

Journal of Computers

ISSN 1796-203X

Volume 7, Number 10, October 2012

Contents

Special Issue: Advances in Information and Computers

Guest Editors: Shifei Ding

Guest Editorial <i>Shifei Ding</i>	2351
SPECIAL ISSUE PAPERS	
Study on Medical Image Processing Technologies Based on DICOM <i>Peijiang Chen</i>	2354
A New Obstacle Avoidance Method Based on Biped Robot <i>Zetao Jiang, Yanru Cui, and Qiang Wang</i>	2362
The PID Controller Based on the Artificial Neural Network and the Differential Evolution Algorithm <i>Wei Lu, Jianhua Yang, and Xiaodong Liu</i>	2368
A Measuring Method for Angular Displacement Based on Correlation Algorithm <i>Ping Yan, Yan He, Runzhong Yi, Fei Liu, and Guorong Chen</i>	2376
Research on Low Power Sigma-Delta Interface Circuit used in Capacitive Micro-accelerometers <i>Yue Ruan, Ying Tang, and Wenji Yao</i>	2383
Non-linear Multi-attribute Based Online Auction Bidding Model and Platform <i>Jie Lin, Jinghua Zhao, and Ling Xue</i>	2390
Virtual Machine-based Intrusion Detection System Framework in Cloud Computing Environment <i>Huaibin Wang, Haiyun Zhou, and Chundong Wang</i>	2397
Scheduling Program of the Track and Field Sports Competition Based on R_Timetable Algorithm <i>Yongxin Wang, Jianying Wang, and Qiufen Wang</i>	2404
Research on a New Kind of Robust Backstepping Filter Derivative Control Method <i>Yue Zhi, Junwei Lei, and Jinyong Yu</i>	2411
Intelligent PID Controller on Soft Computing <i>Honghua Xu, Xiaoqiang Di, Huamin Yang, and Guangcui Cui</i>	2417
An Approach to Interactive Affective Learning Algorithms <i>Chong Su and Hongguang Li</i>	2425
A New Error-Correcting Transmission Method for Dual Ring Fieldbus in CNC System <i>Lei Yang, Hu Lin, Dongfeng Yue, and Tianrong Gao</i>	2432
Numerical Simulation of Heavy Rail Quenching Process <i>Gongfa Li, Yuesheng Gu, Jianyi Kong, Guozhang Jiang, Jintang Yang, Liangxi Xie, Zehao Wu, Siqiang Xu, Yaping Zhao, and Gang Zhao</i>	2439

A Redundant FPGA Based Controller for Subsea Blowout Preventer Stack <i>Zengkai Liu, Yonghong Liu, Baoping Cai, Fei Wang, Zhili Chen, Xiaojie Tian, and Yazhou Wang</i>	2446
A Non-Standard Approach for the OWL Ontologies Checking and Reasoning <i>Yingjie Song, Rong Chen, and Yaqing Liu</i>	2454
Process Goose Queue Methodologies with Applications in Plant-wide Process Optimization <i>Jingwen Huang and Hongguang Li</i>	2462
A New Sub-topic Clustering Method Based on Semi-supervised Learning <i>Xiaodan Xu</i>	2471
Multi-feature Fusion Face Recognition Based on Kernel Discriminate Local Preserve Projection Algorithm under Smart Environment <i>Di Wu, Jie Cao, Jinhua Wang, and Wei Li</i>	2479
Safety Separation Assessment in Free Flight Based on Conflict Area <i>Zhaoning Zhang, Chang Sun, and Peng Zhou</i>	2488
Design of Three-axis ED Milling Machine Based on the PMAC Motion Card <i>Fei Wang, Yonghong Liu, Zhili Chen, Renjie Ji, Xiaojie Tian, and Zengkai Liu</i>	2496
A Case Study of Model Checking Retail Banking System with SPIN <i>Huiling Shi, Wenke Ma, Meihong Yang, and Xinchang Zhang</i>	2503
The Investigation of Fault Diagnosis Based on GA-HPSO-NN <i>Bin Peng, Chang-an Hu, Rong-zhen Zhao, and Yu-qiao Zheng</i>	2511
A Novel Approach to Hardware/Software Partitioning for Reconfigurable Embedded Systems <i>Linbai Cui</i>	2518
A Study on the Control Methods Based on 3-DOF Helicopter Model <i>Junshan Gao, Xinghu Xu, and Chen He</i>	2526
High Level Synthesis using Learning Automata Genetic Algorithm <i>Huijing Yang, Chunying Wang, and Ning Du</i>	2534
Application of Particle Swarm Optimization in Figuring out Non-differentiable Point of Function <i>Yuanbin Mo, Xin-quan Zhao, and Shu-jian Xiang</i>	2542
A Robust Watermarking Against Shearing Based on Improved S-Radon Transformation <i>Minghui Deng, Qingshuang Zeng, and Xiuli Zhou</i>	2549

