordrecht: Kluwer Academic Publishers
- [11] Peeters B 2007 *System identification and damage detection in civil engineering* PhD Thesis of of Leuven Katholieke Universiteit Leuven Belgium
- [12] Belgium 2000 [www.bwk.kuleuven.ac.be/bwm/](http://www.bwk.kuleuven.ac.be/bwm/) 16 March 2014
- [13] Hermans L, van der Auweraer H 1999 *Mechanical Systems and Signal Processing* **13**(2) 193-216
- [14] Teughels A, de Roeck G 2004 *Journal of Sound and Vibration* **278** (3) 589 - 610
- [15] Kramer C, de Roeck G 1999 Z24 bridge damage detection tests *Proceedings of the IMAC XVII Conference* **43**(17) 1023 - 9
- [16] Dooms D, Degrande G 2006 *Engineering Structures* **28** (4) 532-42
- [17] Brewer J W 1978 *IEEE Transactions on Circuits and Systems* **25** 772-81
- [18] Juang J-N 1994 *Applied System Identification* Upper Saddle River, NJ : Prentice-Hall

Authors	
	<p><b>Wensheng Xiao, born in August, 1966, Qingdao County, Shandong Province, P.R. China</b></p> <p><b>Current position, grades:</b> Professor of School of Machinery and Electronic Engineering, China University of Petroleum, China  <b>University studies:</b> Ms.Sc. in Machinery and Electronic Engineering from China University of Petroleum in China, Dr.Sc. from HUAZHONG UNIVERSITY OF SCIENCE and TECHNOLOGY in China  <b>Scientific interests:</b> Vibration analysis of oil equipment, Ocean oil equipment research  <b>Publications:</b> more than 40 papers published in various journals  <b>Experience:</b> teaching experience of 23 years, has completed fourteen scientific research projects</p>
	<p><b>Zhongyan Liu, born in January, 1987, Qingdao County, Shandong Province, P.R. China</b></p> <p><b>Current position, grades:</b> PH.D of School of Machinery and Electronic Engineering, China University of Petroleum, China.  <b>University studies:</b> Ms.Sc. in Machinery and Electronic Engineering from China University of Petroleum in China, Dr.Sc. from China University of Petroleum in China  <b>Scientific interests:</b> Vibration analysis of oil equipment, Ocean oil equipment research  <b>Publications:</b> more than 10 papers published in various journals.  <b>Experience:</b> has completed six scientific research projects</p>
	<p><b>Jian Liu, born in June, 1971, Qingdao County, Shandong Province, P.R. China</b></p> <p><b>Current position, grades:</b> Associate Profess of School of Machinery and Electronic Engineering, China University of Petroleum, China.  <b>University studies:</b> Ms.Sc. in Machinery and Electronic Engineering from China University of Petroleum in China, Dr.Sc. from ZHENJIANG University in China  <b>Scientific interests:</b> Vibration analysis of oil equipment, Ocean oil equipment research  <b>Publications:</b> more than 25 papers published in various journals.  <b>Experience:</b> has completed eleven scientific research projects.</p>
	<p><b>Hanchuan Wu, born in March, 1959, Jingzhou County, Hubei Province, P.R. China</b></p> <p><b>Current position, grades:</b> Chief Designer of SJ Petroleum Machinery Co., China  <b>University studies:</b> Ms.Sc. in Machinery and Electronic Engineering from China University of Petroleum in China  <b>Scientific interests:</b> R&amp;D of oil equipment, Ocean oil equipment research  <b>Publications:</b> more than 8 papers published in various journals  <b>Experience:</b> He has completed twelve scientific research projects.</p>

# Decline in gas pressure influences the deformation and permeability of coal-containing methane

Yunqi Tao<sup>1, 2</sup>, Dong Liu<sup>2, 3</sup>, Jie Cao<sup>3\*</sup>, Jiang Xu<sup>3</sup>

<sup>1</sup>School of Energy Science and Engineering, Henan Polytechnic University, Jiaozuo, Henan Province, China

<sup>2</sup>Yongcheng Coal & Electricity Holding Group Co. Ltd., Yongcheng, Henan Province, China

<sup>3</sup>State Key Laboratory of Coal Mine Disaster Dynamics and Control, Chongqing University, Chongqing, China

Received 17 January 2014, www.tsi.lv

---

## Abstract

The development and utilization of coal-bed methane, as an unconventional gas, is not only beneficial to the reduction of environmental pollution caused by fossil fuels, but also conducive to the prevention of disasters during coal mining. In this paper, a dynamic permeability model of coal body is established and discussed by means of experimental tests, which measure the deformation and gas permeability of coal-containing methane in the process of gas pressure reduction under different temperatures. The results show that, when gas pressure decreases, the strain of coal-containing methane increases linearly. With temperature increases, the variation of radial strain decreases. Under the same temperature, the permeability of coal decreases first and then increases again during gas pressure reduction. The changing point of gas pressure is approximately 1.2 MPa in the study. In the initial stage of gas pressure decrease, the radial strain of coal-containing gas has a significant effect on its permeability.

*Keywords:* Gas Pressure, Gas Permeability, Coal Deformation, Dynamic Permeability Model

---

## 1 Introduction

The “fog” has recently swept through nearly half of China. This occurrence indicates that environment problems are becoming increasingly serious. The energy consumption structure must be reformed; therefore, the exploration and application of natural gas, coal-bed methane (CBM), and shale gas should be strengthened. Among these resources, CBM is an unconventional clean energy that is symbiotic with coal, and it is abundant in China. Furthermore, the reasonable development and utilization of CBM can effectively reduce the risks associated with coal mining.

At present, the United States, Canada, Australia, and China are the main countries engaged in the exploration and development of CBM and have realized CBM industry. During the extraction of CBM, gas pressure of the coal reservoir gradually decreases, which affects reservoir permeability, thus further affects the output of CBM. On the other hand, with the increase in extraction depth, the geothermy is playing an increasingly important role in the extraction of CBM [1]. Therefore, study on deformation and permeability variations in the coal reservoir during the reduction process of gas pressure under different temperatures is significant in the effective exploitation of CBM. Thus far, most of the studies on dynamic deformation and permeability variations during the extraction of CBM are conducted are numerical simulation based on mathematical model. Based on the P&M model, DENG Ze [2] simulated permeability

variations in the reduction process of reservoir pressure, with the background of No. 3 coal reservoirs of Qinshui CBM Field. The results showed that, with the decrease in gas pressure, permeability firstly decreases and then increases. ZHOU Junping [3-4] established a fluid-solid coupling model for CBM considering the matrix shrinkage effects and simulated the changes of permeability during the primary production of CBM. By measuring the macroscopic fracture, mechanical parameters, stress and permeability of the main coal seam in the south-central Qinshui Basin, FU Xuehai [5] established a numerical model to simulate the dynamic change of permeability during the production of CBM. With regard to the physical simulation experiment, the permeability and deformation of coal seam influenced by gas pressure and temperature have been the focus of many studies. ZHAO Yangsheng [6-8] concluded that coal and rock permeability changed parabolically with gas pressure through experiments by keeping the axial pressure and confining pressure constant and increasing the gas pressure and then proposed the concept of critical gas pressure. ZHAO Yangsheng also observed the trend that lignite permeability firstly decreased, then increased significantly, and finally decreased in the experimental study under the condition of different temperatures. CAO Shugang [9] derive a quadratic polynomial relationship between gas pressure and permeability by studying the influence of gas pressure on coal permeability characteristics, whereas the gas pressure range in his study was only confined from 0.3 MPa to 1.5 MPa.

---

\*Corresponding author e-mail: cdcaoj@gmail.com

LIANG Bing [10] analysed the mechanical and nonmechanical mechanisms of the influence of gas on coal and rock deformation by conducting experiments on triaxial compression under the condition of different confining pressures and gas pressures. LI Zhiqiang [11-12] concluded that the relationship between coal permeability and coal temperature was not monotonically increasing or monotonically decreasing under the condition of different effective stresses. A transition zone existed through coal and rock permeability change under the condition of different temperatures and stresses. And the location of the transition zone depended on the ratio of effective stress to thermal stress. XU Jiang [13] conducted coal and rock seepage experiments and mechanical tests under different temperatures and concluded that permeability decreased with the increase in temperature. In addition, the influence of temperature on permeability would decrease with the increase in effective stress and gas pressure. Moreover, XU Jiang concluded that the coal and rock deformation increased with the increase in temperature and that the dependent variables had different changing trends under different temperature ranges. So far, studies regarding the deformation and seepage evolution of CBM reservoir are conducted mostly using numerical simulation. The physical simulation is conducted in the method of increasing gas pressure, ignoring the fact that gas pressure decreases around the borehole in the coal reservoir during the CBM extraction process, which certainly leads to some errors because of adsorption and desorption irreversibility [14-15]. What is more, temperature in those studies was rarely considered. Therefore, this paper focuses on the deformation and permeability evolution of coal-containing methane when gas pressure decreases under different temperatures. The relationship between deformation and permeability could provide support for the CBM extraction program.

## 2 Experimental work

### 2.1 SAMPLE PREPARATION

In this paper, coal specimens were obtained from the Songzao Colliery, Chongqing, China. The thin coal seam exhibits high gas content, as well as relatively developed joints and fissures.

Firstly, the raw coal was crushed into powder and sieved to get pulverized coal with particle diameters between 60 mesh and 80 mesh.

Secondly, the pulverized coal was mixed with enough water and then placed in a mold.

Thirdly, the mixture was formed into a cylindrical specimen ( $\Phi 50 \text{ mm} \times 100 \text{ mm}$ ) with forming pressure of 100 MPa by using the material testing machine.

Finally, samples were dried in a drying basin and desiccated in the vessel.

### 2.2 EXPERIMENTAL APPARATUS

The self-developed triaxial servo-controlled seepage equipment for thermal-hydrological-mechanical coupling of coal-containing methane [16] was used to implement these experiments. This apparatus could test the mechanical properties as well as the flow characteristics of CBM under different axial pressures, confining pressures, gas pressures and temperatures.

### 2.3 TEST PROCEDURE

Seepage experiments under different temperatures (20°C/ 30°C/ 40°C/ 50°C/ 60°C/ 70°C) and different gas (CH<sub>4</sub>) pressures are conducted. The experiments were conducted strictly in accordance with the following test procedures:

- 1) Before the experiments, the coal specimen was fitted into the triaxial pressure chamber, which was lifted into the heated waters to keep the ambient temperature predetermined.
- 2) Applying the axial pressure and confining pressure to 6.0 MPa. And then applying the CH<sub>4</sub> injecting pressure to 3.5 MPa, the status was maintained steady until the coal specimen adsorbed gas sufficiently. Then, the deformation and flow data were recorded.
- 3) The CH<sub>4</sub> gas pressure was adjusted as follows: 3.5→3.0→2.5→2.1→1.8→1.5→1.2→0.9→0.6→0.3 MPa. At each point of the gas pressures, the corresponding data were recorded after sufficient desorption that was directly implied by constant deformation and flow data.
- 4) Each test condition was conducted repeatedly to confirm the reliability of the test results.

## 3 Results and model analysis

### 3.1 THE DEFORMATION CHARACTERISTIC

The strain evolution curves of coal-containing methane with gas pressure decrease under different temperatures (from 20°C to 70°C) are shown in Figure 1. Under constantly external stress (axial stress and confining pressure), the coal sample is gradually compressed while gas pressure decreases. The axial strain and radial strain show a linear relationship with the gas pressure. Under the 20°C condition, the axial strain and radial strain of coal-containing methane increase in a similar way on account of the decrease in gas pressure, and their variations are similar when per unit pressure decreases. As the temperature increases, the variations of axial strain and radial strain with gas pressure show a significant difference. This is because the ability to resist deformation of coal-containing methane is decreased with the increase in temperature. Meanwhile, with the decrease in gas pressure, the effective stress applied to coal-containing methane is increased, which leads to

compressive deformation gradually in the axial orientation and expansion deformation in the radial orientation of the coal sample due to anisotropic

mechanics property. Therefore, under the 70°C condition, a slight decrease in the radial strain is observed when the gas pressure decreases.

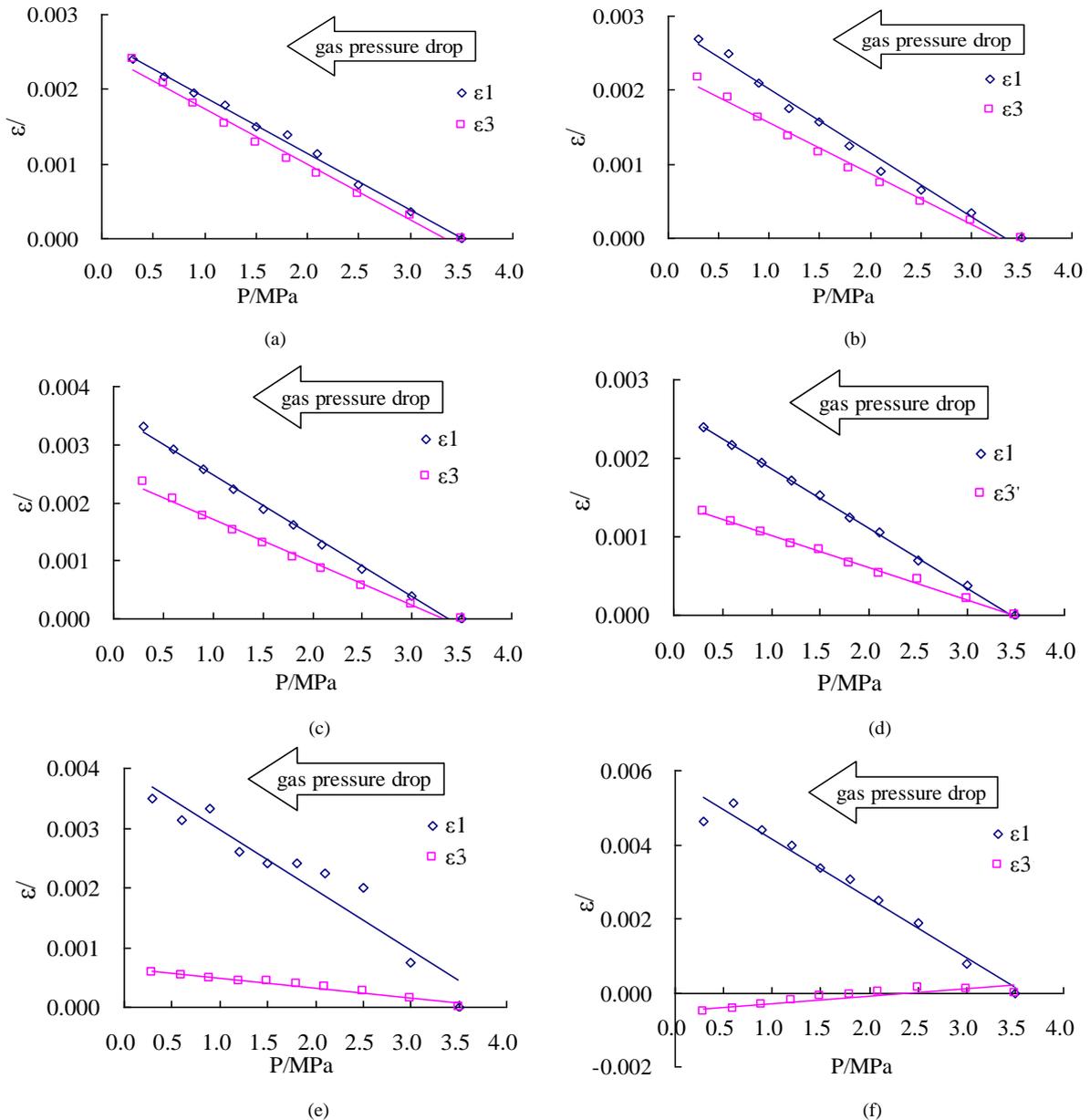


FIGURE 1 Evolutions of strain with gas pressure under different temperatures: (a) 20°C, (b) 30°C, (c) 40°C, (d) 50°C, (e) 60°C, (f) 70°C

Stress-strain relationships for an isothermal gas adsorbing coalbed may be written as [17]:

$$\varepsilon_{ij} = \frac{1}{2G} \sigma_{ij} - \left( \frac{1}{6G} - \frac{1}{9K} \right) \sigma_{kk} \delta_{ij} + \frac{\alpha p}{3K} \delta_{ij} + \frac{\varepsilon_p}{3} \delta_{ij}, \quad (1)$$

where  $\varepsilon_{ij}$  is the component of the total strain tensor,  $\sigma_{ij}$  denotes the component of the total stress tensor,  $\alpha = 1 - K/K_S$ , is the Biot coefficients,  $p$  denotes gas pressure,  $\delta_{ij}$  is the Kronecker delta,  $K$  denotes the bulk modulus of coal and fissure system,  $K_S$  represents the bulk modulus of coal matrixes,  $\sigma_{kk} = \sigma_i + \sigma_j + \sigma_k$ , is the

total bulk stress,  $\varepsilon_p$  is matrix/system strain on account of gas adsorption/desorption. Grey [18] used a linear relationship between the swelling strain  $\varepsilon_p$  and pressure in his permeability model. Eq. (1) indicates a linear relationship between the total strain tensor  $\varepsilon_{ij}$  and gas pressure  $p$  when total stress  $\sigma_{ij}$  remains constant, just as shown in the Fig.1. The otherness of deformation at different temperatures is caused by changes and anisotropy of mechanical properties induced by temperature change.

3.2 THE DYNAMIC PERMEABILITY MODEL AND ITS DISCUSS BY TEST

According to J. Liu [19], the permeability for coal matrix system can be given as:

$$k_{\infty} = \left( 1 + \frac{\alpha}{\phi_0} \cdot \frac{(\Delta\sigma_m + \Delta p)}{K} \right)^3 \cdot k_0, \tag{2}$$

where the mean compressive stress  $\sigma_m$  is denoted by  $\sigma_{kk}/3$ ,  $\phi_0$  indicates initial porosity,  $k_0$  denotes initial Klinkenberg permeability, which could be determined by actual test in the lab,  $k_{\infty}$  is the real-time Klinkenberg permeability.

According to Klinkenberg [20], effective gas permeability at a finite pressure is calculated by the following formula:

$$\begin{cases} k_p = k_{\infty} (1 + b / P_m) \\ b = 4c\lambda P_m / r \end{cases}, \tag{3}$$

where  $k_p$  is the real-time gas permeability,  $b$  is the Klinkenberg factor, dependent on the pore structure of the medium and temperature for the given gas,  $c$  denotes Klinkenberg coefficient acquired by fitting the experimental data observed in the lab,  $\lambda$  is the mean free path of gas molecular,  $P_m$  shows the average gas pressure;  $r$  indicates the effective pore radius;  $\kappa$  is Boltzmann gas constant,  $1.3806505 \times 10^{-23}$  J/K;  $T$  expresses coal temperature,  $d$  is the gas molecular diameter.

Combining Eqs (2) and (3), the following relationship is achieved as:

$$k_p = \left( 1 + \frac{\alpha}{\phi_0} \cdot \frac{(\Delta\sigma_m + \Delta p)}{K} \right)^3 \cdot k_0 (1 + b / P_m). \tag{4}$$

Eq (5) is the dynamic permeability model and reveals that it is a complicated relation between gas permeability and gas pressure, and that the Klinkenberg permeability lessens with the decrease of gas pressure. However, at a finite pressure, the slippage effect gradually dominates gas permeability. Therefore, the gas permeability decreases with the decrease of gas pressure, and then increases, which will be mainly elaborated by means of experimental test.

For the measurement of coal permeability, Darcy's law (Eq (6)) was used to interpret the experimental result [21]:

$$k = \frac{2qp_0\mu L}{A(p_1^2 - p_2^2)}, \tag{5}$$

where  $K$  is the permeability ( $m^2$ ),  $q$  is the gas permeation rate ( $m^3/s$ ),  $\mu$  is the gas kinematic viscosity ( $Pa \cdot s$ ),  $L$  is the length of the coal specimens ( $m$ ),  $A$  is the cross-sectional area of the coal specimens ( $m^2$ ),  $p_1$  is the gas pressure at the upper stream or inlet of the specimens ( $Pa$ ),  $p_2$  is the gas pressure at the downstream or outlet of the specimens ( $Pa$ ), and  $p_0$  denotes the standard atmosphere.

The permeability of coal-containing methane under different gas pressures and temperatures are shown in Table 1. The table shows that, when temperature is less than  $40^{\circ}C$ , the permeability of coal-containing methane shows a tendency to decrease firstly and then increase during the decrease process of gas pressure. If the temperature is greater than  $40^{\circ}C$ , the change in permeability is not evident in the initial stages of gas pressure decrease. Nevertheless, the permeability increases significantly when gas pressure further decreases. The turning point of gas pressure for the permeability is approximately 1.2 MPa. When gas pressure is constant, the permeability of coal-containing methane shows an overall trend that increases firstly and then decreases with the increase in temperature.

TABLE 1 The permeability of coal under different temperatures

Gas pressure (MPa)	$k_p (10^{-3} \mu m^2)$					
	20°C	30°C	40°C	50°C	60°C	70°C
3.5	0.431	0.368	0.287	0.418	0.385	0.511
3.0	0.423	0.365	0.267	0.430	0.388	0.517
2.5	0.404	0.340	0.265	0.437	0.391	0.524
2.1	0.382	0.316	0.216	0.441	0.399	0.530
1.8	0.372	0.286	0.189	0.451	0.383	0.543
1.5	0.366	0.286	0.181	0.430	0.367	0.534
1.2	0.358	0.278	0.202	0.428	0.357	0.519
0.9	0.359	0.279	0.262	0.446	0.423	0.538
0.6	0.383	0.294	0.338	0.600	0.553	0.724
0.3	0.407	0.397	0.344	0.908	0.936	1.100

To analyse the evolution of permeability with gas pressure in an intuitive way, the permeability of coal-containing methane was normalized, using the dimensionless  $K/K_0$  in the analysis, where  $K_0$  is the permeability of coal-containing methane while gas pressure is equal to 3.5 MPa at the corresponding temperature. The contrast curves of radial strain–gas pressure and dimensionless permeability–gas pressure are shown in Figure 2.

As shown in Figure 2, in the initial stage of gas pressure decrease, the radial strain of coal-containing methane reveals a trend of increase, whereas the dimensionless permeability shows an evidently opposite trend under the  $20^{\circ}C$  and  $40^{\circ}C$  conditions. The decrease in gas pressure leads to an increase in effective stress when the external stress is constant at the initial stage. As such, the gas flow channel is contractive because of the increase in radial strain, which eventually leads to the decrease in permeability. When the temperature is  $60^{\circ}C$  or  $70^{\circ}C$ , at the initial stage of gas pressure decrease, we observed little change in the radial strain and dimensionless permeability of coal-containing methane. At the initial stage of gas pressure decrease, the radial strain of coal-containing gas has a significant effect on its

permeability. The relationship curve between dimensionless permeability and radial strain when gas pressure exceeds 1.2 MPa is shown in Figure 3. As

shown in Figure 3, at the initial stage of gas pressure decrease, the permeability of coal-containing methane decreases linearly with the increase in radial strain.

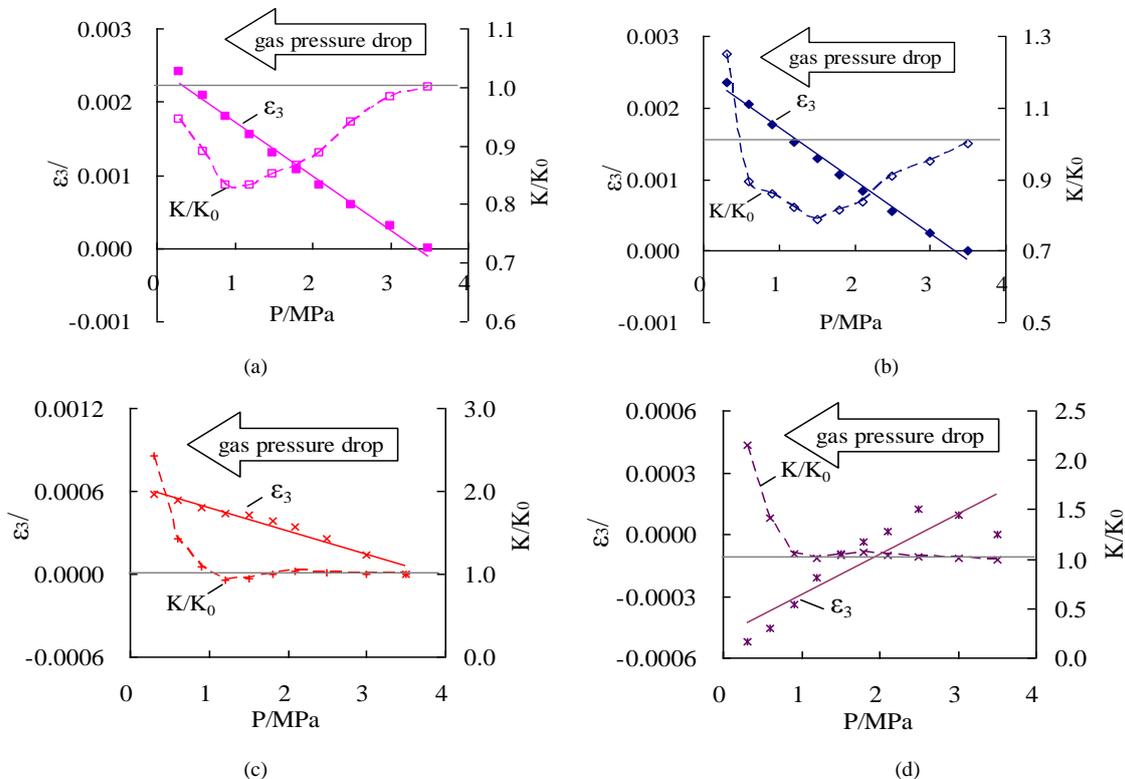


FIGURE 2 Contrast curves of radial strain-gas pressure and dimensionless permeability-gas pressure: (a) 20 °C, (b) 40 °C, (c) 60 °C, (d) 70 °C

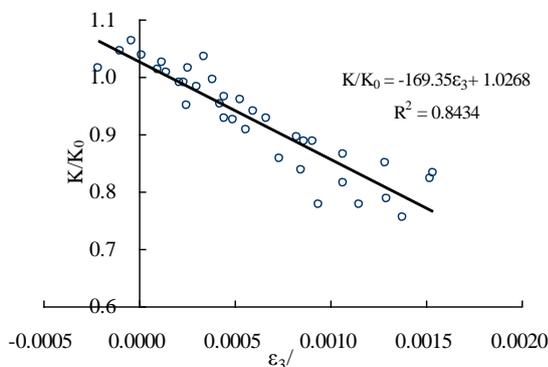


FIGURE 3 The relationship curves of  $K/K_0$  and  $\epsilon_3$  ( $P \geq 1.2$  MPa)

With further decrease in gas pressure, radial strain continues to show a linear increase. However, the permeability of coal-containing methane increases correspondingly because of the increased permeability caused by the slippage effect as shown in the Eq (4). This because the radial strain reflects only the structure deformation of coal-containing methane. Instead of structural deformation, the intensive slippage effect dominants permeability change when the gas pressure decreases less than 1.2 MPa.

With the decrease in CBM pressure, the permeability of coal-containing methane decreases first and then increases because of reservoir compression deformation caused by effective stress and intensive slippage effect.

While the temperature increases (more than 40°C in this paper), reservoir compression deformation decreases and desorption enhances, thus improving the permeability of the coal reservoir.

#### 4 Conclusions

The deformation and permeability characteristics of CBM reservoir are important factors that affect CBM recovery. By means of the experimental study on the deformation and permeability of coal-containing methane by decreasing gas pressure under different temperature conditions, the following conclusions are achieved:

- (1) With the decrease in gas pressure, the strain of coal-containing methane shows a linear increasing trend. As temperature increases, the variation in radial strain has a decreasing tendency with the decrease in gas pressure.
- (2) Under the constant temperature, the permeability of coal-containing methane shows a tendency to decrease firstly and then increase with the decrease in gas pressure. The turning point of gas pressure is approximately 1.2 MPa, less than which the intensive slippage effect dominants permeability change. At the initial stage of gas pressure decrease, the radial strain of coal-containing gas significantly affects its permeability because of effective stress change.
- (3) Elevated temperature within a certain range can alleviate the reservoir compression deformation caused

by gas extraction, as well as promote desorption to improve the extraction efficiency of CBM.

(4) To simplify the test process, gas pressure changes linearly and discontinuously with time, which is different from the gas pressure changes of during actual extraction of CBM. Nevertheless, some predictable rules are obtained through the test results, and more detailed studies on the actual problem will be considered in future work.

## References

- [1] Kong X Y 1999 *Advanced Mechanics of Fluids in Porous Media* Beijing: University of Science and Technology of China Press (in Chinese)
- [2] Deng Z, Kang Y S, Liu H L et al 2009 Dynamic variation character of coal bed methane reservoir permeability during depletion *Journal of China Coal Society* **34**(7) 947–51 (in Chinese)
- [3] Zhou J P, Xian X F, Jiang Y D et al 2009 A permeability model considering the effective stress and coal matrix shrinking effect *Journal of Southwest Petroleum University (Science & Technology Edition)* **31**(1) 4–8 (in Chinese)
- [4] Zhou J P, Xian X F, Jiang Y D et al 2010 A permeability model including effective stress and coal matrix shrinking effect *Rock and Soil Mechanics* **31**(7) 2317–23
- [5] Fu X H, Qin Y, Jiang B, et al 2003 Physical and numerical simulations of permeability of coal reservoirs in central and southern part of the Qinshui basin, Shanxi *Chinese Journal of Geology* **38**(2) 221–9 (in Chinese)
- [6] Jin Z M, Zhao Y S, He J et al 1991 An experimental study on the mechanical properties of gas-bearing coal seams *Chinese Journal of Rock Mechanics and Engineering* **10**(3) 271–80 (in Chinese)
- [7] Hu Y Q, Zhao Y S, Wei J P et al 1996 Experimental study of permeating law of coal mass gas under acting of three-dimensional stress *Journal of Xi'an Mining Institute* **16**(4) 308–11
- [8] Hu Y Q, Zhao Y S, Yang D et al 2010 Experimental study of effect of temperature on permeability characteristics of lignite *Chinese Journal of Rock Mechanics and Engineering* **29**(8) 1585–90 (in Chinese)
- [9] Cao S G, Guo P, Li Y et al 2010 Effect of gas pressure on gas seepage of outburst coal *Journal of China Coal Society* **35**(4) 595–9 (in Chinese)

## Acknowledgments

The research is supported in part by opening fund of the State Key Laboratory of Geo-hazard Prevention and Geo-environment Protection (SKLGP2013K007), in part by the Major Project of National Science and Technology (no. 2011ZX05034-004), in part by the National Natural Science Foundation of China (51304255), and in part by the National “Twelfth Five-Year” Plan for Science & Technology Support (2012BAK04B01).