REGULAR PAPERS

Multiple Model Predictive Control of Component Content in Rare Earth Extraction Process <i>Hui Yang, Rongxiu Lu, Kunpeng Zhang, and Xin Wang</i>	2557
Multi-party Dialogue Games for Dialectical Argumentation <i>Jinpeng Yuan, Li Yao, Zhiyong Hao, and Yanjuan Wang</i>	2564
An Approach of Trustworthiness Evaluation of Software Behavior Based on Multidimensional Fuzzy Attributes <i>Zhen Li and Junfeng Tian</i>	2572
Personalized Web Search Using Clickthrough Data and Web Page Rating <i>Xueping Peng, Zhendong Niu, Sheng Huang, and Yumin Zhao</i>	2578
Study on Learner Modeling in Adaptive Learning System <i>Bing Jia, Yongjian Yang, and Jun Zhang</i>	2585

Multi-satellite Monitoring SST Data Fusion based on the Adaptive Threshold Clustering Algorithm 2593
Hongmei Shi, Han Dong, Lingyu Xu, Cuicui Song, Fei Zhong, and Ruidan Su

Description to Fe-C Alloy Film Fiber Corrosion Sensors by Fractal Corrosion Images 2599
Lihua Cai and Wei Li

Special Issue on Advances in Information and Computers

Guest Editorial

This special issue comprises of 27 selected papers from the 2012 International Conference on Information, Computing and Telecommunications (ICICT 2012). The conferences received 1660 paper submissions from 11 countries and regions, of which 810 were selected for presentation after a rigorous review process. From these 810 research papers, through two rounds of reviewing, the guest editors selected 27 as the best papers on the Information and Computers track of the Conference. The candidates of the Special Issue are all the authors, whose papers have been accepted and presented at the ICICT 2012, with the contents not been published elsewhere before.

Information, Computing and Telecommunications are very hot and important research topics and gain more attention by economy as well as society in recent years. The 2012 International Conference on Information, Computing and Telecommunications (ICICT 2012) was held from Jan 7~8, 2012 in Harbin, P.R. China. This conference is co-sponsored by Harbin University of Science and Technology and International Science and Engineering Research Center, and it is technical co-sponsored by Harbin Engineering University, Northeast Forestry University, Harbin Normal University, Heilongjiang University, Northeast Petroleum University and Harbin University.

“Study on Medical Image Processing Technologies Based on DICOM”, by Peijiang Chen, proposes a general idea of converting DICOM to BMP format, and discusses several medical image processing technologies, such as grayscale processing and edge detection.

“A New Obstacle Avoidance Method Based on Biped Robot”, by Zetao Jiang, Yanru Cui, Qiang Wang, proposes a new obstacle avoidance method for biped robot. The robot could avoid obstacle successfully by crossing the obstacle instead of using traditional method of re-select the path.

“The Artificial Neural Network PID Controller Based on DEA”, by Wei Lu, Jianhua Yang, proposes a new style of PID controller that is based on artificial network and evolutionary algorithm. This new PID controller has more advantages than the traditional one, such as less calculated load, faster global convergence speed, better robust, more independence and adaptability on the plant, etc.

“A Kind of Angular Displacement Measuring Method Based on Correlation Algorithm”, by Ping Yan, Yan He, Runzhong Yi, Fei Liu, Guorong Chen, proposes a kind of angular displacement measuring method based on correlation algorithm is presented with the characteristics of a low manufacturing cost, high precision, anti-noise and anti-partially damaged properties and so on.

“Research on Low Power Sigma-Delta Interface Circuit used in Capacitive Micro-accelerometers”, by Yue Ruan, Ying Tang and Wenji Yao, proposes the chip-level design and implementation of a second order sigma-delta interface circuit used in capacitive micro-accelerometers. The chip layout is derived and parameters such as area and power consumption are shown.

“Non-linear Multi-attribute Based Online Procurement Auction Model and Platform”, by Jie Lin, Jinghua Zhao, Ling Xue, proposes a auction model of multi-attribute online Procurement, which can measure the effect of the attributes in auctions and evaluate suppliers in the procurement managements by normalizing weighted sum. In the end, an application platform of non-linear multi-attribute online procurement auction has been built.

“Virtual Machine-based Intrusion Detection System Framework in Cloud Computing Environment”, by Huaibin Wang, Haiyun Zhou, Chundong Wang, proposes a model detecting the malign intrusion in clouding computing environment. Virtual machine-based architecture, the concept of cloud appliance and the authentication among cooperation agents, all of them proposed realize the intrusion detection effectively and have the hige detection rate and low false alarm rate.

“Scheduling Program of the Track and Field Sports Competition Based on R_Timetable Algorithm”, by Yongxin Wang, Jianying Wang, Qiufen Wang, proposes an automation method based on R-timetable algorithm, which is achieved effectively through the C++ program design, to solve the corresponding problems. The method immensely optimizes the organization and improves the efficiency of arrangement of the sports competition.

“Research on a New Kind of Robust Backstepping Filter Derivative Control Method”, by Yue Zhi, Junwei. Lei and Jinyong Yu, proposes a backstepping design technology which is perfectly integrated with the PID control method . Meanwhile, the relationship between Lyapunov function and transfer function is established, which is an important concept that can be applied in a large family of control systems.

“Intelligent PID Controller on Soft Computing”, by Honghua Xu, Xiaoqiang Di, Huamin Yang, Guangcai Cui, proposes PID control algorithm based on soft computing techniques. With fuzzy system, neural network technique and genetic algorithm, PID controller realizes intelligent self-tuning of controller parameters and better controlling quality.

“An Approach to Interactive Affective Learning Algorithms”, by Chong Su, Hongguang Li, proposes affective computing models (STAM) and affective learning strategies which can gradually grasp essentials in human’s subjective

judgments in decision-making and reduce human's subjective fatigue so as to make the decisions more objective and scientific.

"A New Error-Correcting Transmission Method for Dual Ring Fieldbus in CNC System", by Lei Yang, Hu Lin, Dongfeng Yue, Tianrong Gao, proposes a new transmission scheme based on the Dual Ring Fieldbus in the CNC system, which involves both error detection and self-correction method. The related data format and transmitting structure are also designed.

"Numerical Simulation of Heavy Rail Quenching Process", by Gongfa Li, Jianyi Kong, Guozhang Jiang, Jintang Yang, Liangxi Xie, Yaping Zhao, Gang Zhao, Siqiang Xu, proposes a numerical simulation of temperature field, stress field and various kinds of factors that may affect temperature and stress field distribution. A new quenching method is proposed, the application result indicates that it's effective.

"A Redundant FPGA Based Controller for Subsea Blowout Preventer Stack", by Zengkai Liu, Yonghong Liu, Baoping Cai, Fei Wang, Zhili Chen, Xiaojie Tian, Yazhou Wang, proposes a novel redundant controller based on Field Programmable Gate Arrays for the control system of subsea Blowout Preventers, using triple modular redundancy and multiprocessor system techniques.

"A Non-Standard Approach for the OWL Ontologies Checking and Reasoning", by Yingjie Song, Rong Chen, proposes an approach for OWL ontologies checking and reasoning. With this method, we can find out the inconsistencies of the ontologies and detect the reasons.

"Process Goose Queue Methodologies with Applications in Plant-wide Process Optimization", by Jingwen Huang, Hongguang Li, proposes a novel ideal to implement plant-wide process optimization by decomposing plant-wide process into multi-layer PGQs, where optimization issue is identical with the following and tracking mechanism of goose queue.