- [10] Liang B, Zhang M T, Pan Y S et al 1995 The experimental research on the effect of gas on mechanical properties and mechanical response of coal *Chinese Journal of Geotechnical Engineering* **17**(5) 12–8 (in Chinese)
- [11] Li Z Q, Xian X F, Long Q M 2009 Experimental study of coal permeability under different temperature and stress. *Journal of China University of Mining & Technology* **38**(4) 523–7 (in Chinese)
- [12] Li Z Q, Xian X F 2009 Study on experiment of coal permeability with temperature and stress changing *Journal of Liaoning Technical University* **28**(S1) 156–9 (in Chinese)
- [13] Xu J, Liu D, Peng S J 2011 The impact of temperature on permeability of coal containing methane during deformation and failure process *Disaster Advances* **4**(S1) 63–8
- [14] Long Q M, Zhao X S, Sun D L et al 2008 Experimental study on coal permeability by adsorption *Journal of China Coal Society* **33**(9) 1030–4 (in Chinese)
- [15] Ma D M, Zhang S A, Lin Y B et al 2011 Isothermal adsorption and desorption experimental of coal and experimental results accuracy fitting *Journal of China Coal Society* **36**(3) 477–80 (in Chinese)
- [16] Xu J, Peng S J, Yin G Z et al 2010 Development and application of triaxial servo-controlled seepage equipment for hot-fluid-solid coupling of coal containing methane *Chinese Journal of Rock Mechanics and Engineering* **29**(5) 907–14 (in Chinese)
- [17] Shi J Q, Durucan S 2004 *Transport in Porous Media* **56** (1) 1–16
- [18] Gray I 1987 *SPE Reservoir Engineering* **2**(1) 28–34
- [19] Liu J et al 2010 Evaluation of stress-controlled coal swelling processes *International Journal of Coal Geology* **83**(4), 446–455
- [20] Klinkenberg L J 1941 The Permeability of Porous Media to Liquids and Gases *Drilling and Production Practice, American Petroleum Inst.* 200–13
- [21] Yin G Z, Jiang C B, Wang J G, Xu J 2013 *Transport in Porous Media* **100**(1) 1–16

Authors	
	<p><b>Yunqi Tao, born in March, 1979, Zhengzhou City, Henan Province, P.R. China</b></p> <p><b>Current position, grades:</b> senior engineer from School of Energy Science and Engineering, Henan Polytechnic University in China  <b>University studies:</b> Degree of Engineering Doctor in Safety Technology &amp; Engineering from Chongqing University in China  <b>Scientific interests:</b> His research interest fields include gas disaster and control, recovery and utilization of methane.  <b>Publications:</b> more than 35 papers published in various journals.  <b>Experience:</b> He has completed six scientific research projects.</p>
	<p><b>Yunqi Tao, born in June, 1986, Zhengzhou City, Henan Province, P.R. China</b></p> <p><b>Current position, grades:</b> State Key Laboratory of Coal Mine Disaster Dynamics and Control in China  <b>University studies:</b> Degree of Engineering Doctor in Mining Engineering from Chongqing University in China  <b>Scientific interests:</b> His research interest fields include Gas Disaster and Control, Recovery and Utilization of Methane.  <b>Publications:</b> more than 18 papers published in various journals.  <b>Experience:</b> He has completed two scientific research projects.</p>
	<p><b>Jie Cao, born in August, 1988, Shapingba District, Chongqing City, P.R. China</b></p> <p><b>University studies:</b> M.E. in Mining Engineering from Chongqing University of Chongqing Province in China  <b>Scientific interests:</b> Mine dynamic disaster, Gas migration law  <b>Publications:</b> 5 papers  <b>Experience:</b> has researching experience of 3 years, has involved various scientific research projects</p>
	<p><b>Jiang Xu, born in September, 1960, Chongqing City, P.R. China</b></p> <p><b>Current position, grades:</b> Professor of Chongqing University, China.  <b>University studies:</b> Degree of Engineering Doctor in Mining Engineering from Chongqing University in China  <b>Scientific interests:</b> Rock Mechanics and Engineering, Development and Utilization of Mineral Resources, Mechanism, Monitoring and Forecasting of Geological Disasters.  <b>Publications:</b> more than 100 papers published in various journals.  <b>Experience:</b> has teaching experience of 29 years, has completed 20 scientific research projects</p>

# Income distribution of the bundled transmission of photovoltaic power plant based on DEA game model

Ling Li\*

*School of New Energy Science and Engineering of Xinyu University, Jiangxi, China, 338004*

*Received 6 April 2014, www.tsi.lv*

---

## Abstract

The insufficient utilization capacity of photovoltaic (PV) power has been considered as the bottleneck for the future development of PV power in China. Nowadays, the bundled transmission mode of PV power is regarded as the most advocated solution by many scholars. Under the current unsound fiscal taxation policies in China, however, participators cannot receive the corresponding policy incentives and financial compensation for the additional contributions in constraining the implementation of the bundled transmission mode of PV power. Based on the basic theory of the DEA (Data Envelopment Analysis) Game, the allocation model of excessive profit was established. The feasibility of the developed model was verified by means of an actual case study. Arguably, the paper provided a certain theoretical basis for the quantity and practical form of excessive income distribution concerning the bundled transmission mode of PV power, and offered a solution to the income distribution for the bundled transmission mode of PV power.

*Keywords:* DEA Game Model, Photovoltaic Power Utilization, Photovoltaic Power Transmission, Income Distribution

---

## 1 Introduction

Currently, the insufficient utilization capacity of wind power has become a bottleneck for its development. The average utilization rate of many wind turbine generator systems (WTGS) is less than 30%. Therefore, in order to accelerate the development of PV power in China, appropriate countermeasures must be executed to improve the PV power utilization of power grid [1]. In the opening meeting of Research on Connecting PV Power to Power Grid and Market Utilization, organized by the National Energy Administration (NEA) and held in Beijing on March 30, 2010, this issue has been further augmented. Mr. Liu Qi, the Deputy Director of NEA, presented that [2], “the development of PV power and other new energy industries is an important strategic task in China. Strengthening the coordinative development of power grid and PV power is an important foundation for the rapid development of PV power, wherein, the study on grid connection of PV power and market utilization is a top priority.”

In order to enhance the utilization capacity of PV power, previous researchers proposed to adopt the bundled transmission mode of PV power for expanding its utilization regions and improving the stability of power grid [3-4]. Furthermore, some researches verified its scientificity and feasibility in various aspects. Bai et al. stated that, “the adjustability of other power supplies (such as thermal power and hydropower) is the decisive factor to determine whether the electrical power system can accommodate the scale of PV power or not. As the domestic wind power resources are concentrated, it is

determined that the rapid development and utilization of China’s PV power can be realized only when a firm trans-regional interconnected grid is established [5]. Therefore, effective coordination with other types of power supplies and power grid groups are required to provide the support services for the bundled transmission mode of PV power. The bundled proportion of PV power is related to the peak load regulation capacity of thermal power, hydropower, and so forth. Additionally, it is also related to the output characteristics of PV power and the depth of long-distance power transmission in participating in peak load regulation of receiver grid, etc.

Nowadays, the implementation of the bundled transmission mode of PV power has been restricted by the associated fiscal taxation policies in China. Zhang Yunzhou et al. in their paper “A Study of the Major Issues Related to the Development and Utilization of China’s PV Power” put forward the major problems including: the subsidy standard stipulated for grid connection is too low; No policy is available regarding the investment returns concerning long-distance transmission and transformation project of large-scale PV power bases; no pricing and compensation mechanism for various support services provided by large grid-connected PV power of other power plants and grids is made, etc [6]. Therefore, to ensure the smooth progress of PV power as well as other types of the bundled transmission, the aforementioned problems must be solved and related fiscal taxation policies must be revised.

Therefore, the distribution of income must be well allocated to guarantee the smooth implementation of the

---

\*Corresponding author e-mail: [llingswed@126.com](mailto:llingswed@126.com)

bundled transmission modes of PV power. In other word, all parties involved – hydroelectric power, thermal power, pumped-storage power, nuclear power, biomass power and power grid enterprises should get the corresponding incentive compensation for better coordination with the countermeasure and achieving the objective of enhanced PV power utilization capacity. In this paper, DEA Game model is proposed to solve the distribution of income among all parties involved in the bundled transmission of PV power. This in return provides a basis for establishing corresponding subsidy and incentive mechanism.

**2 DEA Game theory**

Nakabayashi and Tone extended the efficiency analysis of multiplayer game proposed by Golany and Rousseau [7]. This set up the idea of qualitative classification on DMU by means of their coalitions, to consider the cooperative game of DMUs, and construct the DEA Game model. This initiated the application foundation of DEA game though the application of DEA Game was still restricted. This is because that the number of participators and standards considered by Nakabayashi and Tone [8] is relative small and the challenge of getting game solutions is augmented with the increase of the number. Therefore, it is worthy of studying how to design the reasonable game solutions for such a problem. Li et al. [9, 10] took fixed cost proportions among DMUs into account based on DEA and coalition game. They analyzed the allocation of coalition game by the Shapley value and nucleus, respectively. Finally, the linear programming algorithms and the genetic inheritance algorithms were developed to solve the problem.

DEA Game model mainly considers the progress of reaching an agreement by multiplayers under various standards. Assume there are n parties and each person has m standards to distribute the given interests. For private interest, everyone wishes to maximize the standard that is beneficial to himself and to minimize the unfavourable standard. In this way, the given benefits are insufficient for distribution. Therefore, the allocation gets stuck into a dilemma. It will inevitably lead to inconsistency when participators determine the weight values of various indexes. However, DEA can solve such kind of problems. DEA Game is based on the assumption that every party in the game is willing to take part in the game and all participators agree to negotiate together so as to reach a reasonable and fair allocation plan. The coalition and distribution of members are also taken into account.

**2.1 COALITION OF DEA GAME AND THE CHARACTERISTIC FUNCTION**

All players are recorded as N,  $N = \{1,2,3,\dots,n\}$ . Any subset S of N is a coalition, then  $S \subset N$ . When players are conducting profit (cost) distribution, the index that is agreed to be important is denoted as  $i, i = 1,2,3 \dots m$ .

The  $i^{th}$  index score of coalition S is defined as  $x_i(S)$ :

$$x_i(S) = \sum_{j \in S} x_{ij} \quad (i = 1, 2, 3 \dots, m).$$

**2.1.1 DEA maximum game**

The characteristic function of coalition S is defined as  $C(S), (C(\phi) = 0)$ . C(S) is the maximum profit value obtained by coalition S, and C(S) is expressed by the following linear programming:

$$C(S) = \max \sum_{i=1}^m w_i x_i(S),$$

$$s.t. \begin{cases} \sum_{i=1}^m w_i = 1 \\ w_i \geq 0 (i = 1, 2, 3 \dots, m) \end{cases} \quad (1)$$

In the equation,  $w_i$  is the weight of index  $i$  under a certain coalition.

$(N, c)$  is used to represent DEA max game with the participator set of N and the characteristic function of c, which has the transferable utility. The characteristic function of DEA max game has the following properties:

- I. For coalitions of S and T, if  $S \subset N, T \subset N$  and  $S \cap T = \phi$ , then  $C(S \cup T) \leq C(S) + C(T)$ ;
- II.  $C(N) = 1$ .

**2.1.2 DEA Minimum game**

In DEA Minimum game, the characteristic function of coalition S is defined as  $d(S), (d(\phi) = 0)$ .  $d(S)$  is the minimum cost value paid by coalition S, and  $d(S)$  can be expressed by the following linear equation:

$$d(S) = \min \sum_{i=1}^m w_i x_i(S),$$

$$s.t. \begin{cases} \sum_{i=1}^m w_i = 1 \\ w_i \geq 0 (i = 1, 2, 3 \dots, m) \end{cases} \quad (2)$$

where  $w_i$  is the weight of index  $i$  under a certain coalition.  $(N, d)$  is used to represent DEA min game with the participator set of N and the characteristic function of d with the transferable function. The characteristic function of DEA min game has the following properties:

- I. For coalitions of S and T, if  $S \subset N, T \subset N$  and  $S \cap T = \phi$ , then  $d(S \cup T) \geq d(S) + d(T)$ ;
- II.  $d(N) = 1$ .

## 2.2 BENEFIT APPORTIONMENT OF DEA GAME AND THE APPORTIONMENT VECTOR

Benefit (cost) apportionment is the nucleus content of DEA game, as a decisive role in the stability of cooperative alliance. Once a party suffers from unfair treatment, the already formed coalition is at the risk of disintegration. Therefore, to maintain the stability of the cooperation, the benefit (cost) apportionment should satisfy definite rationality.

Assuming that the cooperative cost apportionment can be simplified as a transferable apportionment or utility, namely with side payment (transferable payment), the side payment shall meet the following requirements:

- Every player measures his apportionment with the same utility scale;
- The apportionment of each coalition can be distributed to each participator in any manner. In other words, the apportionment of all players is transferable.

$n$ -dimensional vector  $z = \{z_1, z_2, \dots, z_n\}$  is used to represent shares distributed from the coalition apportionment to each player of DEA game. The following condition shall be satisfied:

- Rationality of individuals:

$$z_j \geq C(j) \text{ or } z_j \leq d(j), (j=1, 2, 3, \dots, n). \quad (3)$$

- Rationality of collectives:

$$\sum_{j=1}^n z_j = C(N) = 1 \text{ or } \sum_{j=1}^n z_j = d(N) = 1. \quad (4)$$

## 3 Construction of income distribution concerning the bundled transmission of PhotoVoltaic (PV) Power Based on DEA Game Model

### 3.1 MODEL PARAMETERS

As an obstacle of recent research, the distribution of income is related to the immediate interests among every participator of the bundled transmission of PV power.

#### 3.1.1 Determination of excessive profit

The distributed interest in the model refers to the excessive profit of the bundled transmission of wind power, considering the practical planning and operation of electric power system. Under the normal operation of the the bundled transmission of wind power, power grid enterprises will settle with each participator according to the existing system. Combined with the actual cost statement delivered from each participator, the power grid enterprises are compared with previous parameters every half a year to estimate the excessive profit of this section. That is to say, the excessive profit of the bundled transmission of wind power ( $M$ ) = profit of utilized wind

power capacity obtained from the bundled transmission (PU) - profit of utilized wind power capacity that is not obtained from the bundled transmission under current grid development condition (PD) + contribution amount of utilized wind power capacity to the social benefit (PE). The contribution amount of the utilized wind power capacity to the social benefit mainly covers environmental benefits including the saved non-renewable energy resources and so on.

#### 3.1.2 Participators of profit apportionment

Wind power plant – provides wind power capacity and bears the cost of wind power generation.

Power grid enterprise – constructs trans-regional power grid and undertakes the assurance of risk and cost of grid stability after the utilized PV power generation.

Grid-connected power station refers to the power plant used for the bundled transmission (including thermal power plant, hydroelectric power station, pumped-storage power station, gas turbine power station and nuclear power station, etc.) – Except for the basic services, they should also be compensated for the provided support services of automatic generator control (AGC), paid peak load regulation, alternative, paid reactive power regulation, black start and so on.

The support services of grid-connected power station can be divided into the basic support services and the compensated support services. The basic support services indicate the support services that must be provided by generator set to ensure safe and stable operation of electric power system and power quality including primary frequency, basic peak load regulation, basic reactive power regulation, etc. For this part of support services, no distribution of excessive profit concerning the bundled transmission of wind power is required. Only the compensable support service, which is provided by other grid-connected power stations and is included in the bundled transmission of wind power, can participate in the distribution.

### 3.2 BASIC MODEL OF INCOME DISTRIBUTION

This paper mainly involves the application of DEA game in the income distribution of the the bundled transmission mode of PV power. A new scheme has been proposed to take advantage of DEA game model for income distribution. Firstly, all the players, namely all the profit sharers of the the bundled transmission mode of PV power, are recorded as  $N$ ,  $N = \{1, 2, 3, \dots, n\}$ . Any subset  $S$  of  $N$  is a coalition, then  $S \subset N$ . When the players are conducting profit (cost) distribution, the index that is agreed to be important is kept down as  $i$ ,  $i = 1, 2, 3, \dots, m$ . And the index score of each player under every index is evaluated together.  $x_{ij}$  is the index score of the  $j^{\text{th}}$  player under the  $i^{\text{th}}$  index,  $j = 1, 2, 3, \dots, n$ . The higher  $x_{ij}$  under

an index reveals that the  $j^{\text{th}}$  player has higher evaluation under this index. The matrix constituted by all index scores is set as  $X$ , namely  $X = \{x_{ij}\}_{m \times n}$ . And matrix  $X$  is standardized, namely  $\sum_{j=1}^n x_{ij} = 1 (i = 1, 2, 3 \dots, m)$ .

The  $i^{\text{th}}$  index score of coalition  $S$  is defined as  $x_i(S)$ :

$$x_i(S) = \sum_{j \in S} x_{ij} (i = 1, 2, 3 \dots, m) \tag{5}$$

Characteristic function of coalition  $S$  is defined as  $C(S), (C(\emptyset) = 0)$ .  $C(S)$  is the maximum profit value obtained by coalition  $S$ , and  $C(S)$  is expressed by the following linear programming:

$$C(S) = \max \sum_{i=1}^m w_i x_i(S),$$

$$s.t. \begin{cases} \sum_{i=1}^m w_i = 1 \\ w_i \geq 0 (i = 1, 2, 3 \dots, m) \end{cases} \tag{6}$$

In the equation,  $w_i$  is the weight of index  $i$  under a certain coalition. Obviously,  $C(N) = 1$ . Unit profit share obtained by each person is expressed by vector  $Z$ :  $z = \{z_1, z_2, z_3, \dots, z_m\}$ .

### 3.3 SOLUTIONS

In the cooperative game, many significant methods can be used to analyze and solve the game, including negotiation set, stabilization set, nucleus, nucleolus and Shapley value. Since the solution of nucleus must be unique and feasible, nucleus is adopted in the paper to solve the model so as to enable the players to get a fair and reasonable profit distribution amount.

Let  $e(S, z) = C(S) - \sum_{i \in S} z_i$ , then  $e(S, z)$  is the difference value between total excessive profit obtained by profit sharers when forming coalition  $S$  and the actual excessive profit obtained by sharers. The greater the difference, take this strategy is not ideal. Since there are  $2^n$  subsets of  $N$ , there are also  $2^n$  subsets of  $e(S, z)$ . They can be restructured into a vector according to the order from small to large  $\theta(z) = (\theta_1(z), \theta_2(z) \dots \theta_{2^n}(z))$ .

The nucleus is defined as:

$$N(V) = \{z \in E(C) / \theta(z) \leq \theta(y), \forall y \in E(C)\} \tag{7}$$

wherein,  $E(C)$  is the set of all allocation vectors.

As known from the above equations, the nucleus is a kind of distribution that allows the minimum excessive vector. All coalitions that might be formed in the cooperation have a definition of excessive value at the nucleus. The nucleus solution can be realized by the following linear programming:

$$\min \varepsilon = e(S, z) = C(S) - \sum_{i \in S} z_i,$$

$$s.t. \begin{cases} \sum_{i \in S} z_i + \varepsilon \geq C(S) \\ \sum_{i \in N} z_i = C(N) \end{cases} \tag{8}$$

In the equation,  $\varepsilon$  is an arbitrary small real number;  $N$  is the set of excessive profit sharers of the the bundled transmission mode of PV power; and  $S$  is the nonvoid subset of  $N$ . This is a standard linear programming problem and it can be solved by virtue of Matlab tool.

### 4 Empirical study on income distribution concerning the bundled transmission of PhotoVoltaic (PV) Power Based on DEA Game Model

To enhance the utilization ability of PV power,  $N$  persons are involved in the bundled transmission mode of PV power to split its excessive profit. It means 4 participators of DEA Game if  $N=4$  is selected. It is assumed that the estimate of excessive profit in the first half year of 2012 is 80 million Yuan according to the equation  $M = PU - PD + PE$ .

#### 4.1 CASE STUDY

Firstly, 4 players are defined as: A – wind power plant participated in the bundled transmission mode of PV power; B – power grid enterprises to maintain the stability of power grid and trans-regional power grid construction after the establishment of the bundled transmission mode of PV power; C – grid-connected thermal power plant that participates in the bundled transmission mode of PV power; and D - grid-connected hydroelectric power station, respectively.

Secondly, indexes that are agreed to be important by 4 players A, B, C and D are listed: power generating capacity, maintenance of power grid stability, superior alternative emergency power supply for peak load regulation, AGC, workload of reactive power regulation and additional investment of basic construction. Experts are employed to score the above 6 indexes according to the Performance Table of Semi-Annual Power Generating Capacity, Power Supply Reserve for Peak Load Regulation and the Use Condition Table, Analysis Table for Workload of Reactive Power Regulation and Table for Additional Investment Condition of Basic Construction delivered by participators. It should be noted that regarding the continuous distribution of excessive profit, the scores of six indexes will vary dependently on different contribution of involved parties. In order to ensure the accuracy of income distribution, the scoring is needed half a year for the changed materials.

According to the materials in the first half year of 2012 provided by four participants, the above six indexes are scored as shown in Table 1.

TABLE 1 Each Index Score of Every Player

	A	B	C	D
Power Generating Amount	0.5	0	0.3	0.2
Maintenance of Power Grid Stability	0.1	0.6	0.2	0.1
Superior Emergency Power Supply Reserve for Peak Load regulation	0.1	0	0.5	0.4
AGC	0.2	0	0.4	0.4
Workload of Reactive Power Regulation	0.1	0	0.5	0.4
Additional Investment of Basic Construction	0.1	0.7	0.1	0.1

Finally, the interest allocations of various alliances are obtained. The characteristic values of function, under different coalitions are calculated by equations (5) and (6), as shown in Table 2.

TABLE 2 All Kinds of Combination Alliances

Method of coalition	Nominal amount of excessive profit	
Independent	A	0.5
	B	0.7
	C	0.5
	D	0.4
Allied by two groups	AB	0.8
	AC	0.8
	AD	0.7
	BC	0.8
Allied by three groups	BD	0.8
	CD	0.9
	ABC	0.9
	ABD	0.9
Allied by four	ACD	1
	BCD	0.9
	ABCD	1

4.2 SOLUTION OF CASE MODEL

The above characteristic values of functions are substituted in (8) and further results are obtained:

$$\min \varepsilon = e(S, z) = C(S) - \sum_{i \in S} z_i, \tag{9}$$

$$s.t. \left\{ \begin{array}{l} z_A + \varepsilon \geq 0.5 \\ z_B + \varepsilon \geq 0.7 \\ z_C + \varepsilon \geq 0.5 \\ z_D + \varepsilon \geq 0.4 \\ z_A + z_B + \varepsilon \geq 0.8 \\ z_A + z_C + \varepsilon \geq 0.8 \\ z_A + z_D + \varepsilon \geq 0.7 \\ z_B + z_C + \varepsilon \geq 0.8 \\ z_B + z_D + \varepsilon \geq 0.8 \\ z_D + z_C + \varepsilon \geq 0.9 \\ z_A + z_B + z_C + \varepsilon \geq 0.9 \\ z_A + z_B + z_D + \varepsilon \geq 0.9 \\ z_A + z_C + z_D + \varepsilon \geq 1 \\ z_B + z_C + z_D + \varepsilon \geq 0.9 \\ z_A + z_B + z_C + z_D = 1 \end{array} \right. \tag{10}$$

The presented procedure is employed to solve the multi-objective programming problem. Since the wind power plant must exist as a participator of the bundled transmission mode of PV power, the abovementioned alliance set thus has invalid coalition (supposing the weight of invalid coalition is 0). To deal with this part of the invalid alliance, the linear weighted sum method of Matlab multi-objective programming problem is adopted. The objective functions for each scenarios and the given weight factors are shown as Table 3.

TABLE 3 The Objective Function and Its Given Weight

Serial No.	Objective Function	Weight
1	$\min \varepsilon = 0.5 - z_A$	0.05
2	$\min \varepsilon = 0.7 - z_B$	0
3	$\min \varepsilon = 0.5 - z_C$	0
4	$\min \varepsilon = 0.4 - z_D$	0
5	$\min \varepsilon = 0.8 - (z_A + z_B)$	0.1
6	$\min \varepsilon = 0.8 - (z_A + z_C)$	0.1
7	$\min \varepsilon = 0.7 - (z_A + z_D)$	0.1
8	$\min \varepsilon = 0.8 - (z_B + z_C)$	0
9	$\min \varepsilon = 0.8 - (z_B + z_D)$	0
10	$\min \varepsilon = 0.9 - (z_D + z_C)$	0
11	$\min \varepsilon = 0.9 - (z_A + z_B + z_C)$	0.15
12	$\min \varepsilon = 0.9 - (z_A + z_B + z_D)$	0.15
13	$\min \varepsilon = 1 - (z_A + z_C + z_D)$	0.15
14	$\min \varepsilon = 0.9 - (z_B + z_C + z_D)$	0
15	$\min \varepsilon = 1 - (z_A + z_B + z_C + z_D)$	0.2

The multi-objective linear programming of Matlab linear weighted sum method is expressed as below:

- Evaluation function of a linear weighted sum method is established as follow,

$$\begin{aligned} \min h(F(x)) = & \lambda_1(0.5 - z_A) + \lambda_2[0.8 - (z_A + z_B)] \\ & + \lambda_3[0.8 - (z_A + z_C)] + \lambda_4[0.7 - (z_A + z_D)] \\ & + \lambda_5[0.9 - (z_A + z_B + z_C)] + \lambda_6[0.9 - (z_A + z_B + z_D)] \\ & + \lambda_7[1 - (z_A + z_C + z_D)] + \lambda_8[1 - (z_A + z_B + z_C + z_D)]. \end{aligned} \tag{11}$$

The corresponding weights into the above equation are substituted, and then the objective function  $\min h(F(x))$  can be sorted out as follow,

$$\min h(F(x)) = 0.875 - (z_A + 0.6 \times z_B + 0.6 \times z_C + 0.6 \times z_D). \tag{12}$$

- The equations are solved with the aid of commercial software MATLAB:  
 $z_A = 0.4162, z_B = 0.2618,$   
 $z_C = 0.1852, z_D = 0.1368.$
- Excessive profits are distributed as: A of 33.296 million Yuan; B of 20.944 million Yuan; C of 14.816 million Yuan; and D of 10.944 million Yuan.

## 5 Model analysis

From the presented model, targeting at the finite non-renewable energy and environmental pollution that may be caused at present, PV power has become the mainstream of future power supply development, and its utilized amount by power grid will be the key factor to determine whether it can be developed into the mainstream or not. Therefore, to make power supply constituents in a sustainable way, various support services provided by participators should receive corresponding compensation incentives to guarantee the implementation of various policies encouraging the utilization of PV power. Income distribution scheme provided previously expressed the distribution principle with three key issues: firstly, the paper provisionally assumes that the materials provided by grid companies combined parties involved are estimated every six months, but the specific estimation method still needs further study. Secondly, in the scoring six key indexes must be determined according to the contribution degree of various participators. Thirdly, due to the existence of subjective judgment in the model, a comprehensive collection of information is required for all parties. The information shall be reasonable and comprehensive. Briefly, the excessive profit distribution of the bundled transmission mode of PV power is a challenging research topic. In the practical applications, the material submitted by participators should be followed strictly to estimate the excessive profit as accurately as possible; and the value of each index should be strictly determined in order to ensure the relative accuracy of the allocation result.

## References

- [1] *Renewable Energy and Power Grid 2009* Translated by New Energy Research Laboratory of China Electric Power Research Institute Beijing: China Electric Power Press 22-140 (in Chinese)
- [2] *Opening Meeting of Research on Connecting PV Power to Power Grid and Market Utilization, Chinese Photovoltaic Materials and Equipment Website* 2010 <http://www.cnwpem.com/22/7/7897.html> / 30 Nov 2010
- [3] Wang Ningbo 2010 The Analysis of the Causes of Wind Power Bottleneck Development and Researches on its Counter-measures *Energy of China* **32**(3) 17-20 (in Chinese)
- [4] Wang Ningbo 2011 Challenge Faced by Jiuquan Kilowatt Photovoltaic Power Base of Gansu and Its Countermeasures. *Power System and Clean Energy* **30** (7) 43-7 (in Chinese)
- [5] Bai Jianhua, Xin Songxu, Jia Dexiang et al 2010 A Study of the Major Issues Related to the Development, Utilization and

## 6 Conclusions

Based on the state of art of PV power, the enhancement of PV power utility capacity is proposed as a key to the rapid development of PV power in future. The most highly praised expansion method of PV power utility capacity is the bundled transmission mode of PV power. In light of the constrained situation by some policy towards the implementation of the bundled transmission mode of PV power, DEA Game is proposed in the paper to solve the problem of excessive profit distribution. Based on the hypothesis that all players are willing to participate in the game and consult together to reach a fair and equitable distribution program, a DEA Game distribution model of excessive profit has been established. Through systematically case studies, the feasibility of the model is proved to provide a theoretical basis for the excessive profit distribution regarding the bundled transmission mode of PV power. It allows the participators to obtain appropriate compensation as an encouragement for them to participate in the bundled transmission mode of PV power in a better way. Then the utility capacity of PV power can be expanded.

In the design of excessive profit distribution model for the bundled transmission mode of PV power, the estimation of excessive profit and the determination of six key index scores are required to be combined with a lot of the relevant materials. More concise and accurate methods still are needed in the further studies.

Transmission of China's Photovoltaic Power *Power System and Clean Energy* **26**(1) 14-7 (in Chinese)

- [6] Zhang Yunzhou, Bai Jianhua, Xin Songxu 2010 A Study of the Major Issues Related to the Development and Utilization of China's PV Power *Energy Technology and Economics* **22**(1) 1-6 (in Chinese)
- [7] Golany B, Rousseau J J 1992 Efficiency evaluation games *F Y Phillips and Rousseau J J eds Systems and Management Science by External Methods* New York: Springer US 327-47
- [8] Nakabayashi K, Tone K 2006 Egoist's Dilemma: A DEA Game *OMEGA International Journal of Management Science* **34**(2) 135-48
- [9] Li Yongjun, Liang Liang 2008 Fixed Cost Apportionment Method Based on DEA and Coalition Game *Systems Engineering – Theory & Practice* **28** (11) 80-4 (in Chinese)
- [10] Li Yongjun, Liang Liang, Ling Liuyi et al 2009 Research on Fixed Cost Apportionment Method Based on DEA Coalition Game Nucleolus Solution *China Journal of Management Science* **17**(1) 58-63 (in Chinese)

## Author



Ling Li, born in August, 1981, Xinyu County, Jiang Xi Province, P.R. China

**Current position, grades:** lecturer of School of Department of New Energy Science and Engineering, Xinyu College, China.

**University studies:** BSc in Physics and electronic of YunNan normal University in China. She received her MSc from Solar Energy Research Institute of YunNan normal University in China.

**Scientific interests:** PV system, solar energy

**Publications:** more than 10 papers

**Experience:** has teaching experience of 6 years, has completed three scientific research projects.

# Optical measurement method for dynamic mechanical testing based on image grey level distribution difference model

Zhiqiang Yin<sup>1, 2\*</sup>, Lei Wang<sup>1, 2</sup>, Haifeng Ma<sup>3</sup>

<sup>1</sup>The Provincial Key Laboratory of mining effects and disasters preventing under deep mining in Anhui, Anhui University of Science and Technology, Huainan, Anhui, China, 232001

<sup>2</sup>School of Mineral and Safety, Anhui University of Science and Technology, Huainan, Anhui, China, 232001

<sup>3</sup>Faculty of Resources and Safety Engineering, China University of Mining and Technology, Beijing, China, 100083

Received 26 October 2013, www.tsi.lv

## Abstract

This study developed an optical measuring system with a gray level distribution difference (GLDD) model, and applied the system to examine the displacement field of a Brazilian disk (BD) split under dynamic loading. The system consists of high-speed (HS) photography, a split-Hopkinson pressure bar (SHPB), a synchronization control system and operation of differential images. First, we captured differential images with a high speed camera (10 frames at a time resolution of 10 $\mu$ s). Next, we established the corresponding relationship between the dynamic fracturing evolution of the disc rock samples and the stress loading process with a synchronization controlling system. Changes in the surface displacement field were calculated with the differential image base method according to the joint probability distribution function of two images. This method takes the image correlation into account and can effectively eliminate the influence of background noise, identify surface displacement and capture cracks and expansion in dynamic Brazilian disk splitting experiments straightforwardly and accurately. Findings can be used for novel measurement of surface displacement fields in Brazilian disk splitting tests under dynamic loads.

*Keywords:* grey level distribution difference model, SHPB, dynamic load, high speed photography, differential image

## 1 Introduction

In recent years, computer digital imagery has been widely used to measure full-field deformation in material and structure mechanics [1]. Recently, new static loading test procedures have been developed to calculate surface displacement and deformation based on digital images of good quality with standard cameras [2]. However, in dynamic loading tests (impact, Hopkinson bar tests), relatively small, higher resolution images are required for the displacement measurement. So far, the development of methods that handle dynamic events is still in its infancy [3].

Traditional contact measurement techniques using sensors such as mechanical extensometers and strain gauges have limitations in frequency response and measuring range. Therefore, they cannot provide sufficient information to explore the rapid variation of dynamic mechanical behaviours. Non-contact, full-field monitoring techniques such as caustics, Moiré, photo elasticity, digital image correlation and coherent gradient sensors have substantial advantages for measuring deformation and stress fields in dynamic loading tests [4]. However, most of these techniques require many optical components. These are often quite expensive, and are only applicable for transparent materials such as organic polymers and inorganic glass. They are not suitable for

deformation measurements in rock dynamic experiments.

Many studies have explored deformation and crack characteristics of rock materials under dynamic loading [5, 6]. High-speed cameras and digital image methods have been developed in recent years. For example, Siviour used a high speed camera for a high strain rate experiment to monitor specimen deformation and analyse field methods for three point bending. Louis Ngai Yuen Wonga used digital image subtraction techniques to inspect cracking processes in Carrara marble specimens containing a single, pre-existing open flaw under dynamic loading.

Digital image subtraction is a common tool to analyse image changes. With this standard tool, most researchers are already familiar with resulting difference images. Taking a simple subtraction between two images is directly equivalent to forming a null hypothesis test statistic, assuming that the expected change is due to uniform noise. Therefore, this can be described by a single distribution across the entire image [7]. However, the background noise in dynamic experiments of digital images is uneven, and is affected by light sources and camera equipment. Therefore, an effective processing technique is essential to remove background noise from dynamic load images.

In this work, Brazilian disc specimens were loaded with a SHPB system, and a high-speed camera was used

\* Corresponding author e-mail: zhqyin@aust.edu.cn

to monitor the full-field dynamic fracture process in the specimens. Surface deformation and failure mode of specimen were analysed, based on the grey level distribution difference (GLDD) model.

**2 GLDD model**

The GLDD model is used to calculate difference images using probability distributions in the normalized scatter gram. This defines a probability that reflects how likely it is that grey level values from corresponding pixels in an image pair are drawn from the same distribution as the rest of the data. Corresponding pairs of pixels from original images are taken, and their grey levels are used to find their coordinates in the normalized scatter gram [7]. Integration is then performed along the vertical cut passing through that point *c*, summing all of the values *F(x, y)* that are smaller than the value of *F(x, c)* at that point.

$$D(x, y) = \sum_c \delta(F(x, c) > F(x, y))F(x, c) . \tag{1}$$

The images of *g<sub>1</sub>(x, y)* and *g<sub>2</sub>(x, y)* are the grey values of the subset in the non-deformed and deformed images, respectively. The variable *m* and *n* are the probability distribution density of the grey values of images, respectively. *S(m, n)* is probability distribution of the pixel gray distribution with *g<sub>1</sub>* and *g<sub>2</sub>*. The image difference *S* can reflect the probability of grey values different between two images in the same region. Therefore, when the image of *g<sub>1</sub>(x, y)* and *g<sub>2</sub>(x, y)* are completely identical, *S*=0. When the background noise of images shows no change, *S* is the direct finite difference, such as:

$$S = abs[g_1(x, y) - g_2(x, y)] . \tag{2}$$

When the background noise of images changes, *S* are equal to stack values of the weighted average background noise and grey. This eliminates the effect of background noise.

**3 Test systems**

The experimental set up for the SHPB impact test in this study is shown in Figure 1. The SHPB system consisted of a gas gun, a conical bullet, and a stress transmission component [8]. A conical bullet was used in the improved test system to eliminate oscillation and obtain a stable half-sine wave loading [9]. The stress transmission component was made up of three elastic bars, including an incident bar (2m in length), a transmission bar (1.5m), and a momentum bar (0.7m). The rock specimen was sandwiched between the incident bar and the transmission bar. The conical bullet and elastic bars made from 50 mm diameter high strength steel (40Cr), and have yield strength of 800MPa. Two sets of strain gauges are glued on the surface of the middle of the incident bar (SG<sub>in</sub>) and the transmission bar (SG<sub>Tr</sub>) respectively. Each set

consisted of two gauges in a symmetrical arrangement. These strain gauges are used to measure the strain histories induced by the stress waves propagating along the elastic bars. The strain history during the test was recorded with a digital oscilloscope (DL 750) though a differential amplifier (super dynamic meter CS-1D) and a Wheatstone bridge connected to strain gauges.

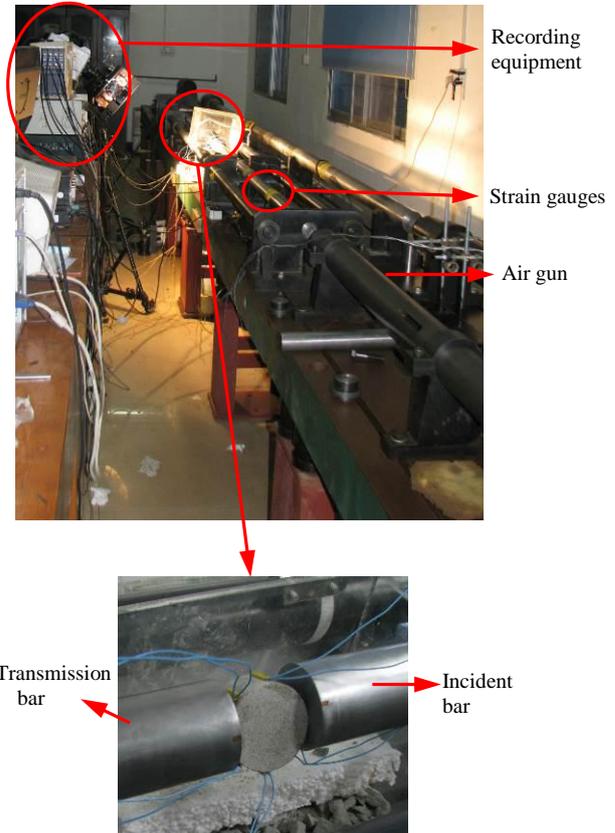


FIGURE 1 Experimental set-up for SHPB at CSU

During a SHPB test, two basic requirements must be satisfied: the specimen is under one-dimensional stress and the specimen is in-loaded evenly [10]. The equation of motion can be written as:

$$\rho \frac{\partial^2 u}{\partial t^2} = E \frac{\partial^2 u}{\partial x^2} , \tag{3}$$

where *u* is the location of the infinitesimal at *x* under stress; *E* is the elastic modulus of bars; *ρ* is the density of the bar; *t* is the time of stress loading; *x* is the location of the infinitesimal.

The half-sine wave produced by the conical bullet lengthens the rise time of the incident stress wave, allowing the specimen to be in-loaded evenly during the rise time. Since the strains in the incident and transmission bars are known, the dynamic loading forces (*P*) at the two elastic bars/ specimen interfaces may be calculated as [11]:

$$P_1 = E_b A_b (\epsilon_1 + \epsilon_R) , P_2 = E_b A_b \epsilon_T , \tag{4}$$

where  $E_b$  is the elastic bar's Young's modulus,  $A_b$  is the cross section area of the elastic bars,  $\varepsilon_i$  and  $\varepsilon_r$  are the incident and reflected strains measured by the strain gauges on the incident bar, and  $\varepsilon_T$  is the transmitted strains measured by the strain gauges on the transmission bar.

If the dynamic forces ( $P_1$  and  $P_2$ ) on both sides of the specimen are almost equal during the entire dynamic loading period, then the specimen is said to be in stress equilibrium, and the applied force on the specimen ( $P(t)$ ) can be calculated as:

$$P(t) = \frac{1}{2}(P_1 + P_2) = \frac{1}{2}E_b A_b (\varepsilon_i + \varepsilon_r + \varepsilon_T). \quad (5)$$

The impact strain  $\varepsilon(t)$  are calculated by one-dimensional stress wave theory using the following formulas:

$$\varepsilon(t) = -\frac{2c_e}{L_S} \int_0^t \varepsilon_r(t') dt', \quad (6)$$

where  $c_e$  is the elastic wave velocity of the bars;  $L_S$  is the length of sample;  $\varepsilon_r(t)$  is the strain of the reflected stress wave at time  $t$ .

Under the condition of stress equilibrium, the dynamic tensile strength ( $\sigma_{td}$ ) can be derived by using the applied dynamic load ( $P(t)$ ) and the time-to-fracture ( $t_f$ ):

$$\sigma_{td} (\sigma_{td}) = \frac{2P(t_f)}{\pi DB}, \quad (7)$$

where  $D$  is the diameter of the BD specimens,  $B$  is the thickness of the BD specimens.

The loading rate ( $\dot{\sigma}_{td}$ ) is determined by the slope of the stress history starting from the time of stress equilibrium ( $t_{equil}$ ) and ending at the time-to-fracture ( $t_f$ ).

#### 4 Sample preparation

The rock material used in this study was fine-grained sandstone found in the Zigong region of Sichuan, China. BD tests were performed to measure surface deformation. BD testing specimens are 50 (diameter)  $\times$  20 (thickness) mm. The non-parallelism and the non-perpendicularity of specimens were both less than 0.02 mm. The specimens were grey and smooth on surface, with no distinct interspaces. The density of specimens was 2.50t/m<sup>3</sup>.

To representatively capture the time-to-fracture of the specimen, three strain gauges were mounted on the specimen surface, perpendicular to the load direction along the specimen centre line with equal spacing, as shown in Figure 2.

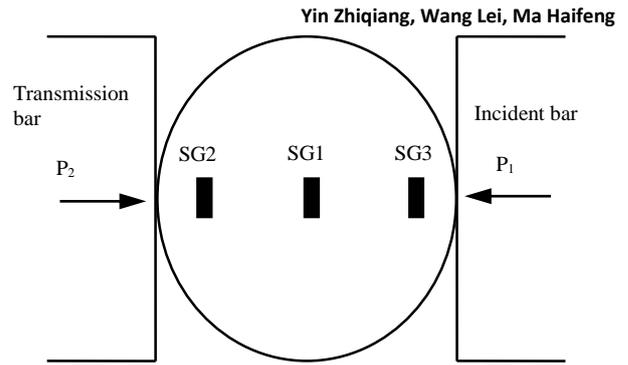


FIGURE 2 Arrangement of strain gauges on specimen

#### 5 High speed camera system

In this study, a high-speed camera (PHOTRON FASTCAM SA1.1) was used to record the full-field dynamic fracture process of specimen photographically during tests, coupled with a high-strength and non-stroboscopic light source (PALLITE VIII), positioned at a safe distance of 0.7 m from the specimen surface. The high-speed camera was operated at the setting for dynamic tests, with a frame-rate of 100,000 frame per second (10 $\mu$ s inter-frame time), and 192 $\times$ 192 pixels for size of 56 $\times$ 56 mm<sup>2</sup>. The specimen was randomly speckled with black and white paint to ensure good contrast in the digital images. The HS image of the dynamic Brazilian disk testing methods is shown in Figure 3.

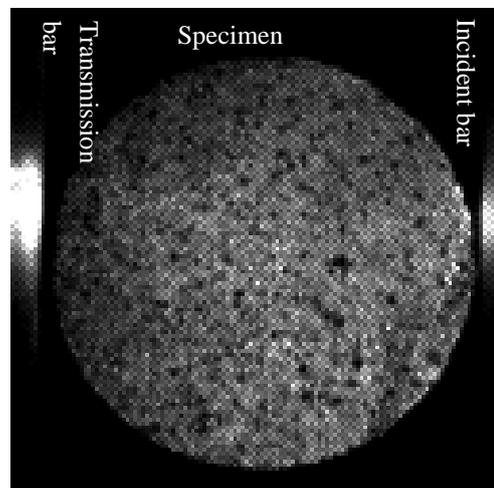


FIGURE 3 Photographic view of dynamic testing

A triggering system is composed of strain gauges on the surface of the incident bar and the oscilloscope. The incident stress wave generated by the impact of the conical bullet will propagate along the incident bar, and a TTL electrical signal will be generated when the oscilloscope detects this stress wave signal by strain gauges [12]. The camera was connected to the oscilloscope using a coaxial cable (about 2m in length). Then the camera was triggered by the TTL pulse from the oscilloscope. This pulse was generated by the strain

gauge on the incident bar, so that the number of captured images at the specimen could be obtained:

$$n = \frac{t_s - t_{in} - t_{TTL}}{t_{frame}} = \frac{t_s - s/c_e - t_{TTL}}{t_{frame}}, \quad (8)$$

where  $t_s$  is the time of the specimen HS image,  $t_{in}$  is the time of the stress wave arriving at the sample,  $s$  is the distance between the specimen and SG<sub>in</sub>,  $t_{TTL}$  is the pre trigger time by a TTL pulse, and  $t_{frame}$  is the time interval of the inter-frame.

In this test, to match the recorded images with the loading steps, the delay time from the loading start time to the triggering start time was determined to be 188 $\mu$ s. This was based on the combined consideration of the travelling time from the strain gauge to the specimen end (wave velocity 5410 m/s, distance 1017mm) and the pre-trigger time by a TTL pulse 26 $\mu$ s, as shown in Figure 4.

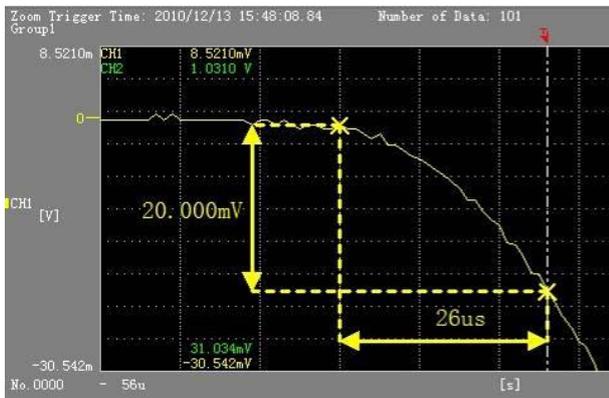


FIGURE 4 The pre trigger time measurement

## 6 Results and discussion

Figure 5 represents a typical testing result with a striking velocity of 3 m/s. Figure 6 indicates that the time of stress equilibrium was approximately 48 $\mu$ s, and the time-to-fracture was about 80 $\mu$ s, according to the time-to-peak of load and time-to-fracture of SG1. The electric potential of SG1 increased gradually from the time of stress equilibrium to the stress peak. The electric potential of SG2 and SG3 showed a similar trend, but lagged behind SG1. Therefore, it can be inferred that the cracks began at the center position (position of SG1) of the sample under dynamic loads, and extended to the two ends of the sample.

Figure 7 shows the tensile stress time history marked with corresponding points of high speed images. The dynamic tensile strength was about 14.2MPa at the loading rate of 202GPa/s.

The first image (0 $\mu$ s) was chosen as the reference image, as shown in Figure 8a. The difference images are shown in Figure 8b. The results of image difference technique reproduced the main deformation process and validate the accurate of BD test under dynamic loads. The

white spots in Figure 8b are regarded as the deformed area on the surface of the rock specimen.

Detailed observations of the patterns before 48  $\mu$ s indicate that the deformation initiated at the right end of the rock disk under dynamic loads at about 2  $\mu$ s. Then the deformation expanded leftwards and arrived at the left end of the rock disk at around 42  $\mu$ s. This result is consistent with the dynamic, photoelastic numerical simulation on the rock specimens [13].

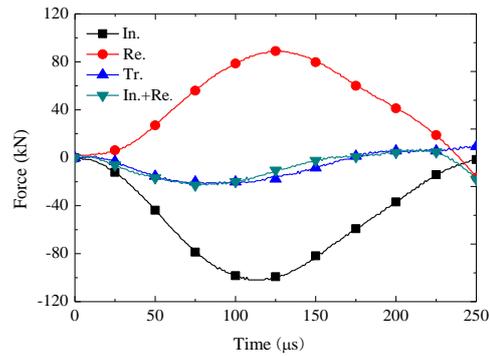


FIGURE 5 Dynamic force balance check for the dynamic Brazilian experiment

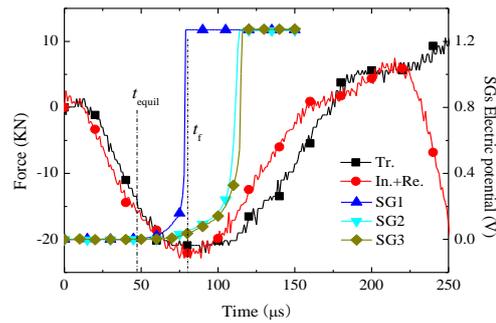


FIGURE 6 Signals from strain gauges on specimen

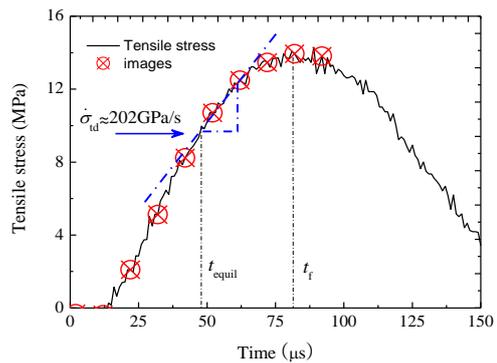


FIGURE 7 Stress-time history with high speed images

Between the time of stress equilibrium (48  $\mu$ s) and the time-to-fracture (80  $\mu$ s), white spots increased with tensile stress (Figure 7). After the time-to-fracture (80 $\mu$ s), observable cracks (block line in Figure 8b) (82  $\mu$ s, 92  $\mu$ s) were initiated at the centre of the disc along the loading line, and the tensile stress decreased with time (Figure 7). The failure pattern at time 80  $\mu$ s closely resembles the experimental results of fine-grained marble obtained by Q.B. Zhang [4].

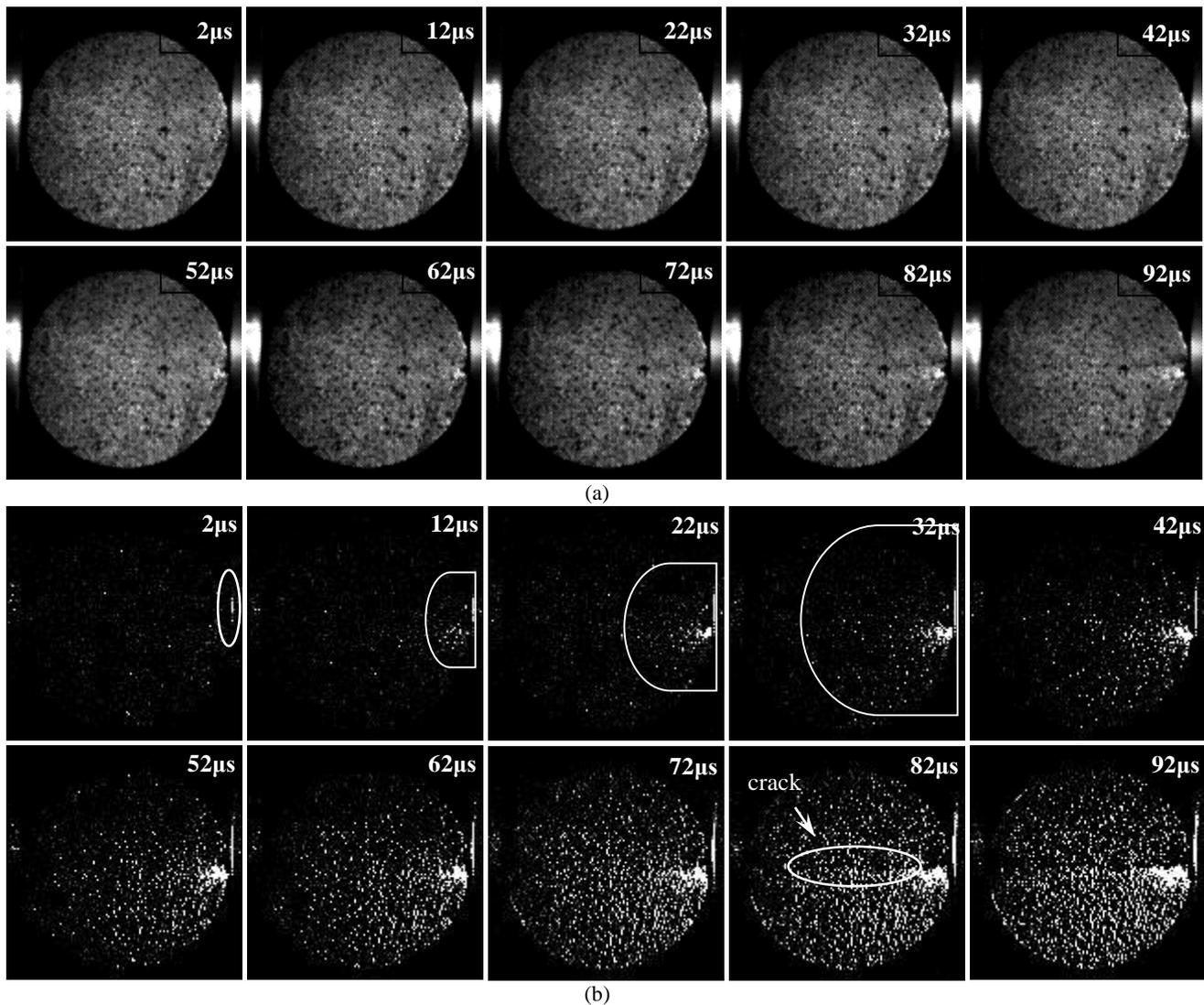


FIGURE 8 High speed images of a specimen under dynamic loading (a) Different stages; (b) Differential images

## 7 Conclusions

This study successfully employed a method combining a SHPB with an HS camera technique to explore the dynamic fracture behaviour of sandstone. In situ images of the surface of sandstone in Brazilian disk test were acquired during the dynamic loading step. The GLDD technology base on probability integration was used in conjunction with HS- photography to measure full-field deformation of specimens. The proposed system facilitates cost-effective, non-contact full-field deformation measurements of specimens in dynamic testing methods. Results demonstrate that dynamic mechanical properties can be well-determined, and that the image difference technology base on probability integration is a reliable, full-field deformation measurement method. This method is likely to gain

popularity with higher speed and higher resolution cameras as well as better computing methods, particularly as sub-pixel techniques become more readily and economically available.

## Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (NO. 51304007, 51104004, 51104068), the State Key Program of National Natural Science Foundation of China (U1361208), China Postdoctoral Science Foundation (NO. 2013M531495) and Scientific Research Fund for Young Teachers of Anhui University of Science and Technology (NO. 2012QNY39) and Scientific Research Fund for Doctor and Master of Anhui University of Science and Technology.

## References

- [1] Peters W H, Ranson W F 1982 Digital imaging techniques in experimental stress analysis *Optical Engineering* **21**(3) 427-31
- [2] Song Y I, Ma S P, Yang X B 2011 Experimental investigation on failure of rock by digital speckle correlation methods *Chinese Journal of Rock Mechanics and Engineering* **30**(1) 170-5 (In Chinese)

[3] Yin Z Q, Li X B, Yin T B 2012 Critical failure Characteristics of high stress rock inducer by impact disturbance under confining pressure unloading *Chinese Journal of Rock Mechanics and Engineering* 31(7) 1355-62 (In Chinese)

[4] Zhang Q B, Zhao J 2013 *International Journal of Rock Mechanics and Mining Sciences* 60(6) 423-39

[5] Siviour C R, Grantham S G, Williamson D M 2009 Novel measurements of material properties at high rates of strain using speckle metrology *The Imaging Science Journal* 57(6) 326-32

[6] Wu M, Wang H, Zhang Zh 2013 *Computer Modelling and New Technologies* 17(4) 229-35

[7] Bromiley P A, Thacker N A, Courtney P 2002 Non-parametric image subtraction using grey level scattergrams *Image and Vision Computing* 20(9-10) 609-17

[8] Li X B, Lok T S, Zhao J 2000 *International Journal of Rock Mechanics and Mining Science* 37(7) 1055-60

[9] Jin J F, Li X B, Zhong H B 2013 Study of dynamic mechanical characteristic of sandstone subjected to three-dimensional coupled static-cyclic impact loadings *Chinese Journal of Rock Mechanics and Engineering* 32(7) 1358-72 (In Chinese)

[10] Li X B, Gu D S, Lai H H 1993 On the reasonable loading stress waveforms determined by dynamic stress-strain curves of rock by SHPB *Explosion and Shock Waves* 13(2) 125-30 (In Chinese)

[11] Dai F, Xia K W, Tang L 2010 *International Journal of Rock Mechanics and Mining Science* 47 (3) 469-75

[12] Zhou Z L, Li X B, Yan X M 2009 Loading condition for specimen deformation at constant strain rate in SHPB test of rocks *Chinese Journal of Rock Mechanics and Engineering* 28(12) 2445-52 (In Chinese)

[13] Zhu W C, Tang C A 2006 *International Journal of Rock Mechanics and Mining Sciences* 43(2) 236-52

Authors	
	<p><b>Zhiqiang Yin, born in November, 1983, Huainan, Anhui, China</b></p> <p><b>Current position, grades:</b> Associate Professor of School of Mineral and Safety, Anhui University of Science and Technology, China.  <b>University studies:</b> D.E. in Mining Engineering at Central South University of Changsha in China.  <b>Scientific interest:</b> Rock Mechanics and Engineering.  <b>Publications:</b> More than 10 papers published in various journals.  <b>Experience:</b> Teaching experience of 10 years, 3 scientific research projects.</p>
	<p><b>Lei Wang, born in July, 1980, Huainan, Anhui, China</b></p> <p><b>Current position, grades:</b> Associate Professor of School of Mineral and Safety, Anhui University of Science and Technology, China.  <b>University studies:</b> D.E. in Safety technology and Engineering at Anhui University of Science and Technology of Huainan in China.  <b>Scientific interest:</b> Mine Dynamic Disaster  <b>Publications:</b> More than 20 papers published in various journals  <b>Experience:</b> Teaching experience of 10 years, 9 scientific research projects</p>
	<p><b>Haifeng Ma, born in August, 1984, Beijing, China</b></p> <p><b>Current position, grades:</b> Ph.D Candidate of Faculty of Resources and Safety Engineering, China University of Mining and Technology, China.  <b>University studies:</b> B.Sc. in Mining Engineering at Anhui University of Science and Technology of Huainan in China. M.Sc. at Anhui University of Science and Technology in China.  <b>Scientific interest:</b> Mine Dynamic Disaster.  <b>Publications:</b> More than 5 papers published in various journals  <b>Experience:</b> Teaching experience of 3 years, 2 scientific research projects</p>

# Research on illumination invariance colour index algorithm based on colour ratio

Fei Cai<sup>1, 2, 3</sup>, Wenjun Wang<sup>1, 2\*</sup>

<sup>1</sup>School of Computer Science and Technology, Tianjin University, Tianjin 300072, China

<sup>2</sup>Tianjin Key Laboratory of Cognitive Computing and Application, Tianjin 300072, China

<sup>3</sup>School of Civil Engineering, Shandong Jianzhu University, Jinan, 250101, China

Received 26 October 2013, www.tsi.lv

## Abstract

Statistics of colour value of each pixel in the image are output in traditional colour histograms. Therefore, though the two same images photographed in different illuminations are consistent in colour content, they have different colour distributions in the histograms. To solve the problem, this paper introduces an illumination invariance colour index algorithm based on colour ratio. According to the colour constancy theory, although colour values of its pixels will be changed once the image is subject to illumination, colour ratios remain unchanged. Colour ratio refers to the ratio between two contiguous pixels. As per colour ratios, colour ratio image may be obtained, which depicts obvious boundaries or margins of the image content so that we statistics of colour ratio histogram can be obtained as an index mechanism to remove illumination effect. Verified by lots of tests, this method can extract useful colour characteristics and remove illumination effect, so that it can be practically used in effective computer recognition of objects in traffic videos.

*Keywords:* colour index, image retrieval, colour constancy, illumination

## 1 Introduction

Images photographed by the traffic monitoring system are affected by illumination to a great extent; as a result, objects can't be positioned and recognized by many traditional methods accurately, causing the intelligent monitoring system less intelligent and limiting its popularity. Human assistance is still needed for recognizing images in monitoring videos. Therefore, how to achieve illumination invariance in colour recognition is a key to guaranteeing the robustness of video monitoring systems.

With the simplest, most direct and effective characteristic, colour has been widely used in computer vision applications. Colour histogram is a simple and effective expression of the colour characteristic. Histogram intersection algorithm raised by Swain et al [1] is rather effective in object recognition. However, colour histogram is easily subject to illumination despite of its robustness in image rotation and affine. To remove illumination effect on colours, RGB colour space may be transformed to RG chroma space, thus making a chroma histogram. Chroma histogram, however, still can't remove the effect of illumination colour variance on image colour. To solve the problem, Colour Constancy Colour Index (CCCI), a descriptor, which is robust to both illumination intensity variance and illumination colour variance, is raised [2-10]. This descriptor removes illumination effect using the derivative of colour

logarithmic space, which is extended by Gevers et al [11]. That further removes the effect of perspective and shade on colours. Since CCCI is based on derivative of colour, it relies on the margin information of images and is likely to be affected by image blurring.

This paper introduces an index algorithm based on illumination invariance. It mainly removes illumination effect, thus affecting the effectiveness of traditional index mechanism to a great extent. Therefore, illumination invariance plays a significant role in improving index technology. The algorithm raised in this paper is well applied in object recognition of traffic monitoring system, so as to assist the intelligent traffic video monitoring system, reduce human assistance and improve monitoring efficiency. Verified by a quantity of tests, this histogram technology based on colour index can effectively remove the effect of objective environmental factors on images, which is the core for solving problems in image recognition technology of intelligent traffic system. Traditional computer recognizes image through pixels of images, but a barrier for recognizing people and vehicles in images is posed by the illumination; colour index can remove the effect of environmental hue on the object recognition of computer.

## 2 Relevant Principle-Physical Characteristics of Colours

The colour of an object is determined by three factors:

\* *Corresponding author* e-mail: wenjwjtj@163.com

surface reflection of the object, illumination distribution and responsivity of colour channels. For a certain wave length  $\lambda$  and a point  $(x, y)$  on the surface, its illumination function is stated as follows:

$$I(x, y, \lambda) = E(\lambda) \cdot S(x, y, \lambda), \quad (1)$$

where in:  $E(\lambda)$  refers to illumination distribution function;  $S(x, y, \lambda)$  refers to surface reflection function of the point  $(x, y)$ .

Colour response of the point  $(x, y)$ , i.e. transformation into visual system through Channel  $k$ . It can be expressed as:

$$\phi_k = \int_{\lambda} I(x, y, \lambda) R_k(\lambda) d\lambda. \quad (2)$$

$R_k(\lambda)$  refers to the responsivity of Channel  $k$ . Illumination mentioned above is independent from surroundings of the object. However, illumination perceived by people is dependent on the surroundings. Both illumination energy distribution and colour channel responsivity are objective factors in physics, so the perceptual colour of object can be approximately expressed as illumination distribution of surface reflection function.

### 3 Illumination Invariance Algorithm Based on Colour Ratio

#### 3.1 IMPROVED STATISTICAL COLOUR HISTOGRAM ALGORITHM

For 24-bit Windows bitmap images of true colours, there is no such a title as histogram. Only statistics of each colour value for grayscale images can be performed and presented in form of histogram. That's why the grey value of a single under a single channel is extracted in the 2<sup>nd</sup> step of the algorithm. Both histogram and colour ratio histogram are concepts designed for images under a single channel. Grey values of grayscale image pixels are stored in a matrix. Therefore, it is easy to draw a colour histogram. Only two nested loop statements are needed to perform statistics of each colour value frequency. They are stored in an array or matrix and drawn using two-dimensional coordinates.

The improved statistical colour ratio histogram algorithm is stated as follows:

Step 1: read in key images from the image library.

Step 2: position sub-blocks in images and calculate sub-block histograms; i.e. users select sub-blocks containing query images.

Step 3: calculate sub-block  $(x, y)$  colour pair tables using "eight-direction adjacent technology" with the calculate colour pair () function.

Step 4: delete the colour pairs whose differences are smaller than a certain threshold, so as to eliminate colour noise.

Step 5: fill the colour pair tables obtained from the

sub-block calculation in the characteristic colour pair table of the image and rank them in descending order. Set a field value, and select the greatest colour pairs in the colour pair table as the representative features of the image.

Step 6: perform colour matching and read the  $N$  images to be compared. Calculate the colour histogram of the compared image, and search for the colour pair table for each sub-block in the target image. Do not use accurate matching. Therefore, colour pairs whose error is smaller than 2% also belong to the matching values.

Step 7: count the number of single matches. Calculate the colour pairs of a sub-block and its surrounding sub-blocks of the target image in order. Inquire the calculated colour pairs in the colour pair tables entered by users. If the difference is smaller than a threshold value, it matches and should be marked with the colour matching logo.

Step 8: if over 60% of colour characteristic colours are matched, then the image is retrieved.

Step 9: present search results.

#### 3.2 CALCULATION OF IMAGE COLOUR RATIO VALUE

Objective brightness of a point should be independent from its surroundings, but the actually perceived colour depends more on its relationship with adjacent points rather than its own amount of light reflection. Therefore, the colour perception amount is subject to the overall situation. That has to say, the adjacent spatial area of the object will be taken into account by human visual system when calculating its colour. Therefore, the invariant colour characteristics may be based on operations to adjacent points; that's, the relationship between a point and its adjacent points should be considered. The following common assumptions are made to image environment: the surface is not smooth, only diffuse reflection exists, and there is no shade on the surface. These assumptions may not always be reached, but it renders analysis to the colour ratio model operative. In computer calculations, these are some typical assumptions. [6] Set  $\alpha\beta\xi$  as the conversion factor and make the following assumptions: illumination function may be expressed using the following linear function:

$$E(\lambda) = \sum_{i=1}^N \alpha_i E_i(\lambda). \quad (3)$$

Surface reflection function may be expressed as follows:

$$S(\lambda) = \sum_{j=1}^M \beta_j S_j(\lambda). \quad (4)$$

Illumination may be described using an expression merely relying on conversion ratio, so may similar reflection functions. Equation (2) can be approximated as:

$$\phi_k = \sum_{\lambda} \sum_{i=1}^N \sum_{j=1}^M \alpha_i \beta_j E(\lambda) S(\lambda) R_k(\lambda) = \sum_{\lambda} \xi_{ijk}(\lambda) \cdot S(\lambda). \quad (5)$$

Despite of illumination changes, adjacent points will receive the same amount of light simultaneously.

For the two different points of  $(x_1, y_1), (x_2, y_2)$  on the surface, simple expressions for colour responsively at Channel  $k$  may be given separately:  $\phi_k^1 = k \xi^1 S^1$ ;  $\phi_k^2 = k \xi^2 S^2$ . The parameter  $\xi$  is used to capture changes in illumination and surface reflection function. Illuminations for the small adjacent areas are the same. At this time,  $\xi$  depends only on surface reflection, thereby substantially dependent on physical properties of the image surface. Based on the same assumptions,  $k$  is also same for the two points. Thus,  $k$  may be removed in obtaining the colour ratio.

These assumptions are relatively weak and are easy to be satisfied in practice. For small adjacent spatial areas, the fourth assumption is obviously correct. When  $N$  and  $M$  is relatively small, such as 3, most changes in ambient light or reflection coefficient can be calculated using the linear function. As mentioned above, even if the illumination is changed, colour responsivities of the adjacent points are the same, unless colour borders exist. Therefore, to obtain the unchanged colour characteristics, we only need to detect how the colour value of a point varies as that of its adjacent point varies. The ratio is obtained using the following formula:

$$\Phi_{x,y} = \frac{\varphi_{x,y} - \frac{1}{m} \sum_{i=1}^m (\varphi_{x,y} - \varphi_{x,y}^i)}{\varphi_{x,y}}, \quad (6)$$

where in:  $\varphi_{x,y}$  means the colour value of the point  $(x, y)$ ,  $0 \leq \phi \leq L$  and  $L$  is the colour grade,  $\varphi_{x,y}^i$  means the colour value 0 the  $i^{\text{th}}$  adjacent point near to the point  $(x, y)$ .

The colour ratio  $\Phi$  of the aforesaid result is a constant physical quantity of adjacent points on the surface. The ratio is invariable compared to the illumination changes. It is essentially a description of the spatial context information of a small area in the image. It is independent from external factors, such as external light. For the three colour channels, three colour ratio figures for the original image are drawn.

Algorithm is designed to obtain colour ratio of the image. First of all, the formula suitable for colour ratio calculation should be located.  $\Phi$  obtained from Equation (6) is the desired colour ratio. But obviously the equation is relatively complicated. To simplify the colour ratio calculation, the colour ratio model is further standardized.

In Equation (6), the colour model is formulated to

emphasize the relationship between its spatial adjacent points. The equation can be simplified as follows:

$$\Phi_{x,y} = \frac{\frac{1}{m} \sum_{i=1}^m \varphi_{x,y}}{\varphi_{x,y}}. \quad (7)$$

Thus, for a given colour channel, colour ratio of a designated point is the average of colour values of its adjacent points. The colour ratio model can be expressed using the following simpler one:

$$\Phi_{x,y} = \sum_{i=1}^m \psi^i(x, y), \quad (8)$$

where:  $\psi^i(x, y)$  is the colour value of the  $i^{\text{th}}$  adjacent point of the specified point. If  $\psi^i(x, y) = \varphi_{x,y}^i / m\varphi_{x,y}$ , Equations (6) and (7) will be used to obtain the colour ratio characteristics. If  $\psi^i(x, y) = \log(\varphi_{x,y}^i / \varphi_{x,y})$ , the method proposed first by Land and McCann and then used by Funt and Finlayson will be adopted to obtain the colour ratio characteristics. This is more in line with the multiplication of adjacent ratios rather than addition. A standard calculation formula for colour ratio model is presented in Equation (8), which is helpful to analyse colour ratio distribution. To facilitate the analysis, only direct colour value is considered; that's,

$$\psi^i(x, y) = \varphi_{x,y}^i / m\varphi_{x,y}.$$

First of all, the colour value for each pixel is extracted from the image and stored in a matrix. Then the colour ratio value is calculated as per Equation (6) and stored in another matrix with the same size. The ratio between the differences of each pixel and its neighbouring 8 points and the sum are calculated. During this process, some particular points need to be treated differently, namely, the pixels at the image edge, separately the matrix middle row, the 1<sup>st</sup> column or the maximum. These pixels can be divided into several categories; and the colour ratio calculation of each category of pixels needs to be treated differently instead of only substituting the pixel value into the formula. The colour ratio calculated in this way is closer to the actual value. Only in this way, can this indexing mechanism have obvious advantages in effect.

#### 4 Experiments

Two images photographed are loaded under different illuminations. Here are two images with the same scene separately photographed at daytime and night-time by the monitoring system, used for experimental purpose. This is one of the practical areas for this technology application.



(a) At daytime



(b) At nighttime

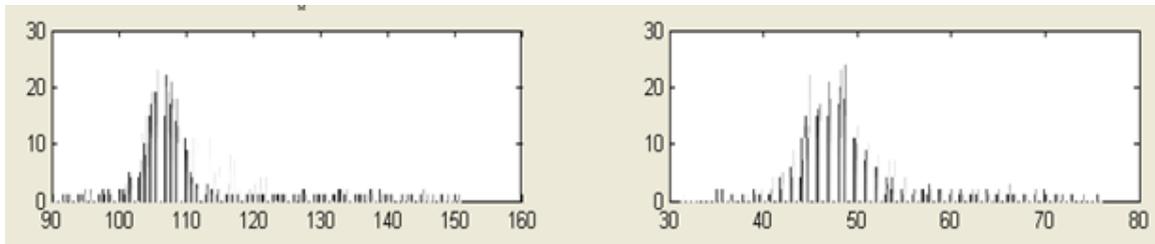
FIGURE 1 Two images with the same scene photographed at daytime and nighttime by the monitoring system

Obviously, the two images have the same scene. However, when performing matching retrieval using traditional colour histograms, they will be deemed to be completely different by computers since their colour histograms vary greatly.

Perform statistics of and draw colour histograms as shown in Figure 2. Through intuitive comparison of the

two colour histograms, it can be seen that the two images are greatly different. Calculate colour ratio values and obtain colour ratio images as shown in Figure 3.

Calculate and obtain colour ratio histograms as shown in Figure 4. It can be seen from the two images that they are quite similar.



(a) Daytime histogram

(b) Nighttime histogram

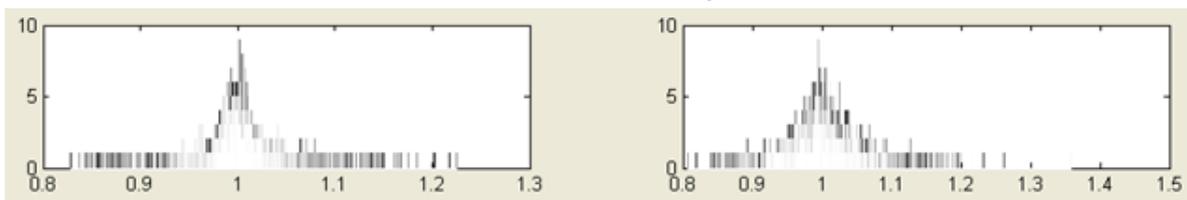
FIGURE 2 Colour histograms



(a) Daytime colour ratio image

(b) Night-time colour ratio image

FIGURE 3 Colour ratio images



(a) Daytime colour ratio histogram

(b) Night-time colour ratio histogram

FIGURE 4 Colour ratio histograms

Another method may be used to display the colour ratio histogram as shown in Figure 5. It can be seen that

the two colour ratio histograms are similar in a more intuitive and obvious manner.

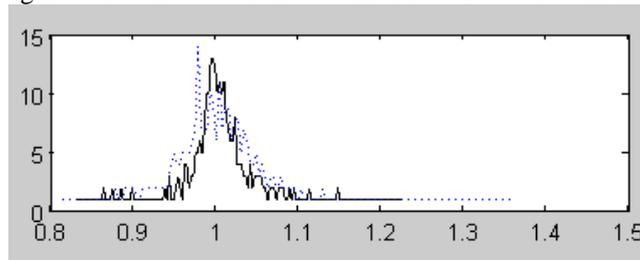


FIGURE 5 Colour ratio histogram

The distribution status of the colour ratio histogram of the image at daytime is drawn by solid black line while that at night-time is drawn by the blue dotted line. In this way, it can be seen that they are substantially the same in a more intuitive manner. When it is used as an index in a retrieval system, correct matching may be achieved with only one threshold value set.

## 5 Conclusions

How to extract effective colour characteristics from the image and remove the illumination effect are mainly studied in the paper. This algorithm can be used for effective computer recognition of target object in the image. This algorithm is relied by target recognition of modern intelligent traffic monitoring system.

Through analysis of current status of computer image processing, such as image retrieval, the application and development status of colour characteristics in image retrieval field, technologies for effective characteristic extract, application of colour indexing in image retrieval and illumination problems, as well as analysis and research on colour processing technologies, one illumination invariance indexing algorithm is proposed and described accordingly; besides, experimental results, specific requirements for the experimental platform,

presentation of experimental results, as well as histograms and colours are detailed.

The algorithm is studied and achieved for the final purpose of applying it in target recognition of traffic monitoring system. After the algorithm is achieved, we should move to target object recognition. One background image and one target image under the same ground are given, just like frame images extracted from the traffic monitoring system. Since images photographed within one day may be subject to illumination differently, the algorithm achieved in this paper instead of traditional colour information is to be used in target recognition.

## Acknowledgements

This work was supported by the National Science-technology Support Plan Project of China (No.2012BAK03B06, No.2013BAK02B06, No. 2012BAK03B00), the National Natural Science Foundation of China (No. 91224009), special funds of marine public welfare industry scientific research projects (201305033), Tianjin informatization project, MOHURD program of China (No.2011-K9-21), Housing and Urban-rural Development project of Shandong Province (Grant No.2011YK026), Soft science research project of Shandong province (No.2013RKB01284).

## References

- [1] Gijsenij A, Lu R, Gevers T 2012 *IEEE Transactions on Image Processing* **21**(2) 697-707
- [2] Gijsenij A, Gevers T, van de Weijer J 2011 *IEEE Transactions on Image Processing* **20**(9) 2475-89
- [3] Gijsenij A, Gevers T 2007 Colour constancy using natural image statistics *IEEE Conf Comput Vis Pattern Recognit* June 2007 1-8
- [4] Gijsenij A, Gevers T, van de Weijer J 2011 Computational colour constancy: Survey and experiments *Image Processing* **1**(99) 1-10
- [5] Gijsenij A, Gevers T, van de Weijer J 2010 *International Journal of Computer Vision* **86**(2-3) 127-39
- [6] Ebner M, 2009 Colour constancy based on local space average colour *Machine Vision and Applications* **20**(5) 283-301
- [7] Bianco S, Schettini R 2012 Colour constancy using faces. *Computer Vision and Pattern Recognition, 2012 IEEE Conference on Biometrics Compendium, IEEE* June 2012, 65-72
- [8] Golz J 2008 *Journal of Vision* **8**(13) 1-16
- [9] Chakrabarti A, Hirakawa K, Zickler T 2012 *IEEE Trans. Patt. Anal. and Mach. Intell.* **34**(8) 1509-19
- [10] Swain M, Ballard D 1991 Colour indexing *International Journal of Computer Vision* **7**(1) 11-32
- [11] van de Sande K, Gevers T, Snoek C 2008 *Pattern Analysis and Machine Intelligence, IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**(9) 1582-96

Authors	
	<p><b>Fei Cai, born in April, 1978, Tianjin, China</b></p> <p><b>Current position, grades:</b> Lecturer of School of civil engineering, Shandong Jianzhu University, China.  <b>University studies:</b> Ph.D. candidate in School of Computer Science and Technology, Tianjin University.  <b>Scientific interest:</b> Image processing, cognitive computing and application.  <b>Publications:</b> More than 20 papers published in various journals.  <b>Experience:</b> Teaching experience of 10 years, 3 scientific research projects.</p>
	<p><b>Wenjun Wang, born in May, 1970, Tianjin, China</b></p> <p><b>Current position, grades:</b> Professor of School of Computer Science and Technology, Tianjin University, China.  <b>University studies:</b> Ph.D. at Peking University in China.  <b>Scientific interest:</b> Intelligent traffic, cognitive computing and application.  <b>Publications:</b> More than 40 papers published in various journals.  <b>Experience:</b> Teaching experience of 9 years, 20 scientific research projects</p>

# Study on feasibility of CORS application in surface movement deformation monitoring in mining areas

**Debao Yuan\***, Xueqian Hong, Shiwei Yu, Liangjian Li, Yanbao Zhao

*College of Geoscience and Surveying Engineering, CUMT, Beijing, 100083, China*

*Received 1 April 2014, www.tsi.lv*

## Abstract

CORS has been widely established in China and abroad, and can be used in geodetic measuring, coordinate system retaining and surveying and mapping of city. However, some issues (stability of base stations, influence of survey precision on horizontal and vertical extension, et al) are still unsolved in Surface Movement Deformation Monitoring in Mining Areas. Based on the CORS, a method and flow for data processing and stability analysis of deformation monitoring network was proposed. Using rank defect free network adjustment, a robust estimation with minimum of first order norm of displacement component was offered to determine displacement of relative stability. The strategy can resolve the confirmation of robust iterative weights and the effects of different reference models. Displacement significance was tested with the normality method. Finally, according to this method, datum of GPS deformation monitoring in Mining Areas in Datong were calculated and preliminarily proved that this method was feasible and effective, providing a new monitoring methodologies terrain monitoring in mining areas.

*Keywords:* CORS (Continuous Operation Reference System), deformation monitoring, rank defect network adjustment, robust estimation, hypothesis test

## 1 Introduction

Coal is the major energy source for China. As the main province of coal resources reserve and production, Shanxi province owns a long history on it. Such a large-scale mining, for one hand, can bring huge economic benefits and create great wealth for the country. However, on the other hand, large-scale underground mining will cause ground deformation and ground settlement, which will affect people's life in mining areas, and produce a serious threat to the safety of people's lives and property. So effective surface movement deformation monitoring is particularly necessary [1].

In a long time, the traditional methods and techniques for mining surface movement monitoring remained unchanged. The traditional methods, such as triangulation measurement and levelling, are still widely used. The weaknesses are also obvious. For example, 1) needs more time and costs; 2) greatly restricted by topography, climate and other external factors; 3) data delay. It is not the real-time response to the information, less efficient [2].

The emergence of GPS solves the defects of traditional measuring method. It is less affected by visibility; topography, climate and other external constraints, and its precision completely meet the requirements. However, there are still some problems with GPS observations: 1) for static GPS observation, there must be at least two known control points within the scope of survey area. Without known monitoring area, GPS survey is obviously unable to meet the requirements

of deformation monitoring; 2) if the baseline length of static GPS observations is too long; the precision will be greatly reduced. For a certain working face, surface monitoring of GPS can meet the requirement, but to a wider regional monitoring, the GPS measurement becomes stretched [3, 8].

Continuously Operating Reference Service (CORS) can be defined as one or several fixed, continuously operating GPS reference station, using modern computer data communication network of LANWAN technology and the internet. CORS can automatically provide different types of GPS observations (such as carrier phase and pseudo-range), corrections, and the state Information of GPS and other relevant Service Projects of GPS to users with different needs and different levels in real time [4].

CORS, which is a revolutionary change of GPS technology, is the orientation of the development of GPS in the future. It not only has the advantages of GPS static measurement, but also perfectly resolved the lack of GPS [5]. First of all, it does not need the scene of known points, fixed reference stations which are more than dozens of kilometres or even hundreds of kilometres far away. This has greatly relieved constraints when designing and make arrangements of observation points. It can better reflect the ground deformation situation. Secondly, the length of baseline is not need to be considered as long as observation points are taken within CORS covered region at any time. In general, CORS reference stations are more than dozens of kilometres or even hundreds of kilometres, which can fully cover the

\*Corresponding author e-mail: Yuandb@cumtb.edu.cn

whole range of mining areas [6]. So, no matter where the monitoring of both working face of surface monitoring and a wider regional monitoring can be taken under CORS. Currently in China, CORS has been applied in domestic construction, urban planning, land surveying and mapping, and many other fields. And it has already achieved fruitful results. However, CORS application in the field of mining surface movement monitoring is still a blank. Therefore, CORS application in this field has a high research value and practical significance [7]. Based on the Project "Research on surface strata movement law of Tashan mine 8,103 and 8,104 working face surface of Datong Mine area", CORS application in the field of mining surface movement monitoring is studied. There are 4 base stations constructed by 115 Exploration Institute of Shanxi Coal Geology, which formed a CORS System. The distance between the 4 CORS base stations is 40 ~ 100 km, entirely covering the Tashan mine. Studied results indicate that the monitoring accuracy of CORS is completely meeting the monitoring requirements and bringing practical guiding significance for the future of mining surface movement monitoring [9].

## 2 Establish monitoring network

Due to mining working face covered by the fourth loess. The terrain is so complicated, and topography are characterized in huangtuliang, plateau and hilly. Based on Tashan mine, when the arrangement of surface movement observation station, we should arrange observation line according to the terrain and landform flexible. Trend line of observation arrange in line with the central line of the ground surface over the 8104 working face. Tendency direction of observation line along the tendency direction of 8103, 8104 working face was layout as a broken line on the plain ground. The details of the arrangement are as follows:

There are two observations line in this program. One of them is a trend line expressed as Z. The length of Z trend observation line is 2,175m, separation distance of the observation point is 25m, the distance between observation point and control point is 50 m respectively and the number of observation station is totally 89, including 3 control points. The tendency line Q represents the observation line: The length of Q observation is

1,356m, separation distance of the observation point is 25m, the distance between observation point and control point is 50m respectively and the number of observation station is totally 59, including 6 control points. The length of the trend observation lines and the tendency observation line is totally 5,041m and a total of 143 points. The distance between observation point and control point is 50m respectively and there are a total of 9 control points [15].

Observation point layout schematic diagram is shown in Figure 1.

## 3 Analytical methods and processes based on Datong CROS land subsidence monitoring network

For high precision GPS monitoring network, in order to obtain better results of the analysis, in addition to the reasonable and correct calculation strategy of the baseline, the choice of benchmark is critical. These benchmarks include the position reference, the scale reference, orientation and time evolution and so on [9]. Stability deformation analysis should be based on a unified, appropriate benchmarks to correctly distinguish between variables and system errors. Therefore, the observation data of GPS monitoring network can be treated scientifically, the study and evaluation method of analysing the stability of data, which is not only the need of theoretical research, but also the practical application.

Tashan surface mine subsidence network stability analysis methods and procedures are as follows: 1) with high precision processing software, choose the right baseline processing strategy to solve the baseline; 2) take the centre of gravity of the monitoring points as a benchmark for rank-defect free net adjustment, get the coordinates of the monitoring points every period and the factor matrix; 3) according to the displacement of two phase calculations and the factor matrix, the displacements of a norm as the minimum robust estimation method to determine the relative stability of the displacement, to obtain the optimum stability of the monitoring array and displacement; 4) construct statistics, then conduct significance test to monitoring displacement, determine the possible instability point; 5) according to each phase of testing the stability of the datum, the use of quasi-stable adjustment calculate the amount of displacement [16, 17].

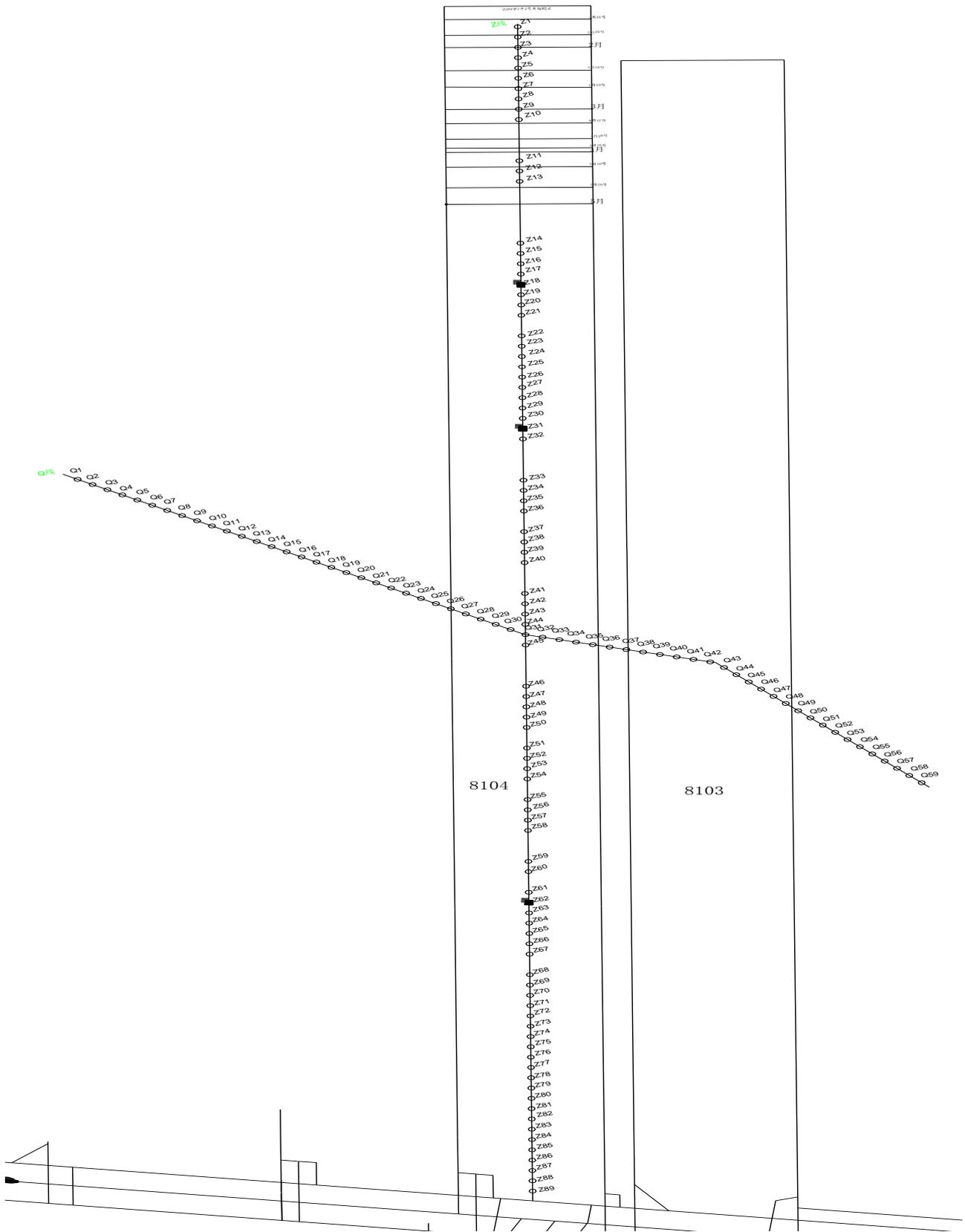


FIGURE 1 Observation station layout

**4 Model construction and Baseline processing**

GAMIT software which is developed by the Massachusetts Institute of Technology and IGS precise ephemeris are used in baseline solution. In baseline processing should consider the earth tide, tidal and polar motion correction; Satellite high Angle is 15°; Combine dual-frequency ionospheres LC as the basic observation; Use Saastamoinen model and parameter estimation to decrease troposphere error, and estimate a gradient troposphere per hour, deal with the daily time relaxation solution as the basic results every day. Baseline accuracy showed that daily coordinates solver accuracy in the north-south direction is 1~3mm, 2~4mm in the east-west direction; elevation accuracy is 8~10mm [10-12].

When processing baseline, double differential observations model is used. Get information of the baseline vector and the variance-covariance matrix with the method of the global combined net. If there are two observation station T1 and T2, the synchronous satellites SJ and SK where SJ is the reference satellite, then the double differential observation function can be expressed as:

$$\nabla\Delta\Phi^k(t) = -\frac{1}{\lambda}[\nabla l_2^k(t) \quad \nabla m_2^k(t) \quad \nabla n_2^k(t)] \begin{bmatrix} \delta X_2 \\ \delta Y_2 \\ \delta Z_2 \end{bmatrix} + \nabla\Delta N^k + \frac{1}{\lambda}[\rho_{20}^k(t) - \rho_1^k(t) - \rho_{20}^j(t) + \rho_1^j(t)] \quad (1)$$

where  $\nabla\Delta N^k$  can be written as:

$$\nabla\Delta N^k = \Delta N^k - \Delta N^j \quad (2)$$

And the vector  $[\nabla l_2^k(t) \quad \nabla m_2^k(t) \quad \nabla n_2^k(t)]$  can be expressed as:

$$\begin{bmatrix} \nabla l_2^k(t) \\ \nabla m_2^k(t) \\ \nabla n_2^k(t) \end{bmatrix} = \begin{bmatrix} l_2^k(t) - l_2^j(t) \\ m_2^k(t) - m_2^j(t) \\ n_2^k(t) - n_2^j(t) \end{bmatrix} \quad (3)$$

If  $\nabla\Delta l^k(t)$  can be expressed as:

$$\nabla\Delta l^k(t) = \nabla\Delta\phi^k(t) - \frac{1}{\lambda}[\rho_{20}^k(t) - \rho_1^k(t) - \rho_{20}^j(t) + \rho_1^j(t)] \quad (4)$$

Then the error equation can be express as:

$$v^k(t) = \frac{1}{\lambda}[\nabla l_2^k(t) \quad \nabla m_2^k(t) \quad \nabla n_2^k(t)] \begin{bmatrix} \delta X_2 \\ \delta Y_2 \\ \delta Z_2 \end{bmatrix} + \nabla\Delta N^k - \nabla\Delta l^k(t) \quad (5)$$

If the number of the synchronous satellite is  $n^J$ , the number of the epoch is  $n_t$ . Then write out  $(n^J - 1) \times n_t$  error equations. The relative error equation can be expressed as:

$$V = (A \quad B) \begin{bmatrix} \delta X_2 \\ \nabla\Delta N \end{bmatrix} + L \quad (6)$$

Therefore, the corresponding normal equation and the solution can be written as:

$$N \cdot \Delta Y + U = 0, \quad (7)$$

$$\Delta Y = -N^{-1}U, \quad (8)$$

where  $\Delta Y$ ,  $N$  and  $U$  can be expressed as:

$$\Delta Y = (\delta X_2 \quad \nabla\Delta N)^T, \quad (9)$$

$$N = (A \quad B)^T P (A \quad B), \quad (10)$$

$$U = (A \quad B)^T P L, \quad (11)$$

where  $P$  is the weight matrix of the double differential observation.

Finally when the number of the satellite of the two observation stations is  $n^J$ , and the number of the epoch is  $n_t$ , the corresponding weight matrix can be written as:

$$P_{\nabla\Delta\phi}(T) = \frac{1}{2\sigma^2} \frac{1}{n^J} \begin{bmatrix} n^J - 1 & -1 & \dots & -1 \\ -1 & n^J - 1 & \dots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \dots & n^J - 1 \end{bmatrix} \quad (12)$$

**5 Instance study**

According to the aforementioned method and steps, carry out rank defect free network adjustment on the two monitoring networks. After six iterations, to get the most reasonable displacement of monitoring point and its stability weight matrix [13, 14]. In the N, E and U direction, stable iteration weight matrix is a  $27 \times 27$  order diagonal matrix, its final iteration results are as follows:

$$\omega = \text{diag}(0.5010 \quad 0.6024 \quad 0.0655 \quad 0.3818 \quad 0.0747 \quad 0.0477 \quad 0.4587 \quad 1.0000 \quad 0.2693 \quad 0.7692 \quad 0.9231 \quad 0.4975 \quad 0.1342 \quad 0.4219 \quad 0.0478 \quad 0.3425 \quad 0.2710 \quad 0.1720 \quad 0.2688 \quad 0.9929 \quad 0.1958 \quad 0.4651 \quad 0.5558 \quad 0.1575 \quad 0.2747 \quad 1.0000 \quad 0.7463)_{27 \times 27}$$

Considering the needs of CORS monitoring data analysis, four points Z78, Z79, Z80, Z81 which are good for the levelling surveying were surveyed under third order levelling while CORS monitoring conducted over the same stage. After error distribution, the maximum levelling closure error W3 of the 6 stage is 5.7mm, the minimum W1 is 3.5mm. Both of them were less than the tolerance requirements of the closure error of third order levelling  $\pm 4\sqrt{n} = 13.9\text{mm}$ . It is shown that the results of levelling measurement are qualified and it can be used to compare the observation data of CORS.

Comparison curve diagrams of each stage are shown in Figures 2-5.

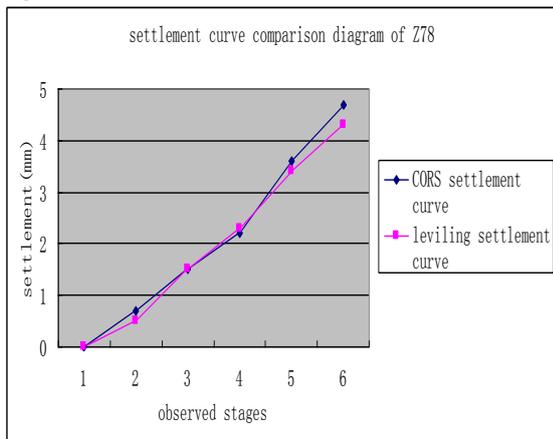


FIGURE 2 Settlement curve comparison diagram of Z78

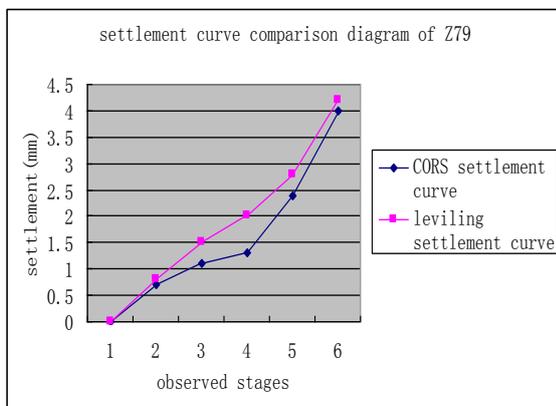


FIGURE 3 Settlement curve comparison diagram of Z79

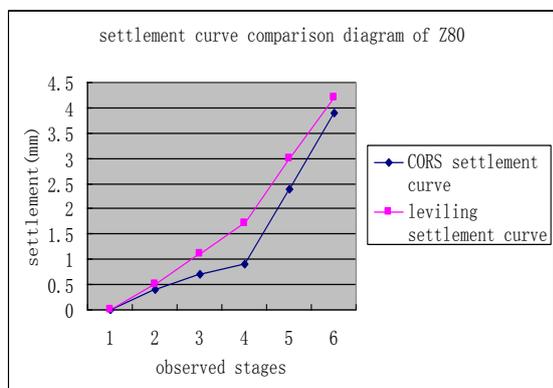


FIGURE 4 Settlement curve comparison diagram of Z80

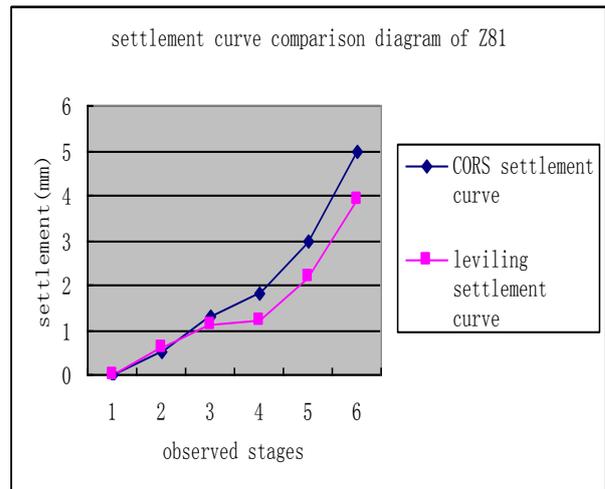


FIGURE 5 Settlement curve comparison diagram of Z81

Comparison of settlement curves between static GPS measurement under CORS and third order levelling surveying showed that the maximum difference between them is 1.3mm, the minimum is 0mm. From the curves, it can be seen that the two settlement curve agree with each other generally.

TABLE 1 Height Difference table of some observed points both by CORS and levelling

Height Difference of stage 1/m			
point	CORS	levelling	CORS-levelling
Z78	6.8069	6.81	-0.0031
Z79	3.7111	3.7105	0.0006
Z80	1.0352	1.041	-0.0058
Z81	-1.1805	-1.1915	0.011
Height Difference of stage 2 /m			
point	CORS	levelling	CORS-Levelling
Z78	6.8069	6.811	-0.0041
Z79	3.7113	3.7105	0.0006
Z80	1.0351	1.039	-0.0039
Z81	-1.1807	-1.19	0.0093
Height Difference of stage 3/m			
point	CORS	levelling	CORS-levelling
Z78	6.8069	6.8135	-0.0066
Z79	3.7114	3.712	-0.0006
Z80	1.0346	1.042	-0.0074
Z81	-1.1807	-1.192	0.0113
Height Difference of stage 4/m			
point	CORS	levelling	CORS-levelling
Z78	6.8071	6.8125	-0.0054
Z79	3.7114	3.7115	-0.0001
Z80	1.0343	1.04	-0.0057
Z81	-1.1809	-1.191	0.0101
Height Difference of stage 5/m			
point	CORS	levelling	CORS-levelling
Z78	6.807	6.8115	-0.0045
Z79	3.7111	3.711	0.0001
Z80	1.0345	1.0425	-0.008
Z81	-1.1811	-1.1915	0.0104
Height Difference of stage 6/m			
point	CORS	levelling	CORS-levelling
Z78	6.8073	6.812	-0.0047
Z79	3.7112	3.7115	-0.0003
Z80	1.034	1.0415	-0.0075
Z81	-1.1802	-1.1905	0.0103

The above tables are the differences of the height of the observation station using CORS measurement method compared with the levelling surveying at the same time. It is not difficult to see from the table that the maximum height error of the two methods is -11.3mm, the minimum is 0.1mm. And height error of the most of the points in all 6 observation stages is within  $\pm 10$ mm, height errors of some points in several stages are less than 1mm.

Height error diagram between CORS and levelling is shown in Figure 6 below.

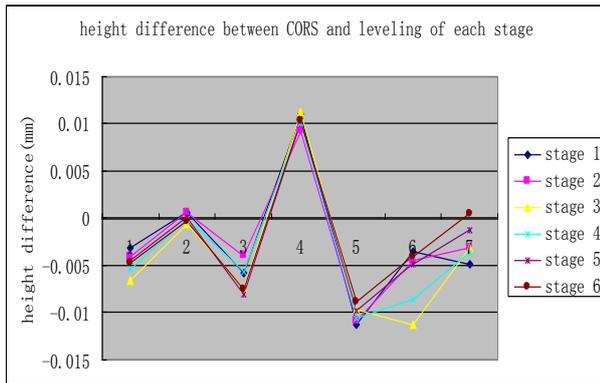


FIGURE 6 Height difference between CORS and levelling of each stage

From the above comparison and analysis, it can be seen that the height observation by CORS and by levelling surveying method is more or less the same, and the trends of changes are also corresponding. Therefore, it can be thought that statistic GPS height observation under CORS can substitute the levelling method. In addition, through the internal fitting accuracy, it can be seen that the accuracy of X, Y is better than H. So it is believed that the accuracy of CORS observations in X, Y,

H direction can meet the requirements of ground deformation completely. Finally, we can conclude that CORS can fully meet the accuracy requirements, and because of its advantages of the high accuracy and there is no need to know the known control points, CORS monitoring can replace conventional methods of measuring ground deformation monitoring [18].

## 6 Conclusions

As mining subsidence problem has increasingly become a popular concern, the ground movement and deformation monitoring has been attracted more and more attention. Strata and surface movement caused by mining process is very complicated, it is caused from geological conditions, hydrology, mining field, topographical conditions and other factors. Therefore, the most effective, reliable and directive method is field observation for conducting mining surface movement deformation monitoring.

The emergence of CORS is a combined product of computer, data communication, Internet technology, and it is the trend of the development of GPS in the future. CORS has been applied in domestic construction, urban planning, land surveying and mapping, and many other fields, and it has already achieved fruitful results. Meanwhile, advantages of high precision and long baseline make CORS for mining deformation monitoring.

Through actual data, it can be proved that the precision of CORS observation can fully meet the requirements of the mining surface movement deformation monitoring. It is a fast and efficient method for ground movement and deformation monitoring of mining area.

## References

- [1] Murr L E, Esquivel E V 2004 *Journal of materials science* **39**(4) 1153-68
- [2] Lovse J W, Teskey W F, Lachapelle G, Cannon M E 1995 *Journal of surveying engineering* **121**(1) 35-40
- [3] Yue J P, Fang L, Li N 2007 Research Advances of Theory and Technology in Deformation Monitoring *Bulletin of Surveying and Mapping* **7** 14
- [4] Snay R A, Soler T 2008 *Journal of Surveying Engineering* **134**(4) 95-104
- [5] Tang W M, Lou Y D, Liu H, Chen R G, Yang Q 2006 Research on positioning precision testing methods in GPS continuously operating reference station system *Journal China Institute of Communications* **27**(8) 73 (in Chinese)
- [6] Kaplan E D, Hegarty C J 2005 *Understanding GPS: principles and applications* London: Artech House 379-454
- [7] Hu G R, Khoo H S, Goh P C, Law C L 2003 *Journal of Geodesy* **77**(5-6) 292-302
- [8] Tomkiewicz S M, Fuller M R, Kie J G, Bates K K 2010 *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**(1550) 2163-76
- [9] Ai G X, Shi H L, Wu H T, Yan Y H, Bian Y J, Hu Y H, Li Z G, Guo J, Cai X D 2008 A positioning system based on communication satellites and the Chinese Area Positioning System (CAPS) *Chinese Journal of Astronomy and Astrophysics* **8**(6) 611 (in Chinese)
- [10] Wendt J, Dietrich R J 2003 *Journal of Geodynamics* **35** 235-46
- [11] King R W, Bock Y 1999 *Documentation for the GAMIT GPS analysis software* <http://www-gpsg.mit.edu/~simon/gtk/GAMIT.pdf> - 2000/ 15 April 2014
- [12] Dongchen E, Biwei Z, Weiping J, Shengkai Z 2005 High-precision GPS data processing by GAMIT/GLOBK *Chinese Journal of Polar Research* **17** (3) 173-82 (in Chinese)
- [13] Cardelli L, Wegner P 1985 *Computing Surveys* **17**(4) 471-523
- [14] Meyers S M 1992 *Effective C++: 50 specific ways to improve your programs and designs* Boston: Addison-Wesley MA 149-92
- [15] Scherzinger B M 2000 Precise robust positioning with inertial/GPS RTK *Proc ION GPS-2000 Salt Lake City* **9** 19-22
- [16] Hilla S, Cline M 2004 *GPS Solutions* **7**(4) 253-67
- [17] Qi F, Liu H 2003 The analysis of applying GPRS in CORS *Gnss World of China* **1** 011 (in Chinese)
- [18] Yuan Z, Zhao L 2014 SINS/GPS Carrier Phase Rate Integrated Navigation System based on Square-root CKF *International Journal of Online Engineering* **10**(3) 29-32

Authors	
	<p><b>Debao Yuan, born in April, 1976, Beijing, China</b></p> <p><b>Current position, grades:</b> lecture of College of Geoscience and Surveying Engineering, CUMT, Beijing, China.  <b>University studies:</b> doctor degree in Surveying Engineering.  <b>Scientific interest:</b> GPS navigation.  <b>Publications:</b> more than 20 papers published in various journals.  <b>Experience:</b> teaching experience of 4 years, 6 scientific research projects.</p>
	<p><b>Xueqian Hong, born in December, 1988, Beijing City, China</b></p> <p><b>Current position, grades:</b> Master of College of Geoscience and Surveying Engineering, CUMT, Beijing, China.  <b>University studies:</b> B.E at Chinese University of Mining and technology 2012.  <b>Scientific interest:</b> GPS navigation.</p>
	<p><b>Shiwei Yu, born in January, 1991, Beijing, China</b></p> <p><b>Current position, grades:</b> the Master of College of Geoscience and Surveying Engineering, CUMT, Beijing, China.  <b>University studies:</b> B.E at Chinese University of Mining and technology 2013.  <b>Scientific interest:</b> GPS navigation.</p>
	<p><b>Liangjian Li, born in December, 1987, Beijing City, China</b></p> <p><b>Current position, grades:</b> Master of College of Geoscience and Surveying Engineering, CUMT, Beijing, China.  <b>University studies:</b> a bachelor degree in Surveying Engineering 2013.  <b>Scientific interest:</b> GPS navigation.</p>
	<p><b>Yanbo Zhao, born in May, 1989, Beijing, China</b></p> <p><b>Current position, grades:</b> Master of College of Geoscience and Surveying Engineering, CUMT, Beijing, China.  <b>University studies:</b> bachelor degree in Surveying Engineering Shijiazhuang University in 2013.  <b>Scientific interest:</b> RS image processing and GPS navigation.</p>

# Study on improved elasto-plastic model for unsaturated soils based on Barcelona Basic Model

Jianjun Dong\*

Department of Civil Engineering College of Mining Engineering Liaoning Technical University, Huludao, China 125105

Received 10 April 2014, www.tsi.lv

## Abstract

The strength and deformation behaviours of unsaturated soil can be approximately described by elasto-plastic constitutive model that was proved by abundance academic and test researches. The Barcelona elastic-plastic model is an excellent model that can simulate the strength and deformation of unsaturated soil. But their calculated result of shear strength is low. So an improved BBM model is settled by using drop-shaped shear yield surface and hardening theory of dual stress. The results show that the improved model can more accurately predict the strength and deformation behaviours of unsaturated soil under suction-controlled triaxial compression stress states.

*Keywords:* shear yield surface, drop-shaped, hardening theory, dual stress, unsaturated soil, elasto-plastic

## 1 Introduction

The constitutive relation is very complex for unsaturated soil, a three-phase material that consists of solids, water, and air, and very little, if any, promising models have been reported. Some representative models on constitutive relation include nonlinear elastic model, elasto-plastic model and structural model [1-3]. Of those models, Fredlund et al. [4] proposed non-linear elastic model for unsaturated soil; Alonso et al. [5-9] proposed the elasto-plastic model, namely, the Barcelona Basic Model (BBM). These proposed models reflect the strength and deformation properties of unsaturated soil in varying degrees. Theoretical and experimental results also show that the strength and deformation characteristics of unsaturated soil can be approximately described by elasto-plastic model. Therefore, BBM can be utilized for applications in engineering practice.

However, experimental studies showed that the generally BBM underestimates the shear strength of unsaturated soils in [6, 10], attributed to the elliptical yield surface of modified Cam-clay model. To address this issue, the drop-shaped yield surface was introduced as the shear yield surface, and the associated flow rule and dual stress hardening theory are used in the framework of the original model. Comparative studies show that the improved model can predict the properties of strength and deformation for unsaturated soil more accurately, under the tri-axial compression test of controlled suction.

## 2 Barcelona basic model

Alonso et al. [5-9] proposed the Barcelona Basic Model

based on the classic modified Cam-clay model. The suction component was introduced and the unified elasto-plastic model was proposed in the generalized stress space ( $q$ ,  $p$ ,  $s$ ) according to the concept of saturated critical state for non-expansive soil and weak expansive soil, in which  $q$ ,  $p$ ,  $s$ , represent the shear stress, the average net stress  $p$  and the suction  $s$ , respectively. When the soil is saturated, the model degenerates to the modified Cam-clay model.

### 2.1 STRESS YIELD SURFACE

According to the following Stress variables, two types of yield functions are defined.

$$p = (\sigma_1 + 2\sigma_3)/3 - u_a, \quad (1)$$

$$q = \sigma_1 - \sigma_3, \quad (2)$$

$$s = u_a - u_w. \quad (3)$$

In Equation (3),  $u_a$  and  $u_w$  are the pore air pressure and pore water pressure.

The first type of yield function is expressed as:

$$f_1(p, q, s, p_0) = q^2 - M^2(p + p_s)(p_0 - p) = 0, \quad (4)$$

$$p_s = ks, \quad (5)$$

$$\frac{p_0}{p^c} = \left( \frac{p_0^*}{p^c} \right)^{[\lambda(0)-k]/[\lambda(s)-k]}, \quad (6)$$

\* Corresponding author e-mail: dongjianjun@gmail.com

$$\lambda(s) = \lambda(0)[(1-r)\exp(-\beta s) + r]. \tag{7}$$

$$d\varepsilon_{vs}^p = \mu_2. \tag{18}$$

Equation (6) is the LC (after loading-collapse) yield surface equation,  $p_0$  and  $p_0^*$  are pre-consolidation stresses, corresponding to suction and saturated conditions.

The second type of yield function is defined as

$$f_2(s, s_0) = s - s_0 = 0. \tag{8}$$

Equation (8) represents the SI yield curve (after suction increase), where  $s_0$  is the maximum suction in the history of soil.

2.2 FLOW RULE

Corresponding to the yield surfaces of  $f_1$  and  $f_2$ , the plastic potential function are assumed as

$$g_1 = f_1, \tag{9}$$

$$g_2 = p. \tag{10}$$

The  $f_1$  and  $f_2$  are used on the associated and non-associated flow rule. The plastic strain increment is  $(d\varepsilon_{vp}^p, d\varepsilon_q^p)$  and the following formulas are obtained

$$d\varepsilon_{vp}^p = \mu_1 n_p, \tag{11}$$

$$d\varepsilon_q^p = \mu_1 n_q, \tag{12}$$

$$n_p = 1, \tag{13}$$

$$n_q = [2q\alpha / M^2(2p + p_s - p_0)]. \tag{14}$$

In Equation (14),  $\alpha$  is a constant that can be obtained by calculating the plastic strain increment in the direction of no lateral deformation, as

$$\frac{d\varepsilon_q^p}{d\varepsilon_{vp}^p} = (2/3)\{1/[1-k/\lambda(0)]\}. \tag{15}$$

In order to simplify the formula,  $d\varepsilon_q^e$  has been assumed to be 0. The stress state to meet  $K_0$  conditions can be determined using Equation (16).

$$(q/p + p_s) = 3(1 - K_0)/(1 + 2K_0). \tag{16}$$

Thereafter,  $\alpha$  can be expressed as

$$\alpha = \frac{M(M-9)(M-3)}{9(6-M)} \{1/[1-k/\lambda(0)]\}. \tag{17}$$

To estimate yield surface of  $f_2$ , the associated plastic strain vector is given as  $(d\varepsilon_{vs}^p, 0)$ , where

$\mu_1$  and  $\mu_2$  can be obtained through the plastic consistency condition, namely

$$\mu_1 = \frac{\frac{\partial f_1}{\partial p} dp + \frac{\partial f_1}{\partial q} dq + \frac{\partial f_1}{\partial s} ds}{\frac{\partial f_1}{\partial p_0^*} \frac{\partial p_0^*}{\partial \varepsilon_v^p}}, \tag{19}$$

$$\mu_2 = \left( \frac{\partial f_2}{\partial s} ds \right) / \left( \frac{\partial f_2}{\partial s_0} \frac{\partial s_0}{\partial \varepsilon_v^p} \right). \tag{20}$$

2.3 HARDENING LAW

The development of yield surface is controlled by hardening parameters  $p_0^*$  and  $s_0$ , which further depend on the total plastic volume strain increment  $d\varepsilon_v^p$ .

$$\frac{dp_0^*}{p_0^*} = \frac{v}{\lambda(0) - \kappa} d\varepsilon_v^p, \tag{21}$$

$$\frac{ds_0}{(s_0 + p_{at})} = \frac{v}{\lambda_s - \kappa_s} d\varepsilon_v^p. \tag{22}$$

2.4 ELASTIC STRAIN

The volumetric elastic strain and shear elastic strain are constituted by the following formula.

$$d\varepsilon_v^e = \frac{k}{v} \frac{dp}{p} + \frac{ks}{v} \frac{ds}{(s + p_{at})}, \tag{23}$$

$$d\varepsilon_q^e = (1/3G)dq. \tag{24}$$

3 Improved model

3.1 SHEAR YIELD SURFACE

BBM with elliptical shear yield surface underestimated the strength of unsaturated soils. The shear yield surface of drop-shaped for saturated soil [1, 11, 12] can be extended to the elasto-plastic model for unsaturated soils Figure 1, with the associated flow rule to improve the BBM.

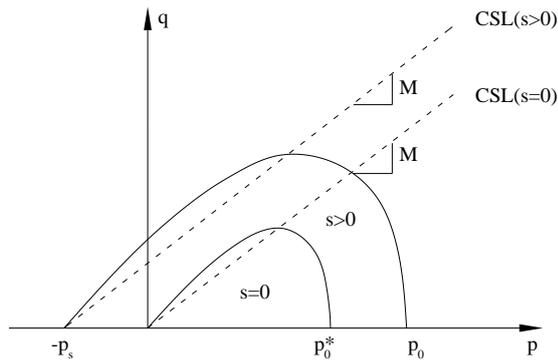


FIGURE 1 Yield surfaces of  $p$ - $q$  space

Equation of drop-shaped yield surface is

$$f_1(p, q, s, p_0) = (p_0 + p_s)q^2 + 2M^2 \cdot (p + p_s)^2 \cdot (p + p_0 + 2p_s) \cdot (p - p_0) = 0 \quad (25)$$

According to

$$\frac{\partial f_1}{\partial p} = \frac{16 \left[ (p + p_s)^2 - \frac{1}{2}(p_0 + p_s)^2 \right] (p + p_s)}{(p_0 + p_s)^4}, \quad (26)$$

$$\frac{\partial f_1}{\partial q} = \frac{4q}{M^2 (p_0 + p_s)^2}. \quad (27)$$

BBM uses non-associated flow rule, while the improved model employs associated flow rule to simplify the calculation. The shear shrinkage rate is

$$\frac{\Delta \varepsilon_q^p}{\Delta \varepsilon_{vp}^p} = \frac{\partial f / \partial q}{\partial f / \partial p} = \frac{q(p_0 + p_s)^2}{4M^2 \left[ (p + p_s)^2 - \frac{1}{2}(p_0 + p_s)^2 \right] (p + p_s)} \quad (28)$$

### 3.2 HARDENING THEORY

With drop-shaped shear yield surface, the predicted shear strength has been significantly improved, compared with the original model, but the results tend to be hard, and the deformation is small. To overcome the deficiency, the dual stress hardening theory is used instead of the iso-surface hardening theory for the original model [1, 13, 14]. Initially, the dual stress theory appeared in the bounding surface model, and its method was to find the dual stress  $q_{CSL}$  that corresponded to the existing stress  $q$  on the boundary surface. Here,  $\delta$  is the distance between  $q$  and  $q_{CSL}$ , hardening modulus  $H$  was assumed to be the function of  $\delta$ .

$$H = H_b + H_0 \frac{\delta}{\delta_0 - \delta} \quad (29)$$

Here  $H = \frac{1}{A}$  and  $A$  is the plasticity coefficient.

Equation (29) shows that  $\delta = 0$  and  $H = H_b$  when the stress point reaches the boundary surface that is equal to the hardening modulus of the boundary surface. When  $\delta = \delta_0$  and  $H = \infty$ , then no plastic strain is produced. If the boundary surface is replaced with the failure surface, the Equation (29) can also be extended to the isotropic hardening model, as

$$H = H_0 \frac{\delta}{\delta_0} \quad (30)$$

Equation (30) shows that  $H = 0$  when reaching the failure state.  $H_0$  is the hardening modulus under the isotropic consolidation (Figure 2).

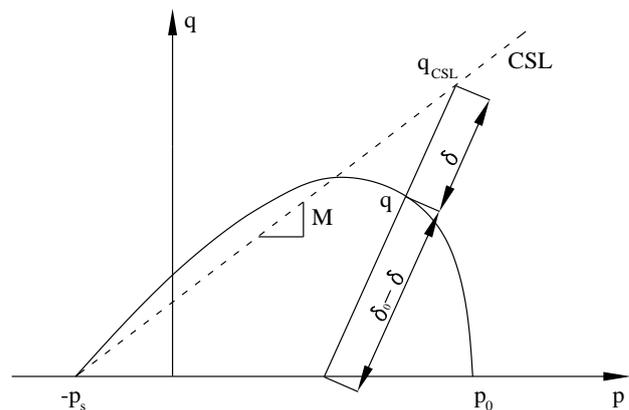


FIGURE 2 Dual Stress of  $p$ - $q$  space

### 3.3 MODEL VALIDATION

Both the improved model and BBM are implemented to simulate the tri-axial compression tests under controlled suction conditions for the silt sand [10], and the simulation is compared with experimental data. The impact and difference of the two elasto-plastic constitutive models for unsaturated soils are compared and the results indicated that our new model is superior to the conventional BBM.

### 3.4 MODEL PARAMETERS

The parameters of numerical model are shown in Table 1. The parameters of improved model and BBM include  $p^c$ ,  $p_0^*$ ,  $\lambda(0)$ ,  $\kappa$ ,  $r$ ,  $\beta$ ,  $s_0$ ,  $\lambda_s$ ,  $\kappa_s$ ,  $G$ ,  $M$  and  $k$ . All parameters can be obtained through three methods of following:

- (1) The isotropic drainage compression test (loading-unloading) under different suction conditions can determine  $p^c$ ,  $p_0^*$ ,  $\lambda(0)$ ,  $\kappa$ ,  $r$  and  $\beta$ ;
- (2) The cycle test of wet and dry under average net stress can determine  $s_0$ ,  $\lambda_s$  and  $\kappa_s$ ;
- (3) The test of shear strength under different suction conditions can provide data to define  $G$ ,  $M$  and  $k$ .

TABLE 1 Parameters of model [10]

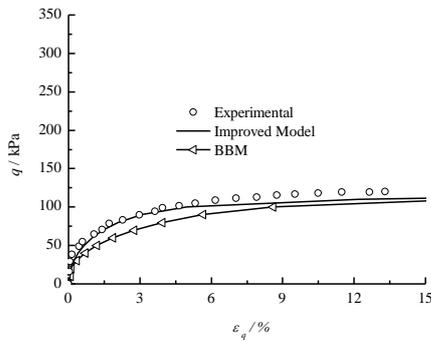
Parameter name	Parameter values	Unit
$\lambda(0)$	0.220	-
$\kappa$	0.011	-
$\beta$	17.89	MPa-1
$r$	0.210	-
$p^c$	36	kPa
$G$	8.800	MPa
$M$	0.982	-
$k$	1.324	-
$p_0^*$	41	kPa

3.5 RESULTS AND ANALYSIS

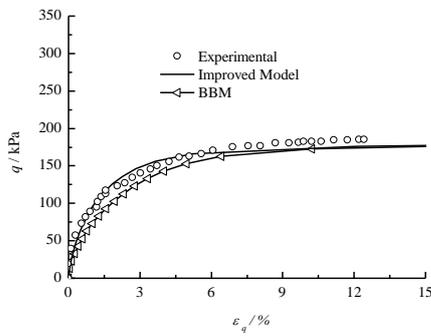
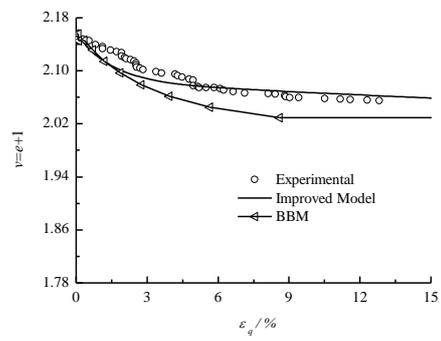
The numerical results of the two types of model, and the comparison with the experimental studies are shown in Figure 3.

The predicted curves show that two predicted models suggest consistent changes and trends with the experimental studies. For the stress-strain curves, the tendency of the improved model is harder than BBM and agrees better with most of the experimental results. However, the theoretical value of final calculated failure stress for two yield surface is equal. For the deformation curve, that is, the specific volume and total shear strain curve, the results of two yield surfaces achieve the same prediction accuracy.

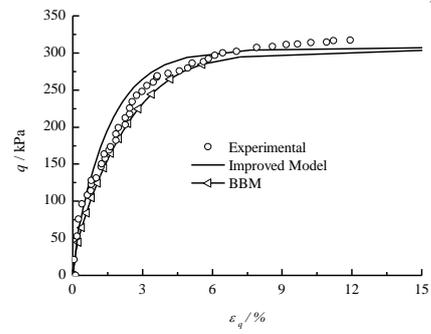
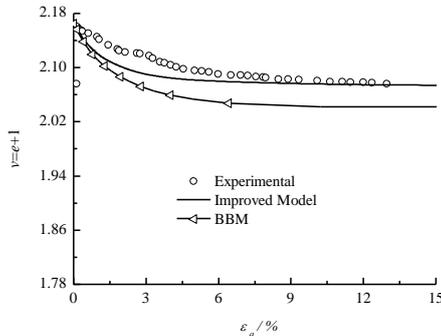
The application of drop-shaped yield surface and dual stress hardening theory in elasto-plastic model for unsaturated soils can provide higher simulation accuracy of the deformation characteristics of unsaturated soil, than the elliptical yield surface-based model



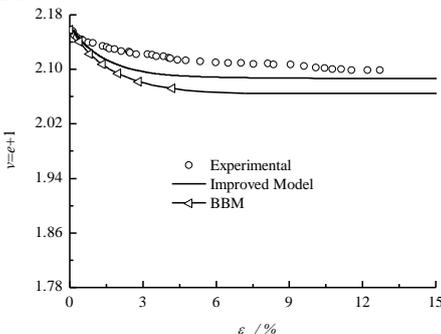
a)  $p_{mi}=50\text{kPa}$  and  $s=50\text{kPa}$



b)  $p_{mi}=50\text{kPa}$  and  $s=100\text{kPa}$



c)  $p_{mi}=50\text{kPa}$  and  $s=200\text{kPa}$



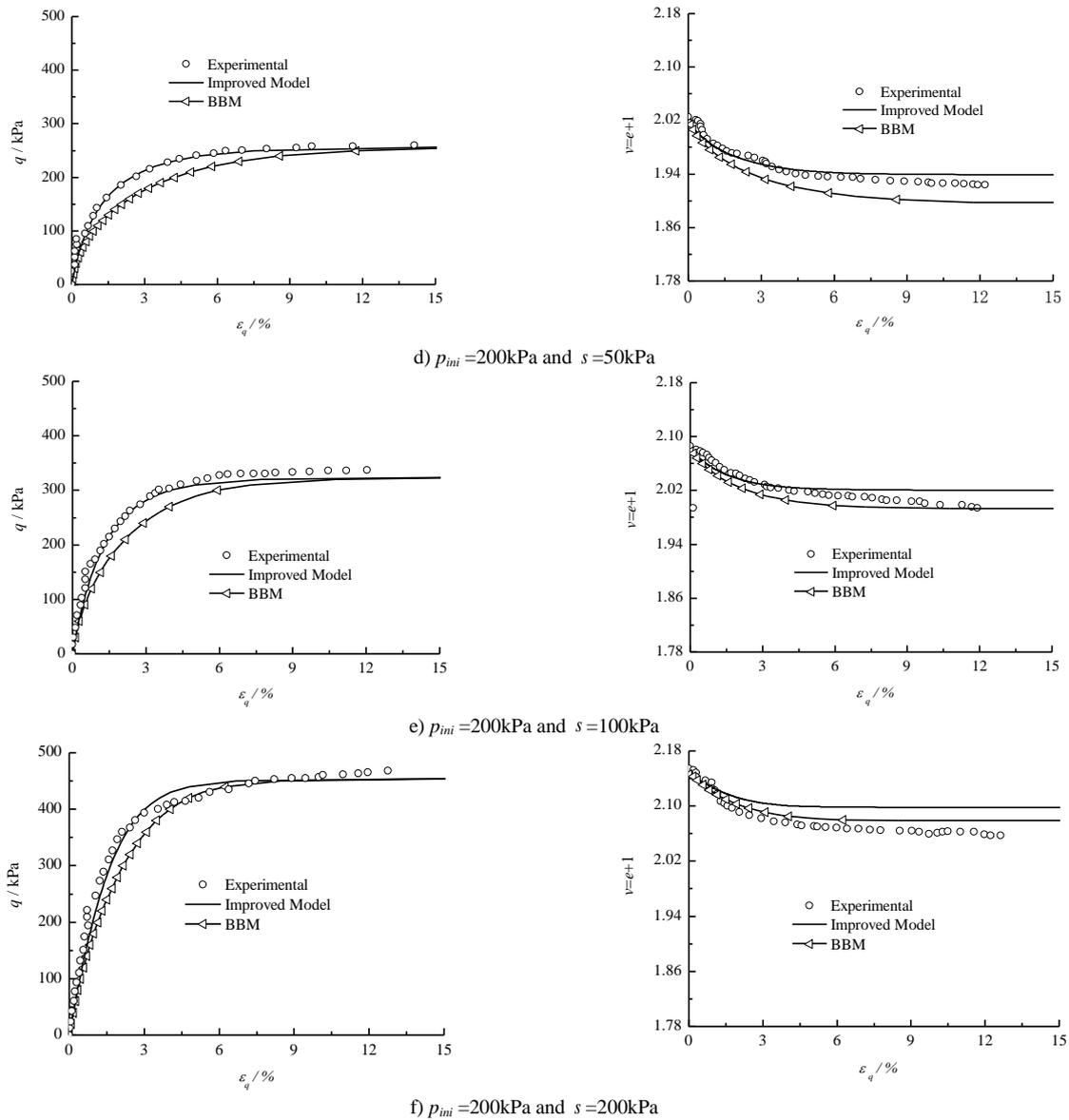


FIGURE 3 Experimental and predicted curve of  $q - \varepsilon_q$  and  $v = 1 + e - \varepsilon_q$

**4 Conclusions**

In this paper, we extended the drop-shaped yield surface to unsaturated soil and proposed the function of drop-shaped shear yield surface, with the associated flow rule to improve BBM. The dual stress hardening theory was then used in instead of the iso-surface hardening theory in the original model for the controlling of deformation in the improved model. The comparison of simulation models and the experimental results show that the improved model can be used to simulate the tri-axial compression test of controlled suction for unsaturated soil, and the results are consistent with the experimental studies. Comparison of two types of calculated results further show that the elasto-plastic constitutive model of

shear yield surface of drop-shaped can provide higher-accuracy prediction than the elliptical yield surface for unsaturated soils. The only disadvantage we observed in our model is that the calculated stress-strain curves of BBM for unsaturated soil was softer.

**Acknowledgments**

This research is supported by the Natural Science Foundation of China (No. 50527803), the Hebei Provincial Natural Science Foundation of China (No. E2009000383) and the Liaoning Technical University Dr. Science Research Foundation of China (Grant No. 13-1061).

## References

- [1] Gens A 2009 Some issues in constitutive modelling of unsaturated soils *Proceedings of 4th Asia Pacific Conference on Unsaturated Soils* 613-26
- [2] D'Onza F, Gallipoli D, Wheeler S, Casini F, Vaunat J, Khalili N, Laloui L, Mancuso C, Masin D, Nuth M, Pereira J M, Vassallo R 2011 Benchmark of constitutive models for unsaturated soils *Geotechnique*, **61**(4) 283-302
- [3] Zhenghan C 2014 On basic theories of unsaturated soil and special soils *Chinese Journal of Geotechnical Engineering* 2014 **36**(2) 201-72 (in Chinese)
- [4] Fredlund D G, Morgenstern N R, Widger R A 1978 The shear strength of unsaturated soils *Canadian Geotechnical Journal* **15**(3) 313-21
- [5] Alonso E E, Gens A and Hight D W 1987 Special problem soils *Proceedings of the 9th European Conference on Soil Mechanics and Foundation Engineering* 1087-146
- [6] Alonso E E, Gens A, Josa A 1990 A constitutive model for partially saturated soils *Geotechnique* **40**(3) 405-30
- [7] Alonso E E, Pereira M J, Vaunat J, Olivella S 2010 A microstructurally based effective stress for unsaturated soils *Geotechnique* **60**(12) 913-25
- [8] Alonso E E, Pinyol N M, Gens A 2012 Compacted soil behaviour: initial state, structure and constitutive modelling *Geotechnique* **63**(6) 463-78
- [9] Jianjun D, Longtan S 2009 Critical state surface in p-q-s stress space based on deformation of middle part of specimens *Chinese Journal of Geotechnical Engineering* **31**(10) 1607-13 (in Chinese)
- [10] Macari E J, Hoyos L R, Arduino P 2003 Constitutive modeling of unsaturated soil behavior under axisymmetric stress states using a stress/suction-controlled cubical test cell *International Journal of Plasticity* **19** 1481-515
- [11] Zhujiang S, Yuanming Z 1991 Experimental study on creep behavior of rockfills *Proceedings of the 6th Soil Mechanics and Foundation Engineering Conference* 443-6 (in Chinese)
- [12] Jianjun D 2008 Study of stress-strain characteristics of unsaturated compacted soil based on digital image measurement *Ph.D. Thesis* Dalian: Dalian University of Technology (in Chinese)
- [13] Wenxi H 1980 The influence of the hardening law on the formulation of the elasto-plastic model of soil *Chinese Journal of Geotechnical Engineering* **2**(1) 1-11 (in Chinese)
- [14] Zhujiang S 1994 Comparison among three kinds of hardening theories *Rock and Soil Mechanics* **15**(2) 13-9 (in Chinese)

## Authors



**Jianjun Dong, born in July, 1978, Huludao, Liaoning Province, P.R. China**

**Current position, grades:** Associate Professor of Department of Civil Engineering, College of Mining Engineering, Liaoning Technical University, China.

**University studies:** Ph.D at Dalian University of Technology in China.

**Scientific interest:** constitutive relations and experimental studies of unsaturated soil, the structural stability analysis of geotechnical engineering.

**Publications:** more than 18 papers published in various journals.

**Experience:** teaching experience of 6 years, 3 scientific research projects.

# Water trickling and roof falling of soft argillaceous roadways with its composite supporting and repairing technology

Wenyu Lv<sup>1, 2\*</sup>

<sup>1</sup>School of Energy Engineering, Xi'an University of Science and Technology, Xi'an, 710054, China

<sup>2</sup>Technology Centre, Sichuan Coal Industry Group Co., Ltd., Chengdu, 610091, China

Received 6 October 2013, www.tsi.lv

## Abstract

Using methods such as the transient electromagnetic method, rock mechanics testing, X-ray diffraction analysis, rock strength weakening experiment, borehole observation technology and grout mix design, etc. to study the mechanism of water trickling and roof falling of soft argillaceous roadways, via model analysis and numerical calculations, we reached many conclusions: Rock breakage is perpendicular to the axis of borehole; expansion of argillaceous swelling rock weakens the surrounding rock, the strength degradation of which occurs in 1h; performance of anchor agent also significantly decreased; uniaxial compressive strength, friction angle and cohesion of surrounding rock decreased by 40%, 16.5% and 11.1%, respectively; water trickling further exacerbated the risk of roof falling in the construction of roadway; according to calculation, the weak plane of the original 12 # I-beam became shear failure, and cannot meet the large initial deformation of roadways, floor heave is large with serious water logging. We presented the technology of "targeted drainage, deep-and-shallow hole grouting, advanced ductile, floor anchor rope, U29-shaped steel, and anchor net spray composite support". Site monitoring on 1# railway cross-cut of Bofang Coal Mine displayed that the maximum amount of deformation was less than 125mm, working resistance of anchor bolt and cable were 72-91kN and 123.3kN, respectively. Water trickling and roof falling were controlled; water logging of floor heave was improved obviously. Numerical simulations show that plastic zone is greatly reduced, and support effects were fairly good.

**Keywords:** soft and argillaceous, X-ray, water trickling and roof falling, composite support, numerical simulation

## 1 Introduction

For a long time, soft rock roadway support has been the difficulty of coal mine roadway support. China has made considerable progress in the research of soft rock control theory and basic theory of soft rock [1]. However, many research results have been concentrated on broken soft rock and high stress soft rock [2-4], while research of soft rock on argillaceous rock under water conditions has not been taken seriously and promoted high enough. Taken 1# railway crosscut of BoFang coal mine as engineering background, this paper studied the water trickling, roof falling, support and repair process of argillaceous roadways.

## 2 Engineering situations

### 2.1 GEOLOGICAL CONDITIONS

Buried depth of 1# railway crosscut is 39 m ~ 150 m and floor-elevation is +1401m, direction of drift driving is parallel with the shaft orientation, and its opening segment is located in the floor of 6-2 coal Seam. 1# railway crosscut passed through each coal seam and its surrounding rocks are mostly silty mudstone, argillaceous siltstone, siltstone and mudstone. The surrounding rocks have developed joints and fissures with strong water

conductivity. No significant water blocking layer overlying, roadway construction is largely affected by trickling water in the roof. Roadway layout near 1# railway crosscut is shown in Figure 1.

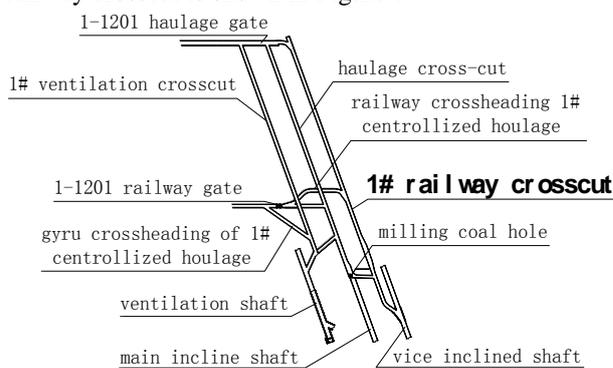


FIGURE 1 Roadway layout nearby 1# railway crosscut

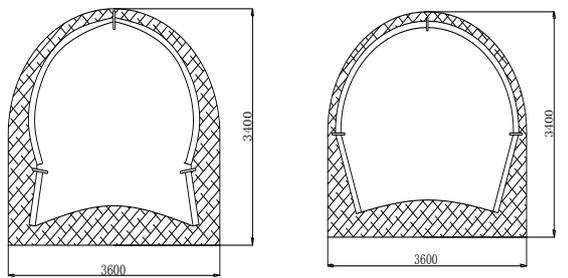
### 2.2 SUPPORTING SITUATION

1# railway crosscut is a roadway with the width 3.6m, height 3.4m, straight wall semi-circular cross-section. Support pattern is 12 # I-beam for mining (shield distance 800mm) + double-network + jet C20 concrete, railway steel (shield distance 800mm) + double-network+ jet C20 concrete. Deformation and damage of 1# railway crosscut is serious in original support. In the 43 days from January

\* Corresponding author e-mail: lvwenyu2816@126.com

20, 2014 to March 4, the largest two sides move quantity of the middle yard segment (70 m~ 110 m from the opening) reached 1800 mm and the kick drum was serious; the largest convergence between roof and floor in this segment reached 1600 mm. Roadway's sides swelling out and roof subsidence were also serious, concrete in vault and shoulders cracked, and connections of steel support broke. Most of the bolt trays were very loose and the anchorage performance of anchor agent significantly reduced [5].

After the destruction of the top support connection, support was like "人" shape under pressure from the two sides. The connections of straight leg segments broke when the pressure from one side was greater. The insertion length of the leg reached 800 mm when pressure from roof was greater. Sketch of deformation and failure in 1# railway crosscut is shown in Figure 2.



a) leg of support broken      b) leg of support move inward

FIGURE 2 The sketch map of the deformation and failure

### 2.3 GENERAL SITUATION OF WATER TRICKLING, ROOF FALLING AND GEOPHYSICAL EXPLORATION

Roof collapse happened when 1# railway crosscut dug 124 meters in. The caving length was 7.8 meters, the largest caving height was 5.6 meters, and the caving position was 12.3 meters away from the driving face. While water trickling was serious before and rear the caving zone, but there was not a lot of water during excavation [6, 7]. This shows that serious water trickling was mainly due to the fractures generated when digging roadway, and these fractures were connected to the water bodies.

Considering that drilling method will produce a great impact on roadway excavation, after the roof falling, the geophysical prospecting has been done in the working face of 1# railway cross-cut with an YCS160 intrinsically safe transient electromagnetic instrument (TEM). There are 39 dates from 1 (working face) × 13 (view angle) × 3(direction) and the three directions were bottom, top and along driving direction, detection was along rock and floor direction and probing direction was 45° upward the level of the roof. Its apparent resistivity contour lines pseudosection (achievement) was shown in Figure 3.

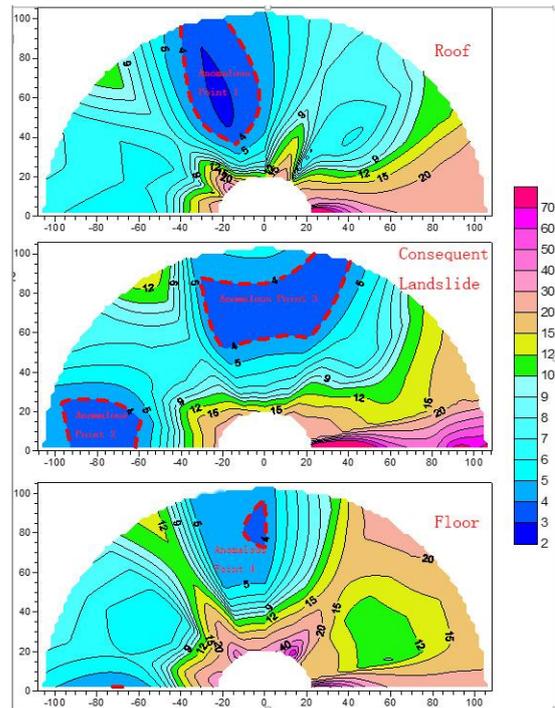


FIGURE 3 Apparent resistivity contour lines pseudosection (achievement) map of 1# railway cross-cut

In this figure, the values in the contours represent the conductive strength of underground rock. The higher the value is, the weaker the conductivity is. The smaller the value is, and the stronger the conductivity is. Detection conclusions are as follows based on the detection results: Roof direction detection results, abnormal point 1: abscissa -35 ~ -5m, ordinate 35 ~ 95m. Detection results along driving direction, abnormal point 2: abscissa 90 ~ -60m, ordinate 3 ~ 25m; abnormal point 3: abscissa -30 ~ 40m, ordinate 50 ~ 95m. Bottom direction detection results, abnormal point 4: abscissa -10 ~ 5m, ordinate 70 ~ 90m. Apparent resistivity is relatively low and water-abundance is better. We hypothesized that the water is from the water bearing structure, and also may be caused by water in the drilling process. Apparent resistivity is relatively high in other detection zone. In order to provide reliable data for repair scheme, surrounding rock components analysis and borehole observation also have been done in the following.

### 3 Surrounding rock components analysis

#### 3.1 X-RAY DIFFRACTION EXPERIMENT

The mineral composition of soft rock is the key factor that determines its mechanical properties, according to the engineering characteristics of large water and roof fall project in 1# railway crosscut. Mechanical test and X-ray detection on representative rock samples in roof falling zone have been done. X-ray diffraction instrument was D/Max-3B. A diffraction pattern is shown in Figure 4, the main components are indicated and concrete components are given in Table 1.

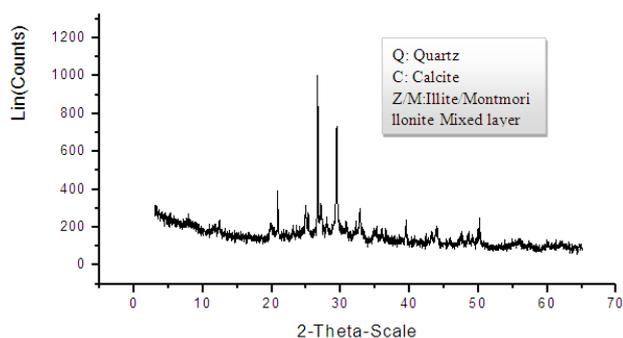


FIGURE 4 X-ray diffraction pattern

TABLE 1 experimental results of X ray diffraction

Sample No.	1#	2#	3#	4#	5#
Main Components	I/S, Q, C	I/S, Q, K, S	Q, K, D, S	I/S, Q, D, K	Q, S, K

Table: I/S: illite/montmorillonite mixed layer; Q: quartz; C: calcite; K: kaolinite; S: siderite; D: dolomite

### 3.2 SURROUNDING ROCK COMPONENTS ANALYSIS

The main components of the surrounding rock are: illite/montmorillonite mixed layer, quartz, calcite, kaolinite, siderite and dolomite [8], among which the kaolinite and illite belong to clay shale swelling rock, volume of this two kinds of rocks increase smaller when influenced by water, water can only weaken the connection force between particles. And montmorillonite expands when there is water, its volume change greatly resulting in the weakening of roadway surrounding rock.

1# railway crosscut is layout in the aquifer, joint and fracture of its surrounding rocks developed and water conductivity of the rock is also very good. Deformation characteristics are sensitive to moisture content [9]. The water content of this natural rock is 16.82% and saturated moisture content of this kind of rock is only 17.03%, though moisture content increased only 0.21% from natural state to saturation state, the uniaxial compressive strength decreased 40 percent, friction angle decreased 16.5%, and cohesion decreased 11.1%, as shown in Figure 5.

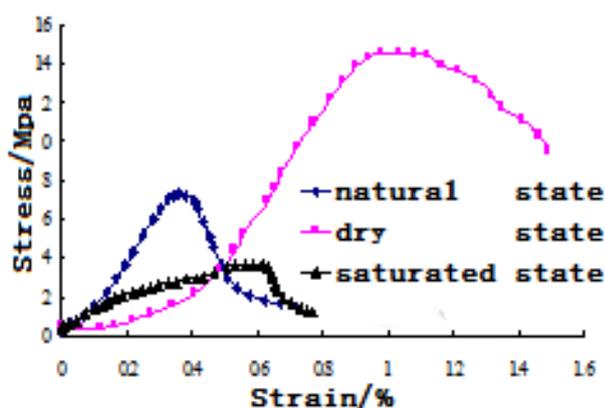


FIGURE 5 The effect of weakly cemented soft rock stress-strain curve

### 3.3 WEAKEN DEGREE OF MUDSTONE STRENGTH IN WATER

Take mudstone and argillaceous siltstone samples without flooding from the roadway to do the strength test in soaking in water.

In order to study the degree and speed of the impact of water on the mudstone rocks, the time interval is 60min and the results is shown in Figure 6.

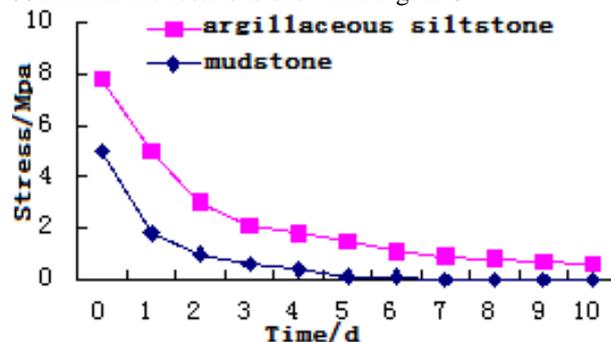


FIGURE 6 Attenuation curve of rock immersion strength

As seen from the figure, strength of the mudstones soaked in water decreases rapidly, and the water can decay the mudstones within a few hours. Strength degradation occurs mainly within the first 1h, less strength degradation occurs in 1-4h and strength will remain at 0.4MPa after 6 ~ 7h. At the same time, minerals that has cementation functionary in original rock will drain with the leaking water, and the anchorage performance of anchor agent significantly reduced.

### 4 1# railway crosscut surrounding rock structure detection

In order to detect the lithology, thickness, faults, fractures and other geological structure and their changes in roadway fracture development ring, the structure of 1# railway crosscut was probed with a YTJ20 rock drilling detector. The rock destruction within each borehole was shown in Figure 7 along the drilled axially from the inside out, in which drilling peep of 2# drilling was shown in Figure 8.

1) 1# railway crosscut is located in the stress concentration area which is in the transition section of topsoil to bedrock. By the drilling peep we can know the surrounding rocks are mostly silty mudstone, argillaceous siltstone, siltstone and mudstone. The surrounding rock has developed joints and fissures, and the structure is also very bad and rock cleats are disordered. Small structures were found during roadway excavation, no significant water blocking layer overlying, water trickling further exacerbated the danger of roof cave.

2) Through comprehensive analysis of the situation of each borehole peep, we discovered that within the range of 1.2 ~ 1.5m depth surrounding rock was severely broken. This range can be considered as broken loose body with lower bearing strength.

3) By monitoring the greatest damage depth from drilling peep, we know that damage within 3.65m was more serious; and with the increase of the drilling depth, rock destruction is easing. Therefore, it can be considered that the failure depth of roadway surrounding rock is 3.6m.

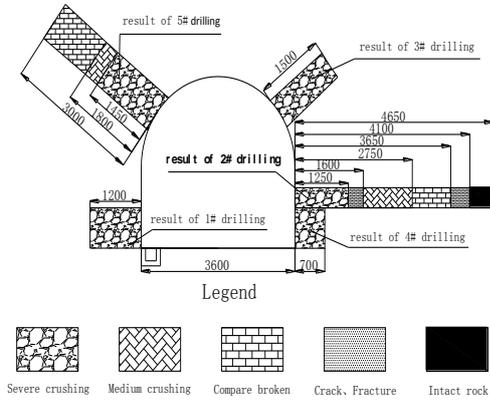


FIGURE 7 Distribution map of broken rock zone

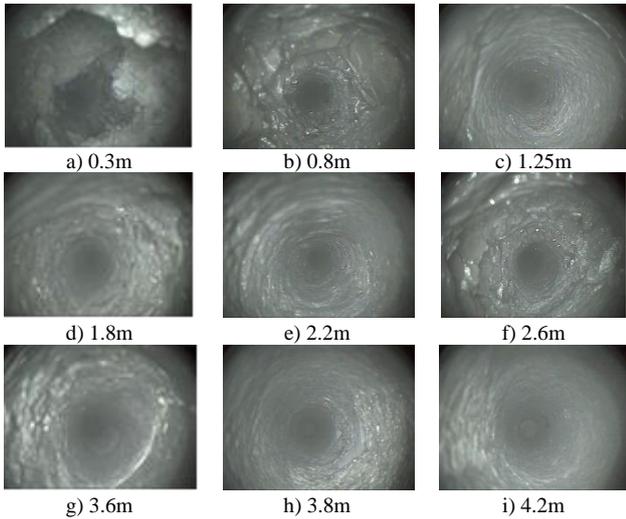


FIGURE 8 Picture of borehole peep

4) Through detection, we know that fracture is mainly caused by roadway excavation and lack of supporting strength. Most of the fractures are perpendicular to the borehole axis direction. The original support cannot guarantee the stability of surrounding rock, so the original support is failure.

**5 1# railway crosscut repair scheme**

**5.1 SUPPORT PARAMETERS BASED ON MATHEMATICAL MODEL**

**5.1.1 Failure depth of the two walls and roof**

By natural equilibrium arch theory, failure depth of the two walls was calculated by Equation (1):

$$C = \left( \frac{K_{cx} \gamma HB}{10^4 f_y} - 1 \right) h \tan \frac{90^\circ - \varphi}{2}, \tag{1}$$

where

$K_{cx}$  – the surrounding rock extrusion stress concentration coefficient, determined by the cross-section shape and aspect ratio of roadway;

$\gamma$  – the average gravity density from the roadway to the surface,  $kN/m^3$ ;

$H$  – depth from the roadway to the surface, (m);

$B$  – dimensionless parameter represents the influence degree of mining;

$f_y$  – coefficient of rock hardness, (m);

$\varphi$  – internal friction angle of surrounding rock, ( $^\circ$ ).

According to the relevant data of BoFang coal mine, the roadway side failure depth  $C=0.56m$ . The damage depth of roof strata  $b$  can be calculated according to Equation (2):

$$b = \frac{(a + C) \cos \alpha}{k_y f_n}, \tag{2}$$

where

$a$  – half width of roadway, (m);

$\alpha$  – coal seam dip angle;

$k_y$  – stability coefficient of the to be anchored rock;

$f_n$  – coefficient of hardness of the anchored rock;

According to the relevant data of BoFang coal mine, failure depth  $b=0.92m$ .

**5.1.2 Roof pressure**

The calculation formula for the roof pressure according to the normal of rock beddings is:

$$Q_H = 2\gamma_n a b B = 83.13 \text{ kN/m}. \tag{3}$$

Roadway width is 3.6 m, and then the pressure of bracket is:

$$P = 2a Q_H = 299.25 \text{ kN}. \tag{4}$$

**5.1.3 Carrying capacity**

Calculation of bearing capacity of rail steel support and 12# I-steel support in 1# railway crosscut:

$$P_{18} = \sum_{i=1}^n Q_i L_i = Q_a \times \frac{\theta_2}{180} \times 3.14 + 2 \times Q_b L_b = 111.8 \text{ kN}. \tag{5}$$

$$P_{12} = \sum_{i=1}^n Q_i L_i = Q_a \times \frac{\theta_2}{180} \times 3.14 + 2 \times Q_b L_b = 223.02 \text{ kN}. \tag{6}$$

Calculation of bearing capacity of U29 type steel bracket under the same conditions:

$$P = \sum_{i=1}^n Q_i L_i = Q_a r \times \frac{180}{180} \times \pi + 2 \times Q_b t = 406.89 \text{ kN}. \tag{7}$$

The results show that, the original 12# I-steel support and 18 kg/m rail steel support are all rigid metal support, the bearing capacity is lower than the required load and cannot adapt to the muddy roadway initial large

deformation, support welds are weak and prone to shear failure. U29-type steel support is easy to connect, more flexible and easy in fabrication, the load-carrying capacity is 1.82 times of 12# I-steel and 3.64 times of 18 kg/m rail steel support and this can achieve the required load.

5.1.4 Repair scheme

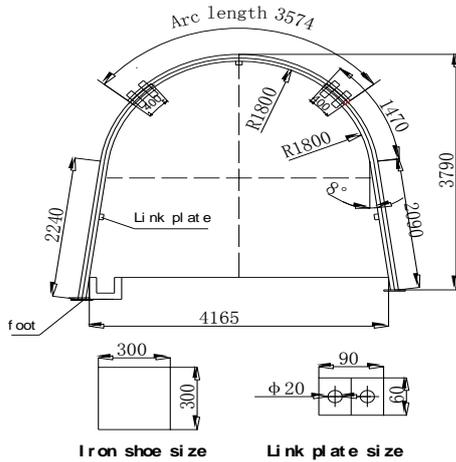


FIGURE 9 Design drawing of U-steel support

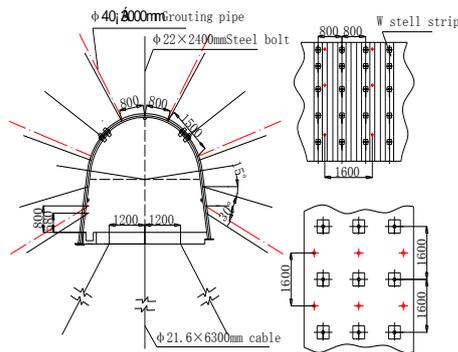


FIGURE 10 Orthographic views of combined support

TABLE 2 Tables of support parameters

Name	Parameters	
Bolt	Specifications	bolt specifications sinistral thread steel bolt with no longitudinal reinforcement $\Phi 22 \times 2400$ mm
	Length	1500mm
	Resin cartridge	K2335 one roll, Z2350 one roll
	Row & Line Space Number	800mm $\times$ 800mm 11
Cable	Specifications of trays	200mm $\times$ 200mm $\times$ 10mm
	Cable specification	$\Phi 21.6 \times 6300$ mm
	Row & Line Space Number	1200mm $\times$ 1600mm 3
Grouting Pipe	Specifications of trays	300mm $\times$ 300mm $\times$ 15mm
	Parameters	$\Phi 40 \times 3000$ mm, 1500 $\times$ 1600mm
Steel Strip	Model	W-type steel band (BHW-280-3.00)
Mesh Reinforcement	Material	$\Phi 6$ steel, 1000mm $\times$ 1500mm, mesh Size 100mm $\times$ 50mm

Through analysing the failure mechanism of 1# railway crosscut, we know that in the original support rock is broken, water trickling is serious with roof falling, roadway deforms severely and carrying capacity is low. The floor heave and waterlogging are serious, which is the sally port of the railway deformation. The two sides first broke under the stress caused by rock expansion, as the destruction continues the roof would also damage and finally the whole roadway broke. So we presented the technology of “targeted drainage, deep-shallow hole grouting, advanced conduit grouting, floor cable, U29-shaped steel anchor and net spray composite support” [10-13]. Cross-section design and orthographic views are shown in Figures 9 and 10, and bolt net support parameters is shown in Table 2.

5.2 GROUTING PARAMETERS

5.2.1 Slurry ratio parameter determines

In order to determine the ratio of grouting material, 425 # ordinary portland cement and 40 degrees Baume water glass (Formula  $\text{Na}_2\text{O}_n\text{SiO}_2$ ) were used in laboratory experiment. The model is 7.07  $\times$  7.07  $\times$  7.07 cm cubes and stripping time is about 24 hours. After numbering the test blocks, these blocks would be maintained at constant temperature and humidity (about 20 Celsius), strength test was done on test blocks with a servo universal tester according to different age groups. The results are shown in Table 3.

TABLE 3 Grout setting time and intensity in different proportions

Groups	W:C	S:C	Setting time (min)	Strength (MPa)	
				The seventh day	The fourteenth day
I	1:1	1:0.06	216	3.9	7.5
	0.9:1		191	4.3	10.4
	0.8:1		179	6.1	12.7
	0.7:1		156	8.2	16.1
	1:1		293	3.8	8.6
II	0.9:1	1:0.05	245	4.1	10.7
	0.8:1		229	6.3	14.3
	0.7:1		210	8.3	16.7
	1.0:1		327	5.1	9.6
III	0.9:1	1:0.04	289	6.1	11.9
	0.8:1		261	8.1	17.7
	0.7:1		245	9.6	20.9
IV	1.0:1	1:0.03	378	4.3	8.8
	0.9:1		337	5.2	11.2
	0.8:1		325	7.9	17.1
	0.7:1		301	9.1	18.4

When cement and sodium silicate ratio S:C is 0.04:1 and water cement ratio is 0.7:1, the strength of test blocks is high. Considering strength, gel time, slurry pump and other factors, cement and sodium silicate is identified as S: C 0.04:1, water cement ratio (0.7~1).

5.2.2 Depth and row & line space of grouting hole

Empirical equation of rock fracture degree:

$$r_y = (0.78 + 2.13\gamma H / R_c) a, \tag{8}$$

where:

- $\gamma$  – rock density;
- $H$  – buried depth of roadway;
- $R_c$  – strength of rock mass;
- $a$  – roadway radius.

Cement grout can be only injected in the rock crack, which is more than 0.2 mm in width. Through calculation, the fissure development range of 1# railway crosscut is 3.7 m, the depth of grouting hole should be not less than 1.5~2.0m considering the permeability of rock. In addition, considering the permeability region of grout determining the depth of grouting hole 3500 mm

For the convenience of construction, grouting hole spacing is often designed the integer times to bolt spacing or metal support spacing. And penetrating distance of two grouting holes should have some overlap.

Grout diffusion distance is 0.5~3.0 m in field test of grouting, the design of grouting hole spacing and single hole diffusion distance roughly the same, and here is 1.5 m.

5.2.3 Determining the grouting pressure

A large number of experiments show that grouting pressure is generally less than 2 Mpa if grouting hole depth is limited and rock fissures are developed. To prevent rock splitting damage caused by grouting, grouting pressure cannot exceed 1/10 of rock's compressive strength in soft rock. Comprehensive consideration suggests that the grouting pressure was 3 MPa.

5.2.4 Targeted drainage and deep-shallow hole grouting

Targeted drainage: grouting before drainage to close the cracks, and after to cement fractured rock timely, crowd out the remaining water, fill the fractures caused by rock dehydration, reduce the attenuation of rock strength after water, improve the cohesion, angle of internal friction and tensile strength of rock, and finally improve the overall performance of rock compressive to maintenance the stability of roadway.

Shallow hole grouting: low-pressure grouting was used in shallow hole in order to consolidate the large fissures in shallow first and form a grout layer in the rock surface [14].

Deep grouting: deep hole grouting was conducted after forming a grout layer in shallow, this will enhance the bearing capacity of the roadway in a large range.

5.3 THE SUPPORTING EFFECT MONITORING

After the application of the control technology in BoFang coal mine, observation of roadway surface displacement and stress testing of anchor and cable were done. The results were shown in Figures 11 and 12.

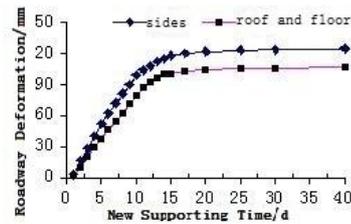


FIGURE 11 Surface displacement of roadways

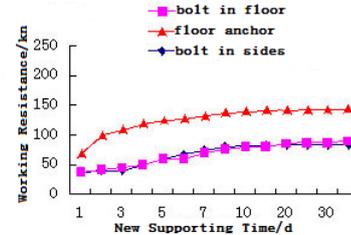


FIGURE 12 Resistance detection of bolt and anchor cable

1) Surface displacement of roadway [15, 16]. The maximum deformation of roadway sides was 124mm after the adoption of new supporting technologies and the maximum deformation of the roof and floor was 102mm; roadway deformation rate is relatively large in the former 12 days of support, the deformation rate of roadway was significantly lower 12 days later, and 15 days later the roadway was basically stable. 40 days after the introduction of new support technology, the surrounding rock has been stabilized.

2) Working resistance of cable and bolt. Initial support resistances of bolts and cables were about 36kN and 71kN respectively. The bolt support resistance grew significantly in the previous five days, and cable resistance had more growth in the previous 10 days. Surrounding rock was stable in 30 days, and support resistances of bolts and cables are 72-91 kN and 123.3 kN respectively. Floor support was strengthened, the two sides and roof deformation was prevented and the water condition has remarkable improved.

6 Numerical calculations based on FLAC3D model

The deepest point of 1# railway crosscut is 150m. Due to the complex geological conditions incurred by faults and complicate mechanical calculation of the stability of the surrounding rock [17], it is unable to carry on accurate analysis of the stability. This section mainly uses FLAC3D model to establish the expanding brush repair model (U-shape steel+ top slope bolt+ floor cable+ grouting model) [18, 19] to simulate plastic zone distribution of roadway surrounding rock in order to test the support effect. The result is shown in Figure 13.

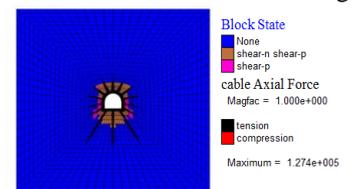


FIGURE 13 Plastic zone distribution of roadway

By calculating we concluded that the boundary of the upper part of the plastic zone is 1.37 m from the roof, the boundary of the lower part is 2.69m from the floor, boundary of the left part is 1.35m from the left side and boundary of the right part is 1.30m from the right part. The width of plastic zone of the two sides is shorter, and the scope of plastic zone is greatly reduced after the use of bolts and cables. Force of bolts and cables is relatively reasonable after using the repair program, in which the largest axial force of bottom cable is 127.4 kN. And the tails of anchoring segment are in the outside of the plastic zone of surrounding rock.

## 7 Conclusions

1) The surrounding rock of 1# railway crosscut contains mainly illite, montmorillonite and kaolinite, so the rock is easy to soften, fragmentation, collapse and swelling in water, resulting in the weakening of rock structure.

## References

- [1] He Man-chao, Jing Hai-he, Sun Xiao-ming 2000 Research Progress of Soft Rock Engineering Geomechanics in China Coal Mine *Journal of engineering geology* **1** 46-62 (In Chinese)
- [2] Kang Hong-pu, Wang Jin-hua, Lin Jian 2007 High pretensioned stress and intensive bolting system and its Application in deep roadways *Journal of China Coal Society* **12** 1233-8 (In Chinese)
- [3] Sun Jun 1998 Research of rock mechanics on the joint of the century *Proceedings of the 5th Science conference of Chinese society for rock mechanics and engineering Chinese Science and Technology Press: Beijing* (In Chinese)
- [4] He Man-chao 1993 Survey of soft rock tunnel engineering Xuzhou: China university of Mining and technology press (In Chinese)
- [5] Zeng You-fu, Wu Yong-ping, Lai Xing-ping 2009 Analysis of roof caving instability mechanism of large-section roadway under complex conditions *Journal of Mining & Safety Engineering* **26**(4) 423-7 (In Chinese)
- [6] Xu Xing-liang, Zhang Nong 2007 Study of control process deformation behavior and of soft rock drift under rich water condition *Journal of China University of Mining & Technology* **36**(3) 298-302 (In Chinese)
- [7] Gou Pan-feng, Chen Qi-yong, Zhang Sheng 2004 Influence analysis of the anchor-hold of the resin bolt by the draining water in the drill hole *Journal of China Coal Society* **29**(6) 680-3 (In Chinese)
- [8] Liu Chang-wu, Lu Shi-liang 2001 Research on mechanism of mudstone degradation and softening in water *Rock and Soil Mechanics* **21**(1) 28-31 (In Chinese)
- [9] Gamboa E, Atrens A 2003 Environmental influence on the stress corrosion cracking of rock bolts *Engineering Failure Analysis* **10**(5) 521-58
- [10] Williams J H, Johnson C D 2004 A coustic and optical borehole-wall imaging for fractured-rock aquifer studies *Journal of Applied Geophysics* **55**(1) 151-9
- [11] Jiang Yao-dong, Zhao Yi-xin, Liu Wen-gang 2004 Research on floor heave of roadway in deep mining *Chinese Journal of Rock Mechanics and Engineering* **23**(14) 2396-2401 (In Chinese)
- [12] Pan Yi-shan, Zhu Ping, Wang De-li 1988 The creep simulation test and numerical analysis on deep roadway floor heave and its Prevention *Journal of Heilongjiang Mining Institute* **8**(2) 1-7 (In Chinese)
- [13] Marti J, Cundall P A 1982 Mixed discretization procedure for accurate solution of plasticity problems *International Journal for Numerical Methods in Engineering* **6** 129-39
- [14] You Chun-an 2004 Mechanics analysis of anchorage segment of pressure-type cable *Chinese Journal of Geotechnical Engineering* **26**(6) 828-31 (In Chinese)
- [15] Yang Xi-an, Hung Hong-gei, Liu Bao-wei 2000 Study on the deformation and support of weak rock mass roadway in the Jinchuan nickel mine *Journal of Xiangtan Mine Institute* **15**(3) 12-7 (in Chinese)
- [16] Wang Lian-guo, Li Ming-yuan, Wang Xue-zhi 2006 Study on mechanisms and technology for bolting and grouting in special soft rock roadways under high stress *Chinese Journal of Rock Mechanics and Engineering* **25**(16) 2889-94 (in Chinese)
- [17] Kang Hong-pu, Lin Jian 2001 New development in geomechanics measurement and test technology of mine roadway surrounding rock *Coal Science and Technology* **29**(7) 27-30 (In Chinese)
- [18] Lv Jun, Hou Zhong-jie 2000 Factors that influence the behavior of ground pressure under shallow conditions *Ground Pressure and Strata Control* **2** 39-43 (In Chinese)
- [19] Li Shu-qing, Wang Wei-jun, Pan Chang-liang 2006 Numerical Analyses on Support Structure of Rock around Deep Roadway *Chinese Journal of Geotechnical Engineering* **28**(3) 377-81 (In Chinese)

## Authors



**Wenyu Lv, born in May, 1981, Xian City, Shanxi Province, China**

**Current position, grades:** the Lecturer of School of Xi'an University of Science and Technology, China.

**University studies:** Ph.D. degree in Mining Engineering from China University of Mining & Technology Beijing in China.

**Scientific interest:** Computer application in coal mining, backfill mining technology.

**Publications:** More than 20 papers published in various journals.

**Experience:** teaching experience of 3 years; 3 scientific research projects.

Authors' index					
An Jing	213	Li Ling	348	Tian Cunzhi	52
Baydalina L	261	Li Qingmei	173	Wang Bin	218
Cai Fei	360	Liu Dong	342	Wang Honghai	24
Cao Jie	342	Liu Haiqiang	299	Wang Jianxiang	240
Chen Chunmei	321	Liu Hong	281	Wang Lei	354
Chen Dashan	159	Liu Honghai	186, 192, 290	Wang Qinghui	113
Chen Dongchao	294	Liu Jia	192, 290	Wang Wenjun	360
Chen Jianjun	101, 106	Liu Jian	335	Wang Xiaoxiong	142
Chen Wenqiang	330	Liu Kejian	153	Wu Han-chuan	335
Chen Yixiang	15	Liu Kunliang	125	Wu Hongjie	132
Cheng BingHua	179	Liu Long	304	Wu Jinzhao	44
Dong Jianjun	372	Liu Meiling	44	Wu Nan	89
Dou Yong	7	Liu Shengli	304	Xi Xuefeng	132
Fan Honghui	72	Liu Shue	218	Xiao Wen-sheng	335
Gao Xiang	304	Liu Zhong-yan	335	Xiao Wentao	290
Grekul V	261	Lu Minyan	89	Xin Zhenghua	37
Han Qingbang	72	Lu Yaqin	52	Xu Bo	213
He Chengbing	294	Lu You	132	Xu Chonghuan	208
He Zhenfang	225	Luo Qiuyue	255	Xu Jiang	342
Hong Xueqian	365	Lv Ming	299	Xu Pan	304
Hu Rongdong	7	Lv Wenyu	381	Xu Yuhui	255
Hua Ze	132	Ma Haifeng	354	Yan Xu	299
Huang Chunmei	65	Ma Xiaoping	37	Yang Xiani	52
Huang Jinming	125	Mikel Luján	7	Yang Xiuxia	311
Huang Wangyuan	113	Nie Xin	299	Yin Zhiqiang	354
Jia Liang	79	Niu Caihong	119	Yu Shiwei	365
Jiang Chunmao	65	Pan Huachen	299	Yu Xing	204
Jiang Guozhang	186, 192, 290	Peng Jiansheng	147	Yuan Debao	365
Jiang Hesong	321	Qin Shuijie	166	Zhang Dajian	89
Jiang Hong	321	Qu MingCheng	65	Zhang Juan	321
Jiang Jingfei	7	Qu Xiao-Yuan	97	Zhang Ni	132
Jiang Xueling	166	Ren Xiangyang	173	Zhang Yi	311
Jin Jian	240	Ren Yu	173	Zhang Yikun	192
Keshabyan-Ivanova I	59	Shao Fei	179	Zhang Zhijun	281
Lei Chongwu	186	Shi Hao-ran	153	Zhao Cuirong	24
Li Dashe	218	Song Kai	119	Zhao Yanbao	365
Li Gongfa	186, 192, 290	Song Liangrong	247	Zheng Qiuhong	247
Li Hong	37	Sun Nigang	79	Zhou Weiwei	311
Li Hongyan	271	Tang Houxing	197	Zhou Yulan	233
Li Huiyong	15	Tang Weidong	44	Zhu Hongjin	72
Li Liangjian	365	Tao Yunqi	342	Zhu Weiliang	31

## Cumulative Index

## Mathematical and Computer Modelling

**Jingfei Jiang, Rongdong Hu, Luján Mikel, Yong Dou** Accuracy evaluation of deep belief networks with fixed-point arithmetic*Computer Modelling & New Technologies 2014 18(6) 7-14*

Deep Belief Networks (DBNs) are state-of-art Machine Learning techniques and one of the most important unsupervised learning algorithms. Training DBNs is computationally intensive which naturally leads to investigate FPGA acceleration. Fixed-point arithmetic can be used when implementing DBNs in FPGAs to reduce execution time, but it is not clear the implications for accuracy. Previous studies have focused only on accelerators using some fixed bit-widths. A contribution of this paper is to demonstrate the bit-width effect on various configurations of DBNs in a comprehensive way by experimental evaluation. Explicit performance changing points are found using various bit-widths. The impact of sigmoid function approximation, required part of DBNs, is evaluated. A solution of mixed bit-widths DBN is proposed, fitting the bit-widths of FPGA primitives and gaining similar performance to the software implementation. Our results provide a guide to inform the design choices on bit-widths when implementing DBNs in FPGAs documenting clearly the trade-off in accuracy.

*Keywords:* deep belief network, fixed-point arithmetic, bit-width, FPGA

**Huiyong Li, Yixiang Chen** On denotational semantics of the complex event query language STeCEQL*Computer Modelling & New Technologies 2014 18(6) 15-23*

With the complex event processing technology has been widely used in processing the information of the internet of things, many scholars have proposed a lot of event query languages(EQL) for different scenarios. Early scholars generally study the operational semantics of EQL. Recently, many researchers begin to pay attention to the correctness of the operational semantics of the EQL. Some researchers have shown the correctness of the operational semantics by proven the equivalence between the denotational semantics and the operational semantics of EQL. The internet of vehicles is an important research branch of internet of things and it has a very wide range of applications. STeCEQL is a spatial and temporal constraint EQL for the internet of vehicles. In this paper, we focus on the correctness of the operational semantics of STeCEQL. We mainly establish the denotational semantics of STeCEQL. Finally, we prove the equivalence between the two semantics of STeCEQL. Therefore, the operational semantics of STeCEQL are correct.

*Keywords:* Complex Event Query Language, Internet of things, Mobile System, Denotational Semantics, Operational Semantics

**Cuirong Zhao, Honghai Wang** Analysis and research on the simulation and output of discrete event system with fuzzy parameters*Computer Modelling & New Technologies 2014 18(6) 24-30*

This paper discusses discrete event system simulation output analysis method with fuzzy input parameter. For a classic discrete event model, which contains randomness, once stimulation running is only a sampling according to systematic behaviour, which could not represent all features of the system. Hence, there should be a systematic analysis method, under the guidance of which to apply multi-times stimulation of model and analyse output data of stimulation. This paper provides a solution and introduces random fuzzy theory at last to improve traditional output analysis method. Result of stimulation experiment proves that the method could improve the reliability of stimulation output analysis.

*Keywords:* Discrete event system, Fuzzy Parameters

**Weiliang Zhu** Research of simulation techniques based on rough set theory*Computer Modelling & New Technologies 2014 18(6) 31-36*

Rough set theory can effectively analyse and deal with incomplete information in simulation techniques. This paper

studied the knowledge reduction problem and discrete continuous attributes and improved the BP neural network on the basis of rough set theory. Firstly, methods of attribute reduction of classical are analysis. This paper proposes a heuristic algorithm for reduction of knowledge based on information entropy. Subsequently, it studied the combination algorithm of rough set and neural network. Pre-treatment of sample data based on rough set in dealing with imprecision and uncertainty issues on the edge. The decision rules obtained after reduction in order to map to the training sample of neural network. Finally, the neuron number of hidden layer of neural network and hidden layer makes the neural network more logical. The simulation results show that the simulation technique of rough set and neural network has obvious complementary and reduce the time to train the neural network. It improved the training accuracy and generalization ability simulation techniques achieved satisfactory results.

*Keywords:* simulation techniques, rough set, knowledge reduction, BP neural network

**Hong Li, Xiaoping Ma, Zhenghua Xin** A new formal representation of granules based on features

*Computer Modelling & New Technologies 2014 18(6) 37-43*

In order to provide the unified representation of granules under Granular Computing, this paper studied the granules' basic meaning, descriptions and relationships between the features of granules, and described that the four elements of granules were closely related to granular features. Then the paper presented four-tuples formal representation on data levels based on granular features. The representation includes the object set, feature set, relationship sets and constraints set. Then the representations of several special granules were presented. They showed the unity and advantage of the representation. At last, this paper gave a specific example. By the example, the formal representation has important significance in methodology to solve the granular representation on data levels based on granular features well. And this method is conducive to solve problems and study of granular computing theory.

*Keywords:* granule, feature, formalization, representation, granular computing

**Weidong Tang, Jinzhao Wu, Meiling Liu** Interleaving semantics and action refinement in event structures

*Computer Modelling & New Technologies 2014 18(6) 44-51*

An event structure acts as a denotational semantic model of concurrent systems. Action refinement is an essential operation in the design of concurrent systems. But there exists an important problem about preserving equivalence under action refinement. If two processes are equivalent with each other, we hope that they still can preserve equivalence after action refinement. In linear time equivalence and branching time equivalence spectrum, interleaving equivalences, which include interleaving trace equivalence and interleaving bisimulation equivalence are not preserved under action refinement [9-11, 14, 16, 21]. In this paper, we define a class of concurrent processes with specific properties and put forward the concept of clustered action transition, which ensures that interleaving equivalences are able to preserve under action refinement.

*Keywords:* event structure, action refinement, concurrency, interleaving equivalence, clustered equivalence

**Xiani Yang, Yaqin Lu, Cunzhi Tian** Research on growth opportunity and liquidity monitoring by mathematical optimization

*Computer Modelling & New Technologies 2014 18(6) 52-58*

Based on the assumption of variable-investment, this paper introduces growth opportunity into the model of liquidity needs (Tirole, 2006). Through the establishment of mathematical optimization model, we analyse the influence of growth opportunities on liquidity needs and liquidity investment decisions. Both of mathematical derivation and numerical simulation show that, the entrepreneur tends to overinvest in illiquid assets if the growth opportunity is small; otherwise, he will overhoard of liquid assets. In addition, the agency costs due to information asymmetry may also affect the entrepreneur's decisions of liquid assets investment.

*Keywords:* Growth opportunity, liquidity monitoring, variable-investment model, mathematical optimization

**I Keshabyan-Ivanova** A computer-assisted analysis of literary texts: a sample study

*Computer Modelling & New Technologies 2014 18(6) 59-64*

The overall aim pursued in this work is to demonstrate how quantitative data and a range of different corpus-based analytical techniques can be used in assessing an author's literary originality in relation to his texts' structures and meanings. With this in mind, the present study provides a sample of quantitative analysis of the two literary texts – Shakespeare's *Hamlet* [1685] and Sumarokov's *Gamlet* [1787]. Prior research has explored *Hamlet* and *Gamlet* in terms of historical, philosophical, language-based, etc. approaches that have existed to date. Taking into consideration the aforementioned visions of both plays, a special perspective on *Hamlet* and *Gamlet* is adopted herein. Given the importance assigned to computer-assisted analysis of literary texts, the current study is based on the idea that the texts under examination contain a certain number of particular characters that are distributed in a special way within and among the acts and intervene with a particular frequency specified by the authors. To achieve this aim, the texts are closely read and, then, computational and quantitative resources are applied. In general, the relevant findings unveil substantial structural deviations of the presence and interventions of all main characters, leading to noticeable diversions in the role and weight assigned by the authors to them per different acts inter-plays.

*Keywords:* Corpus-based, computational, quantitative, presence, interventions

### Computer and Information Technologies

**Chunmei Huang, Chunmao Jiang, MingCheng Qu** Design of software error detection system based on SPARC V8 and research on the key technology

*Computer Modelling & New Technologies 2014 18(6) 65-71*

This paper analyses the key issues confronted when SIHFT is implemented on the SPARC V8 platform, gives the algorithm to solve the problem and the corresponding technology solutions. Software error detection technology system was designed based on SPARC V8, and software signature control flow error detection technology was implemented, the system is based on the architecture of SPARC V8, uses software signature control flow error detection technology and copy instruction error detection technology as the prototype, it's a software system which detects transient faults induced by space radiation and was developed through research, analysis and transformation, with availability, modifiability, portability, maintainability, readability, scalability and other features. The error detection coverage rate of software error detection technology suitable for target platform was tested through simulation experiments. The result data of experiments conducted in the emulator TSIM shows that on the basis of given average performance overhead, the system had high error detection coverage rate when brought in register injected fault and memory injection fault. This proved the SIHFT technology is feasible and effective.

*Keywords:* SIHFT, architecture, SPARC V8

**Honghui Fan, Hongjin Zhu, Qingbang Han** Image reconstruction of concrete based on Filtered Backprojection method using ultrasonic time of flight data

*Computer Modelling & New Technologies 2014 18(6) 72-78*

This research aims to recognize the defect of concrete materials using an ultrasonic computed tomography imaging technique. Filtered Backprojection method was used to reconstruct concrete images in this paper. Ultrasonic time of flight data was measured to reconstruct computer tomography images. 306 data paths were obtained in total by manual scanning for one computer tomography image. We examined the effect of the interpolation data as the density of time of flight data has a considerable effect on image quality. The feasibility of concrete reconstruction system and time of flight data interpolation were examined in detail using numerical and concrete phantoms.

*Keywords:* Image reconstruction, Time of flight, Filtered Backprojection, Interpolation, Concrete

**Liang Jia, Nigang Sun** A line segment detection algorithm based on statistical analyses of quantified directions in digital image

*Computer Modelling & New Technologies 2014 18(6) 79-88*

Line segment detection is a typical image processing problem with constantly evolving solutions. Following the line segment detect (LSD) by Grompone von Gioi, two branches of algorithms merged. The first branch aimed to improve

its speed at the cost of lower accuracy; the second applied in the opposite way. We investigated the philosophies of these methods, and attempt to develop a line segment detection algorithm based on statistical analyses of quantified directions (LSDSA) to achieve better accuracy and faster speed. We utilize a statistical approach estimating the distributions of pixels with direction values approximating the direction changes when traversing along the edges given by any edge detector. It efficiently reduces the dimension of the input data, and incurs limited increasing in computation time for validation process. The simulation results show that the proposed algorithm achieves better performance compared to the existing typical LSD algorithms. The experiment using industrial data in noisy cases also exhibits excellent performance.

*Keywords:* Line segment detection, Hough transformation, Image processing, Pattern recognition

**Dajian Zhang, Minyan Lu, Nan Wu** A model-based assurance case construction approach for system control software

*Computer Modelling & New Technologies 2014 18(6) 89-96*

As the massive damage caused by the failures of system control software becomes increasingly prominent, people pay more attention to the construction of assurance case to demonstrate the dependability level of system control software. In this paper, a new assurance case construction approach for system control software is proposed. Based on the metamodel of modular GSN, we give the basic procedure and tree structure deductive algorithm of the approach, and verify our work using Brake Control software used in an aircraft. The results show that the approach can develop assurance case effectively and efficiently.

*Keywords:* Software, Dependability, assurance case, GSN, modularization

**Xiao-Yuan Qu** A noise estimation approach by assembling fast edge detection and block based methods

*Computer Modelling & New Technologies 2014 18(6) 97-100*

Noise estimation is one of the most important research topics in image processing. Aishy and Eric had proposed a variance estimation method used in Gaussian white-noise, in which, a measure was provided to determine the homogeneous blocks and an analyser was used in calculating the homogeneities. The approach should present two shortcomings corresponding to structures and textures. One is that the blocks with edge textures should be considered as intensity-homogeneous blocks that could have an effect on estimation accuracy. The other is that some special blocks with high variance but low homogeneity could result in over estimation. In order to avoid the two shortcomings, in this paper we have proposed an improved noise estimation approach by combining fast edge detection and block based methods. The blocks hold continuous points were firstly excluded rejected by using fast edge detection method. The experimental results indicated that our method can avoid over estimation effectively in special conditions and can obtain more accurate results than the Aishy and Eric's method did.

*Keywords:* Guess Noise, noise variance estimation, edge detection, image processing

**Jianjun Chen** Supervised orthogonal tensor neighborhood preserving embedding for face recognition

*Computer Modelling & New Technologies 2014 18(6) 101-105*

The deficiency of supervised discriminant information is the problem of Orthogonal Tensor Neighborhood Preserving Embedding (OTNPE) proposed recently for face recognition. So a dimension reduction algorithm called Supervised Orthogonal Tensor Neighborhood Preserving Embedding (SOTNPE) is proposed in the paper. On the basic of OTNPE, the algorithm achieves neighborhood reconstructions within the same class, preserving supervised class label information and neighborhood reconstruction information. Experiments on AR and YaleB face datasets show our proposed algorithm is efficient.

*Keywords:* face recognition, dimensionality reduction, orthogonalization, tensor neighborhood preserving embedding, neighborhood reconstruction

**Jianjun Chen** Constraint-based sparsity preserving projections and its application on face recognition

*Computer Modelling & New Technologies 2014 18(6) 106-112*

Aiming at the deficiency of supervise information in the process of sparse reconstruction in Sparsity Preserving Projections (SPP), a semi-supervised dimensionality reduction method named Constraint-based Sparsity Preserving Projections (CSPP) is proposed. CSPP attempts to make use of supervision information of must-link constraints and cannot-link constraints to adjust the sparse reconstructive matrix in the process of SPP. On one hand, CSPP obtains the high discriminative ability from supervised pairwise constraint information. On the other hand, CSPP has the strong robustness performance, which is inherited from the sparse representation of data. Experimental results on UMIST, YALE and AR face datasets show, in contrast to unsupervised SPP and existing semi-supervised dimensionality reduction method on sparse representation, our algorithm achieves increase in recognition accuracy based on the nearest neighbour classifier and promotes the performance of dimensionality reduction classification.

*Keywords:* semi-supervised dimensionality reduction, pairwise constraint, sparse representation, sparse reconstruction, sparsity preserving projections, face recognition

**Qinghui Wang, Wangyuan Huang** An effective time of advent-based scheme for mitigating the influence of the non-line-of-sight propagation

*Computer Modelling & New Technologies 2014 18(6) 113-118*

Aiming at the problem that Time of Advent-based wireless location is easily influenced by non-line-of-sight, a set of effective wireless ranging system for mitigating the influence of non-line-of-sight has been proposed in this paper. At first, the Kalman filter model has been established to eliminate the influence of non-line-of-sight, as well as the random interference. Secondly, it carries out the off-line filter simulation with test data to verify the effectiveness of the model. Finally, a set of ranging system has been designed by taking the ATmega1280 microprocessor as the controller and the nanoPAN5375 as the radio frequency chip, and it can carry out the real test on the improved ranging platform. According to the results, the designed system can accomplish the real-time dynamic filter, which can reduce the non-line-of-sight error with a high measurement precision, and it can also be applied in the location under non-line-of-sight environment directly.

*Keywords:* time of advent, wireless location, non-line-of-sight, Kalman filter, nanoPAN5375

**Kai Song, Caihong Niu** Incomplete character recognition technology in the license plate recognition system

*Computer Modelling & New Technologies 2014 18(6) 119-124*

The incomplete characters recognition in the license plate, characters can be divided into linear character and curve character. For the curve characters, used a method of statistic the number of character holes as the character characteristics for feature extraction, which extended seed filling algorithm in the Computer Graphics. For linear characters, we proposed a method of extracting a character conversion slope of the line feature by Hough transform, which had a good effect on linear character recognition.

*Keywords:* license plate recognition; extraction of character feature; incomplete character recognition

**Kunliang Liu, Jinming Huang** Research and implementation of 3D reconstruction base on multi-contours

*Computer Modelling & New Technologies 2014 18(6) 125-131*

3D Reconstruction often faces to a serial of 2D contour lines but not to volume data, which as we often processed, so study of 3D reconstruction based on multi-contours has important practical values. In the process of 3D-reconstruction based on multi-contours, contours correspondence, contours splicing, branch problem, and terminal contours closing are all its key technologies. In this paper we give the concrete solutions on every step of 3D construction of multi-contours. According to winding issue of contours we provide means of gauging sum of angles of contour's edges adjacent to each other, which avoided error judging of winding of contours. As to branch issue of one contour corresponding to several contours, we give the way of splitting contour based on ratio in circumference of corresponding contours. We also give the mean of maximal field angle to reduce the calculation time on triangulating terminal contours. The solution we provided can give correct result of contours splicing under any kind of contours. It proves that every step of the solution is correct and effective. The solutions we designed are more general than other

solutions.

*Keywords:* contour, 3D reconstruction, Delaunay triangulate, convex hull

**You Lu, Xuefeng Xi, Ze Hua, Hongjie Wu, Ni Zhang** An abnormal user behaviour detection method based on partially labelled data

*Computer Modelling & New Technologies 2014 18(6) 132-141*

Detecting abnormal user behaviour is of great significance for a secured network, the traditional detection method, which is based on machine learning, usually needs to accumulate a large amount of abnormal behaviour data for training from different times or even different network environments, so the data gathered is not in line with practical data and thus affects accuracy, and that increases overhead for data labelling. In light of these disadvantages, this paper proposes the detection method based on collaborate learning, it uses under-sampling method based on distance and distribution to generate training sample from imbalanced data, and semi-supervised learning method combined by ensemble classifying method to reduce demand for labelled data, it also uses differentiated member classifiers based on mixed perturbation method for collaborate training and selectively build ensemble classifier according accuracy to detect abnormal user behaviour. Experiments based on data from simulation and real network showed that this method can effectively detect abnormal behaviour and outperform traditional methods in several evaluating indicators.

*Keywords:* abnormal user behaviour detection, collaborative learning, support vector machine

**Xiaoxiong Wang** Landscape based on three-dimensional SketchUp modelling to get visualization applications

*Computer Modelling & New Technologies 2014 18(6) 142-146*

Landscape designing CAD software are now mostly in two-dimensional draft stage in domestic, and three-dimensional visualization of designed landscape models and applications are still in the initial stage. The essay based on precise three dimensional solid modelling SketchUp software, introduces SketchUp software functions and features, and focuses on three aspects of garden design, hand-painted TIN terrain design, planting design and planning and design of garden objects. Through engineering examples to import the landscape of two-dimension vector data into SketchUp software, use SketchUp techniques flexibly to solve different details and avoid redundant and diverse work. After rendering and post processing, establishing a realistic visualization of three-dimensional entity model.

*Keywords:* SketchUp software, landscape, three dimensional solid model, engineering examples

## Operation research and decision making

**Jiansheng Peng** The robot path optimization of improved artificial fish-swarm algorithm

*Computer Modelling & New Technologies 2014 18(6) 147-152*

The robot path optimization solution is seek a collisionless path from starting point to end point to make the robot get the shortest route go along with planning path. Let robot path optimization problem map to mathematical model TSP (Traveling Salesman Problem) to resolve it, and make the corresponding algorithm realize robot path planning optimization was introduced in this paper. According to existent insufficiency of traditional artificial fish-swarm algorithm, using improved artificial fish-swarm algorithm optimizes the robot path planning, and stands out the superiority of improved artificial fish-swarm algorithm. The main improved aspect of artificial fish-swarm algorithm is increases examine link in foraging behaviour.

*Keywords:* robot path optimization, Traveling Salesman Problem, improved artificial fish-swarm algorithm, foraging behaviour

**Hao-ran Shi, Kejian Liu** Research on lead-time reduction of two-stage supply chain based on Stackelberg game

*Computer Modelling & New Technologies 2014 18(6) 153-158*

In the two-stage supply chain, under the model of lead time reduction management cost shared by upstream and downstream based on Stackelberg Game, when suppliers have the priority of decision right rather than retailers, it is

more advantageous to reduce the cost and the lead time and can reach the maximum profit for the whole supply chain.

*Keywords:* Supply chain, Lead time, Decision order, Cost sharing

### **Dashan Chen** Effect analysis of speed guidance on traffic demand and driver compliance

*Computer Modelling & New Technologies 2014 18(6) 159-165*

Active traffic management is method of increasing capacity and smoothing traffic flows. As one of the most important methods, speed guidance control could be used to improve operational efficiency and reduce accident rates. This paper aims to consider a variety of factors to determine the best traffic management services. Driver obedience for the speed guidance value affected the effect directly. The effect might also lose even play a negative role when the traffic demand reached a certain level. Simulation was carried out depending on different traffic demand and driver compliance rate through abstract urban expressway model. Six kinds of traffic demand under different obedience level were analysed comparatively. Speed guidance control has the positive effect about safety and efficiency when the traffic demand is low or medium. When the traffic demand is high, the effect on safety and efficiency both are negative, and different driver compliance rate affected the effect level to some extent.

*Keywords:* Urban expressway, speed guidance, conflict, compliance rate, effect analysis

### **Xueling Jiang, Shuijie Qin** Simulating a crowd with dynamic emotional transmission based on Hidden Markov Model

*Computer Modelling & New Technologies 2014 18(6) 166-172*

Crowd simulation has been widely applied in computer animation and graphics rendering technology. However, the social communication and emotional characteristics are often unrecognized in crowd simulation. In psychology, there are two kinds of emotional factors for humans: the internal one from the individual, and the external one from the neighbours. To this end, in this work, we propose simulate a crowd using affective computing with dynamic emotional transmission. Specifically, we use Hidden Markov Model (HMM) to model the emotions for individuals with consideration of personality, and to capture the internal emotion state transfer. Besides, we abstract a two-layer transmission process to quantify the impact from highly active neighbours. In addition, we conduct some simulation experiments to evaluate our proposed model.

*Keywords:* Crowd simulation, Hidden Markov Model, Affective computing

### **Xiangyang Ren, Yu Ren, Qingmei Li** Uncertain demand of farming-enterprise supply chain coordination based on the option contract

*Computer Modelling & New Technologies 2014 18(6) 173-178*

The paper sets up a single-cycle and two-level supply chain model of single rural cooperative and multiple retail enterprises, based on subsidies of option price to combined contract, with the profits maximization of the whole farming-enterprises' supply chain and each member as the goal. It gives the best order quantity and the profits of the whole supply chain and each member under two circumstances, centralized decision and decentralized decision. It also obtains conditions to achieve supply chain coordination. Finally, the paper verifies the model through an example.

*Keywords:* Farming-enterprise supply chain, Option contract, Price subsidy contract, Supply chain coordination

### **Fei Shao, BingHua Cheng** Traffic driven epidemic spreading in weighted networks with different routes

*Computer Modelling & New Technologies 2014 18(6) 179-185*

How does traffic processes in weighted networks impact on the dynamics of epidemic spreading have attracted increasing attention. It is of great importance to reduce the epidemic spreading velocity and increase the critical epidemic threshold of those real world networks. In this paper, the traffic driven epidemic spreading behaviour in BBV weighted networks is investigated. Formulas to describe the infected density and the epidemic threshold of weighted networks are derived and validated by simulations. The infected density and the epidemic threshold are found to undergo a corresponding change when packets are forwarded through different routes according to the different tuneable parameter. By simulations, the optimal route is explored which is better to control the epidemic spreading.

*Keywords:* weighted network, BBV network, epidemic spreading, traffic flow

**Guozhang Jiang, Chongwu Lei, Honghai Liu, Gongfa Li** Planning and scheduling model of production process in iron and steel enterprise

*Computer Modelling & New Technologies 2014 18 (6) 186-191*

According to the analysis of the characteristics of production process in iron and steel enterprise, production planning and scheduling model of multi-stage hybrid procedure and match of production scheduling mode and time were studied. Then the production planning model of production process in iron and steel enterprise was set up. Simulation analysis of production planning and scheduling was carried out using the production procedure of an iron and steel enterprise as an example. The good simulation results show that the established models are correspond to the actual conditions.

*Keywords:* production planning and scheduling, iron and steel enterprise, production process, multi-stage, simulation

**Gongfa Li, Yikun Zhang, Guozhang Jiang, Honghai Liu, Jia Liu** Production procedure optimization in iron and steel enterprise

*Computer Modelling & New Technologies 2014 18(6) 192-196*

Green manufacturing is an effective way of realizing the sustainable development strategy. From the view of evolution of production chain and goods value, energy-saving, cleaner production and green manufacturing of iron and steel industry is discussed and the importance of system optimization of steel manufacturing process is stressed. Connotation of green manufacturing for the iron and steel was explained, the function of steel production process for green manufacturing was discussed and the content system of implementation of green manufacturing for the iron and steel enterprise was established. Finally, the steel production process was optimized. The function of iron and steel manufacturing procedure are broaden-manufacturing function of steel product, function of energy conversion and function of waste recycling, which will enhance enterprise's competitiveness and sustainability.

*Keywords:* production process optimization, iron and steel enterprise, green manufacturing, resource, environment protection

**Houxing Tang** Spatial effect of knowledge spillover on regional economic development: an empirical study from China

*Computer Modelling & New Technologies 2014 18(6) 197-203*

In order to measure the spatial effect of knowledge spillover on regional economic development, a spatial Durbin model, which contains the dependent variable GDP and independent variables Capital, Labour & Knowledge stock, was constructed based on C-D production function. And then an empirical study with 31 provinces of mainland China from year 2000 to 2011 was conducted. The results show that, firstly, the Capital, Labour and Knowledge-stock all have significant positive correlation with GDP. In other words, these three factors have an important impact on the local regional economic development, but the effect of Capital is the greatest and the Labour & Knowledge-stock follow suit. Secondly, the Capital and Labour have a negative spillover effect, but the Knowledge-stock has much more positive effect. Consequently, the governments of developing regions should make full use of the spatial effect of knowledge spillover from developed regions to promote the economic restructuring and great-leap-forward development, especially when they are lack of sufficient funds to support local R&D activities.

*Keywords:* spatial effect, knowledge spillover, knowledge sharing, regional economic development, spatial durbin model

**Xing Yu** A fuzzy set approach for a multi-period optimal portfolio selection model

*Computer Modelling & New Technologies 2014 18(6) 204-207*

Due to portfolio decision deals with future events and opportunities, the market information is uncertain. This paper aims to propose a fuzzy multi-period portfolio selection model to hedge against the uncertainty. A new transformation method based on qualitative possibility theory is developed to transfer the model to a crisp programming, which can be solved by an optimization technique. An example is used to illustrate our approach.

*Keywords:* Multi-period portfolio selection, Fuzzy sets, Optimization, Qualitative possibility theory

**Chonghuan Xu** A novel method of user interest drift detection engaging in individual background factors*Computer Modelling & New Technologies 2014 18(6) 208-212*

Personalized service tends to be an emerging challenge in the field of interest mining on e-commerce platform, the issues of which include how to integrate the user's individual background factor, to hiddenly attain portal user interest behaviour, and to mine interest drift pattern. According to user interest drift problem of personalized service in network, this paper explains the user interest through an integration of individual background factor, user behaviour and interest. Meanwhile, it recommends the fuzzy logic thought to explain its impact factor weights comprehensively in order to reflect the level of the user interest on theme. And, it establishes the Hidden semi-Markov Model via user browsing path to detect whether the interest is drifted or not. Finally, the method is proved to be accurate through the experiment analysis.

*Keywords:* user interest, HSMM, background factors, drift detection

**Jing An, Bo Xu** Empirical study of C2C logistics customer satisfaction based on AHP and FCE*Computer Modelling & New Technologies 2014 18(6) 213-217*

Evaluation system of C2C logistics customer satisfaction was mainly studied here. Firstly, three primary indexes and nine secondary indexes were constructed as the components of the evaluation system. Then the integrated use of AHP and FCE was carried on taobao.com as the empirical research. The result indicates that the overall logistics customer satisfaction of taobao.com is good. But there is still space for improvement. Through this study, logistics customer satisfaction of taobao.com is to improved. In addition, references are provided for other C2C enterprises and logistics companies. The common development of them is therefore promoted.

*Keywords:* C2C logistics customer satisfaction, AHP, FCE

**Bin Wang, Dashe Li, Shue Liu** Research and development of comprehensive communication experiment teaching system*Computer Modelling & New Technologies 2014 18(6) 218-224*

This paper has developed a set of management and experiment system for communication laboratories. Its development platform is Visual Basic 6.0, using access database, winsocket programming and multicast technology. Including several sets of software and using C/S architecture, the system can jointly work with data network configuration system so as to achieve the communication experiment of data network. It can conduct VLAN isolation and IP filter using right management switch, which can make multiple servers connect to the device simultaneously and completely control the numbers of computers entering the system at some point. By off-line configuration technology and shared online database technology, the system can make the device resources be assigned automatically to solve the basic problems of many people doing experiment at the same time.

*Keywords:* comprehensive communication experiment teaching system, access, winsocket, multicast

**Zhenfang He** Artificial neural network model of forecasting relative humidity in different humid and arid areas of China*Computer Modelling & New Technologies 2014 18(6) 225-232*

The objective of the present study is to build different models forecasting the daily mean relative humidity (MRH) values in China with the help of the meteorological parameters. A back-propagation artificial neural network (BPANN) models was employed to identify the relationship between meteorological factors and the relative humidity in China. Weather data 1-day lag was the input layer variables, including (1) the highest atmospheric pressure, (2) the lowest atmospheric pressure, (3) the average atmospheric pressure, (4) the average temperature, (5) the highest temperature, (6) the lowest temperature, (7) precipitation, (8) the average wind speed, (9) the maximum wind speed (the average wind speed over 10 minutes), (10) the utmost wind speed, (11) hours of sunlight, (12) the relative humidity. Experimental results: in the validation period for 1-day lead, the comparison of the prediction performance efficiency of the BPANN models indicated that the BPANN models with trainbr algorithm was superior to the remaining two ones (trainlm and traingdx) in forecasting the relative humidity time series in term of correlation coefficient (R). During the training and

testing periods for 1-day lead, the best performance for the given problem was arid area, followed by semi-arid area, semi-humid area, and humid area respectively. The possible cause for the results was that the impact of these factors on the relative humidity in arid area was the largest, followed by semi-arid area, semi-humid area, and humid area, respectively. From the prediction results of MRHextrema, humid area was the first; semi-arid area was the second; semi-humid area was the third; and arid area was the fourth. From the prediction results of MRHextrema, trainbr algorithm was the best in arid area, semi-humid area, and humid area; but trainlm was the best in semi-arid area. So trainbr algorithm was further employed to predict MRH for 2, 3 or 4-day lead at Urumqi City. From the training and testing effects, 1-day lead was the best, followed by 2, 3 or 4-day lead respectively. In the prediction results of MRHextrema, the best was 2-day lead; the second was 3-day lead; the third was 1-day lead; and the fourth was 4-day lead. The BPANN model results will assist researchers determining meteorological parameters to forecast MRH.

*Keywords:* Meteorological Parameters, Humid and Arid Areas, Artificial Neural Network Model, Relative Humidity, Training Algorithms

#### **Yulan Zhou** Dynamic pricing model of monopolistic manufacture based on the after-sale service

*Computer Modelling & New Technologies 2014 18(6) 233-239*

Dynamic pricing is concerned by business and academia as a pricing method, and has also made extensive research in this field. But the dynamic pricing theory with multi period is not mature considering monopolistic environment and after-sale service of manufacture. Because consumer not only pays attention to the product itself, but more emphasis on after-sale service of product with the changes of consumption concept and increasingly fierce market competition. Therefore good after-sale service is an important reason for consumers to purchase repetitively, which has become the key to the success of manufacture. This paper puts forward to the demand function with learning character, and constructs the multi-period dynamic pricing model on account of monopolistic manufacture and after-sale service level. Then it has important theoretical and practical significance when the conclusions are applied to the monopolistic manufacture. The research findings show the product price of manufacture fluctuates with oscillation both in the short and long term. But it is gradually reduced to a constant value for the magnitude or extent of the price oscillation with certain rate convergence in the long term. Finally, the price may tend to consumer's reservation price or unit operating cost of manufacture.

*Keywords:* Dynamic Pricing Model, After-sale Service, Monopolistic Manufacture

#### **Jian Jin, Jianxiang Wang** Study on local government public expenditure and multi-factor productivity in china based on instrument variable model

*Computer Modelling & New Technologies 2014 18(6) 240-246*

Based on the conventional C-D Production Function Model, this paper adopted Instrument Variable Model to measure the multifactor productivity growth of 223 cities at prefecture level and above in China, and probed into its relationship with local government public expenditure. It is shown that relationship between total public expenditure of local government and city multifactor productivity growth in China is significantly negative, which does not mean that local government public expenditure in China is inefficient, but because a considerable part of it is put into social security, health and medical care, and other public services. Further research by different productivity levels show that the faster productivity grows, the more deeply market-driven is the economics, the weaker is the negative correlation of local government public expenditure and productivity growth. Science & technology and educational expenditure of local government positively affect multifactor productivity growth in China cities significantly, however in varying degrees.

*Keywords:* instrument variable model, local government public expenditure, multi-factor productivity, decision reference

#### **Qihong Zheng, Liangrong Song** Study on pre-loan evolutionary stable strategy of bank-enterprise for preventing moral hazard

*Computer Modelling & New Technologies 2014 18(6) 247-254*

Based on the information asymmetry between enterprises and banks at the stage of loan application, a replicator dynamic model of bank-enterprise evolution at the pre-loan stage was established and analysed by using the

evolutionary game theory (EGT) and the stability theory of nonlinear differential equation. A numerical simulation was also performed in details, which displayed intuitively how banks and enterprises achieved stable cooperation through long-term evolution. The results showed that, to effectively prevent pre-loan moral hazard, it was vitally important for commercial banks to improve their screening ability, increase disguised costs of enterprises, and formulate proper sanctions and appropriate amounts of penalty as per the local loan atmosphere also with the profit which the loan investment projects will make.

*Keywords:* information asymmetry, evolutionary game, duplicator dynamic model, decision-making

**Yuhui Xu, Qiuyue Luo** Study on the city planning for geological disasters defence based on the model of safe city planning

*Computer Modelling & New Technologies 2014 18(6) 255-260*

It is difficult to take geological disasters defence into city planning, which can greatly prevent lands for construction from geological disasters and ensure the rationality, safety and high-efficiency of land-use. Based on the model of safe city planning and discussing from the angle of safe land arrangement, this paper proposes the system of the city planning for geological disasters defence. It focuses on the defence of mutant geological disasters and takes the geological hazards assessment as the foundation. It makes the detailed regulations of the “specific control and management”, which guarantees both the defence and the control of disasters from technology aspect. Besides, it puts forward methods with highly couple of different measures to defend different disasters in city planning system which can enhance the operability to answer the reality.

*Keywords:* city planning for geological disasters defence, safe city planning, mutant geological disasters, specific control and management, city planning system

**V Grekul, L Baydalina** A forecasting model of the economic efficiency of data centre construction project

*Computer Modelling & New Technologies 2014 18(6) 261-270*

Today, a data centre is the main instrument for providing flexible, scalable IT services to business on the basis of distributed or cloud computing technologies. Building a data centre is always expensive and resource-intensive project. Therefore, during the development of the concept of this project it is extremely important to estimate accurately its economic efficiency. This article represents the model for the analysis of the effectiveness of investment in the data centre construction. Our model comprises several regressions that show correlations between main characteristics of the project (capital and operational expenditures, Net Present Value) and parameters of the data centre under construction (data centre area and the number of racks). The model is based on the results of the analysis of the current state and trends in the data centre market in Russia.

*Keywords:* datacentre, price, project, CAPEX, OPEX

## NATURE PHENOMENA AND INNOVATIVE ENGINEERING

**Hongyan Li** An improved energy-efficient distributed storage system

*COMPUTER MODELLING & NEW TECHNOLOGIES 2014 18(6) 271-280*

Energy consumption has increasingly become a serious problem in contemporary data centres. The electricity bill contributes a significant fraction of the Total Cost of Ownership (TCO), and it is predicted to increase at an even faster pace in the following years as extremely large volume of data are being generated on a daily basis which would necessitate corresponding storage capacity to hold them. As a profitable work-around step toward the energy problem within the cloud infrastructure, in this paper, we propose REST, an energy-efficient cloud storage, which is built upon a cluster-based object store similar to GFS. It achieves high energy-efficiency by cleverly exploiting the redundancy already present in the system without compromising the inherently well-established schemes for consistency, fault-tolerance, reliability, availability, etc., while maintaining a reasonable performance level. By modifying slightly the data-layout policy, REST can safely keep a large amount of the storage nodes in standby mode or even powered off entirely most of the time. Deploying a sophisticated monitor, it also provides the flexibility to power up sleeping or powered down nodes when necessary to accommodate to the variations in workloads. Trade-offs between energy

efficiency and performance can be conveniently made by simply adjusting a trade-off metric in REST. The FileBench and real world workload experimental results demonstrate that power savings can reach 29% and 33%, respectively, while still providing comparable or even surprisingly better performance.

*Keywords:* cloud storage, data centre, distributed storage, energy efficient, power saving

**Zhijun Zhang, Hong Liu** Research on ontology-based literature retrieval model

*Computer Modelling & New Technologies 2014 18(6) 281-289*

Proper understanding of textual data requires the exploitation and integration of unstructured and heterogeneous scientific literature, which are fundamental aspects in literature retrieval research. The traditional literature retrieval is based on keyword matching, and the retrieval results often deviating from the users' needs. In this paper from the perspective of ontology, we built shareable and relatively perfect medical enzyme ontology, which is the foundation of the study of domain ontology constructing method. The ontology-based full text retrieval algorithm is put forward, and a document retrieval system based on medical enzyme semantics is designed and implemented, which can not only implement intelligent literature retrieval, but also improve the recall significantly while keeping high precision. This system can employ in particular area moreover it can be used in different areas of the semantic retrieval, which can provide intelligent foundation for the expert systems in medical enzymes field, information retrieval and natural language understanding, etc. The experimental results on the public medical enzyme domain dataset show that our approach performs better than the state-of-the-art methods.

*Keywords:* Ontology, Literature Retrieval, Semantic Web, Ontology Construction, similarity computation

**Gongfa Li, Jia Liu, Guozhang Jiang, Honghai Liu, Wentao Xiao** Case-based reasoning intelligent prediction model of rotary kiln temperature

*Computer Modelling & New Technologies 2014 18(6) 290-293*

Temperature is a key technical index in rotary kiln combustion process, which is so difficult to measure directly online. The offline analysis has large-time delay and poor precision. An intelligent prediction model of rotary kiln temperature based on case-based reasoning was developed, which consists of four modules: data collection and pre-treatment, prediction, online modification and effect estimate. The practical data of some rotary kiln were simulated. The industrial application results show that the prediction model can reflect the actual operation condition and meet the requirement of real-time control. Its effectiveness is proved evidently.

*Keywords:* case-based reasoning, intelligent prediction, temperature control, rotary kiln

**Chengbing He, Dongchao Chen** Study on subsynchronous resonance problem in series-compensated transmission system

*Computer Modelling & New Technologies 2014 18(6) 294-298*

The series capacitor compensation technique can improve the transmission capability of transmission system; however, it may cause subsynchronous resonance (SSR) which will seriously influence the safe operation of large turbo-generator shafts. Aiming at a series capacitor compensation transmission system, the transient torques of SSR is simulated in different compensation degrees and unit load levels. Based on the modified segmentation Prony algorithm (MSPA) proposed in this paper, the characteristic frequencies and damping parameters are identified and analysed. The MSPA, complex Morlet wavelet and time-frequency contour map are integrated and used in analysis of transient torques of SSR in time-frequency field. The results of study show that the transient torques develop towards divergence direction and the frequencies of SSR (converted to the rotor side) in negative damping area move towards into low frequency area with the increasing of the compensation degree. With the increasing of the power output, the torque values of each shaft section will increase, but the dominant oscillation mode will not change. The increasing of power output can make the damping effect slightly change and the trend of the change develops towards the direction that is not conducive to restrain SSR.

*Keywords:* series-compensated transmission, subsynchronous resonance, modified segmentation Prony algorithm, complex Morlet wavelet, time-frequency contour map

**Ming Lv, Xu Yan, Xin Nie, Huachen Pan, Haiqiang Liu** Experimental studies of differential heating for artificial upwelling*Computer Modelling & New Technologies 2014 18(6) 299-303*

PIV experiments were carried out to study the mechanism and characteristics of a new artificial upwelling technology via differential heating named “Differential- Heating-Liquid-Upwelling” (DHLU). Results show that there is a small scaled area with high temperature around the modelling heat point source in DHLU system. The existence of this high temperature area is suggested as the power source of upwelling. Obvious upwelling flows upon the heating source were observed in DHLU system. The stream tube of upwelling for DHLU system is just like an upside-down cone. The max ascending velocity of horizontal layer increases firstly and decreases then as the height increases. A few of fluid masses with high ascend velocities were observed and the mechanism of the DHLU is revealed.

*Keywords:* artificial upwelling, differential heating, mechanism

**Shengli Liu, Xiang Gao, Pan Xu, Long Liu** Method of multi-feature fusion based on SVM and D-S evidence theory in Trojan detection*Computer Modelling & New Technologies 2014 18(6) 304-310*

According to the low accuracy and low stability of the single feature-based method for Trojan detection, a multi-feature fusion method based on SVM and DS evidence theory is proposed. First, three types of flow features such as session, upload data of session/download data of session, distribution of data packet size are extracted from the data stream. Then the SVM classification results of each single feature are used as evidences to construct the basic probability assigned (BPA). Finally, we use DS combination rule of evidence to achieve the decision fusion and give the final detection results by fusion results. The experimental results showed that the accuracy of multi-feature fusion method was 97.48%, which has good performance on accuracy and stability compared with the single feature method in Trojan detection.

*Keywords:* Trojan detection, support vector machine, DS evidence theory, multi-feature fusion

**Yi Zhang, Xiuxia Yang, Weiwei Zhou** UAV trajectory optimization generation based on Pythagorean hodograph curve*Computer Modelling & New Technologies 2014 18(6) 311-320*

The study of autonomous planning of UAVs (Unmanned Aerial Vehicles) flyable on-line path to adapt unstructured environment and improve manoeuvring warfare capability has an important practical significance. A path planning algorithm on-line of UAV based on Pythagorean Hodograph (PH) curve is put forward, which can consider the kinematics and dynamic constraints. The effect of the key parameters on the trajectory generation are analysed, and the appropriate value range that satisfy the constrains are given. To overcome the blindness iteration, the method of improved estimation of distribution algorithm (EDA) is used to trajectory optimization. In the estimation of distribution algorithm, the probability selection mechanism of global elite individual based on interval selection is raised, which improve the speed and precision of the path generation. When the obstacle is detected by the UAV, only if the position and the direction of the interrupt point is given, the trajectory can be replanned online. Simulation results for UAV trajectory optimization generation prove the validity and practicability of the algorithm.

*Keywords:* unmanned aerial vehicle, PH curve, estimation of distribution algorithm, trajectory generation, backstepping

**Juan Zhang, Hesong Jiang, Hong Jiang, Chunmei Chen** Reliable UDP over the air transfer in digital radio system*Computer Modelling & New Technologies 2014 18(6) 321-329*

Reliability is very important in digital radio point-to-point transmission system, especially for bulk data transfer in narrow band channel. Currently most applications are based on raw UDP service, which does not guarantee the reliability, and existing reliable UDP transfer protocols do not satisfy the performance expectations. The article presents R2UDP(Reduced Reliable UDP) over the air transfer suitable to radio system, SBACK (selective bundled ACK) and smart probe improves the transfer efficiency and saves the bandwidth, and also minimizes the impact of bulk data transfer to other traffic on the shared channel.

*Keywords:* R<sup>2</sup>UDP, bundled ACK, smart probe, digital radio, narrow band channel

**Wenqiang Chen** A genetic algorithm for the vehicle routing optimization problem of logistics park distribution

*Computer Modelling & New Technologies 2014 18(6) 330-334*

The Vehicle Routing Problem of Logistics park distribution (VRPLPD) is an extension of the vehicle routing problem, which deals with simultaneous distribution of goods to customers. With the increasing importance of logistics activities, it is of great theoretical and practical significance to determine efficient and effective vehicle routes for simultaneous delivery activities. The study attempts to propose a genetic algorithm approach to tackle this problem. Numerical example is presented with parameter settings in order to demonstrate the applicability and feasibility of the proposed approach. The simulation is carried out in Simulink package of MATLAB. It is shown that Genetic Algorithms are highly effective in optimizing vehicle routing problem.

*Keywords:* Vehicle routing problem (VRP), Logistics park distribution, Genetic algorithms

**Wen-sheng Xiao, Zhong-yan Liu, Jian Liu, Han-chuan Wu** A novel method for identifying system modal parameters using stabilization diagram

*Computer Modelling & New Technologies 2014 18(6) 335-341*

Modal parameters of a structure are important for system identification. In order to identify modal parameters of a structure more accurately, this paper proposes a parameter identification method combined with stabilization diagram. Stochastic subspace identification (SSI) is a recently developed method for identifying a linear system. Combining SSI and the proposed method can easily confirm system order. However, the proposed method has difficulty in distinguishing spurious modals. Therefore, the proposed method must be revised to ensure that the spurious modal can be detected and the SSI can be used to improve identification accuracy. Finally, a simulation is conducted on a fracturing pump truck, when the damping ratio increases from 10% to 40%, those spurious modals disappear. The results indicate that this method performs precise identification.

*Keywords:* Parameter Identification, Modal Analysis, Stochastic Subspace Identification, Stabilization Diagram, Fracturing Pump Truck

**Yunqi Tao, Dong Liu, Jie Cao, Jiang Xu** Decline in gas pressure influences the deformation and permeability of coal-containing methane

*Computer Modelling & New Technologies 2014 18(6) 342-347*

The development and utilization of coal-bed methane, as an unconventional gas, is not only beneficial to the reduction of environmental pollution caused by fossil fuels, but also conducive to the prevention of disasters during coal mining. In this paper, a dynamic permeability model of coal body is established and discussed by means of experimental tests, which measure the deformation and gas permeability of coal-containing methane in the process of gas pressure reduction under different temperatures. The results show that, when gas pressure decreases, the strain of coal-containing methane increases linearly. With temperature increases, the variation of radial strain decreases. Under the same temperature, the permeability of coal decreases first and then increases again during gas pressure reduction. The changing point of gas pressure is approximately 1.2 MPa in the study. In the initial stage of gas pressure decrease, the radial strain of coal-containing gas has a significant effect on its permeability.

*Keywords:* Gas Pressure, Gas Permeability, Coal Deformation, Dynamic Permeability Model

**Ling Li** Income distribution of the bundled transmission of photovoltaic power plant based on DEA game model

*Computer Modelling & New Technologies 2014 18(6) 348-353*

The insufficient utilization capacity of photovoltaic (PV) power has been considered as the bottleneck for the future development of PV power in China. Nowadays, the bundled transmission mode of PV power is regarded as the most advocated solution by many scholars. Under the current unsound fiscal taxation policies in China, however, participators cannot receive the corresponding policy incentives and financial compensation for the additional contributions in constraining the implementation of the bundled transmission mode of PV power. Based on the basic

theory of the DEA (Data Envelopment Analysis) Game, the allocation model of excessive profit was established. The feasibility of the developed model was verified by means of an actual case study. Arguably, the paper provided a certain theoretical basis for the quantity and practical form of excessive income distribution concerning the bundled transmission mode of PV power, and offered a solution to the income distribution for the bundled transmission mode of PV power.

*Keywords:* DEA Game Model, Photovoltaic Power Utilization, Photovoltaic Power Transmission, Income Distribution

**Zhiqiang Yin, Lei Wang, Haifeng Ma** Optical measurement method for dynamic mechanical testing based on image grey level distribution difference model

*Computer Modelling & New Technologies 2014 18(6) 354-359*

This study developed an optical measuring system with a gray level distribution difference (GLDD) model, and applied the system to examine the displacement field of a Brazilian disk (BD) split under dynamic loading. The system consists of high-speed (HS) photography, a split-Hopkinson pressure bar (SHPB), a synchronization control system and operation of differential images. First, we captured differential images with a high speed camera (10 frames at a time resolution of 10 $\mu$ s). Next, we established the corresponding relationship between the dynamic fracturing evolution of the disc rock samples and the stress loading process with a synchronization controlling system. Changes in the surface displacement field were calculated with the differential image base method according to the joint probability distribution function of two images. This method takes the image correlation into account and can effectively eliminate the influence of background noise, identify surface displacement and capture cracks and expansion in dynamic Brazilian disk splitting experiments straightforwardly and accurately. Findings can be used for novel measurement of surface displacement fields in Brazilian disk splitting tests under dynamic loads.

*Keywords:* grey level distribution difference model, SHPB, dynamic load, high speed photography, differential image

**Fei Cai, Wenjun Wang** Research on illumination invariance colour index algorithm based on colour ratio

*Computer Modelling & New Technologies 2014 18(6) 360-364*

Statistics of colour value of each pixel in the image are output in traditional colour histograms. Therefore, though the two same images photographed in different illuminations are consistent in colour content, they have different colour distributions in the histograms. To solve the problem, this paper introduces an illumination invariance colour index algorithm based on colour ratio. According to the colour constancy theory, although colour values of its pixels will be changed once the image is subject to illumination, colour ratios remain unchanged. Colour ratio refers to the ratio between two contiguous pixels. As per colour ratios, colour ratio image may be obtained, which depicts obvious boundaries or margins of the image content so that we statistics of colour ratio histogram can be obtained as an index mechanism to remove illumination effect. Verified by lots of tests, this method can extract useful colour characteristics and remove illumination effect, so that it can be practically used in effective computer recognition of objects in traffic videos.

*Keywords:* colour index, image retrieval, colour constancy, illumination

**Debao Yuan, Xueqian Hong, Shiwei Yu, Liangjian Li, Yanbao Zhao** Study on feasibility of CORS application in surface movement deformation monitoring in mining areas

*Computer Modelling & New Technologies 2014 18(6) 365-371*

CORS has been widely established in China and abroad, and can be used in geodetic measuring, coordinate system retaining and surveying and mapping of city. However, some issues (stability of base stations, influence of survey precision on horizontal and vertical extension, et al) are still unsolved in Surface Movement Deformation Monitoring in Mining Areas. Based on the CORS, a method and flow for data processing and stability analysis of deformation monitoring network was proposed. Using rank defect free network adjustment, a robust estimation with minimum of first order norm of displacement component was offered to determine displacement of relative stability. The strategy can resolve the confirmation of robust iterative weights and the effects of different reference models. Displacement significance was tested with the normality method. Finally, according to this method, datum of GPS deformation

monitoring in Mining Areas in Datong were calculated and preliminarily proved that this method was feasible and effective, providing a new monitoring methodologies terrain monitoring in mining areas.

*Keywords:* CORS (Continuous Operation Reference System), deformation monitoring, rank defect network adjustment, robust estimation, hypothesis test

### **Jianjun Dong** Study on improved elasto-plastic model for unsaturated soils based on Barcelona Basic Mode

*Computer Modelling & New Technologies 2014 18(6) 375-380*

The strength and deformation behaviours of unsaturated soil can be approximately described by elasto-plastic constitutive model that was proved by abundance academic and test researches. The Barcelona elastic-plastic model is an excellent model that can simulate the strength and deformation of unsaturated soil. But their calculated result of shear strength is low. So an improved BBM model is settled by using drop-shaped shear yield surface and hardening theory of dual stress. The results show that the improved model can more accurately predict the strength and deformation behaviours of unsaturated soil under suction-controlled triaxial compression stress states.

*Keywords:* shear yield surface, drop-shaped, hardening theory, dual stress, unsaturated soil, elasto-plastic

### **Wenyu Lv** Water trickling and roof falling of soft argillaceous roadways with its composite supporting and repairing technology

*Computer Modelling & New Technologies 2014 18(6) 381-387*

Using methods such as the transient electromagnetic method, rock mechanics testing, X-ray diffraction analysis, rock strength weakening experiment, borehole observation technology and grout mix design, etc. to study the mechanism of water trickling and roof falling of soft argillaceous roadways, via model analysis and numerical calculations, we reached many conclusions: Rock breakage is perpendicular to the axis of borehole; expansion of argillaceous swelling rock weakens the surrounding rock, the strength degradation of which occurs in 1h; performance of anchor agent also significantly decreased; uniaxial compressive strength, friction angle and cohesion of surrounding rock decreased by 40%, 16.5% and 11.1%, respectively; water trickling further exacerbated the risk of roof falling in the construction of roadway; according to calculation, the weak plane of the original 12 # I-beam became shear failure, and cannot meet the large initial deformation of roadways, floor heave is large with serious water logging. We presented the technology of “targeted drainage, deep-and-shallow hole grouting, advanced ductile, floor anchor rope, U29-shaped steel, and anchor net spray composite support”. Site monitoring on 1# railway cross-cut of Bofang Coal Mine displayed that the maximum amount of deformation was less than 125mm, working resistance of anchor bolt and cable were 72-91kN and 123.3kN, respectively. Water trickling and roof falling were controlled; water logging of floor heave was improved obviously. Numerical simulations show that plastic zone is greatly reduced, and support effects were fairly good.

*Keywords:* soft and argillaceous, X-ray, water trickling and roof falling, composite support, numerical simulation