"A New Sub-topics Clustering Method Based on Semi-supervised Learning", by Xiaodan Xu, proposes a new method for sub-topic clustering of multi-document based on semi-supervised learning, which can improve the accuracy of clustering by getting the best value of k according to the characteristics of the data.

"Multi-feature Fusion Face Recognition based on Kernel Discriminate Local Preserve Projection Algorithm under Smart Environment", by Di Wu, Jie Cao, Jinhua Wang, Wei Li propose a new face recognition method based on kernel discriminative local preserve projection and Multi-feature fusion under smart environment, The experiments on the AMI database indicate the proposed method can enhance the accuracy of the recognition system effectively.

"Safety Separation Assessment in Free Flight Based on Conflict Area", by Zhaoning Zhang, Chang Sun, Peng Zhou, proposes the collision risk model in free flight based on conflict area, which considers the influence of CNS performance to the probability of overlap and introduces the probability of the failure of controller's monitoring.

"Design of Three-axis ED Milling Machine Based on the PMAC Motion Card", by Fei Wang, Yonghong Liu, Zhili Chen, Renjie Ji, Xiaojie Tian, Zengkai Liu, proposes the design of three-axis electrical discharge milling machine including the motion system, control system, working fluid supplying system and EDM pulse generator.

"A Case Study of Model Checking Retail Banking System with SPIN", by Huiling Shi, Wenke Ma, Meihong Yang, Xinchang Zhang, proposes a case study of model checking the business flow of retail banking System , including a specific approach to effectively abstract Retail Banking System to build the PROMELA model, the system properties definition, and the analysis of actual verification with SPIN."

"The Investigation of Fault Diagnosis Based on GA-HPSO-NN", by Bin Peng, Chang-an Hu, Rong-zhen Zhao, proposes a novel algorithm GA-HPSO combining with the advantages of genetic algorithm, simulated annealing and particle swarm optimization. The novel algorithm is applied to train neural network using proportion and blend methods.

"A Novel Approach to Hardware/Software Partitioning for Reconfigurable Embedded Systems", by Linhai Cui, proposes a new approach to Hardware/Software partitioning, including an innovative partition algorithm and a system task schedule algorithm by combining the genetic algorithm and the Tabu search algorithm together.

"A Study on the Control Methods Based on 3-DOF Helicopter Model", by Junshan Gao, Xinghu Xu, Chen He, proposes methods to analysis and design 3-DOF helicopter system, such as how to establish mathematical model, how to design PID controller and fuzzy-PID controller for the Helicopter system, and analyzed the performance of systems with two controllers.

"High Level Synthesis using Learning Automata Genetic Algorithm", by Huijing Yang, Chunying Wang, Ning Du, proposes a new method Learning Automata Genetic Algorithm is used in high level synthesis to deal with scheduling and allocation simultaneously. It can produce area and performance optimized designs.

"Application of Particle Swarm Optimization in Figuring out Non-differentiable Point of Function", by Mo Yuanbin, Zhao Xinquan, Xiang Shujian, proposes an algorithm to figure out the non-differentiable point of function by PSO. The result of numerical calculation testifies the validity of the algorithm.

"A Robust Watermarking Algorithm Based on Improved S-Radon Transformation", by Deng Minghui, Zeng Qingshuang, Zhou Xiuli, proposes a robust image watermarking algorithm in improved S-Radon transformation domain and the chirp signals are used as watermarks which are invisible and robust against geometric distortion.

We would like to take this opportunity to thank the authors for the efforts they put in the preparation of the manuscripts and for their valuable contributions. We wish to express our deepest gratitude to the program committee members for their help in selecting papers for this issue and especially the referees of the selected papers for their thorough reviews under a tight time schedule. Last, but not least, our thanks go to the editorial board of the "Journal of Computers" for the exceptional effort they did throughout this process.

In closing, we sincerely hope that you will enjoy reading this special issue.

Guest Editor:

Shifei Ding, China University of Mining and Technology, China



Prof. Shifei Ding was born in Qingdao, China, in 1963, received his bachelor's degree and master's degree from Qufu Normal University in 1987 and 1998 respectively. He received his Ph.D degree from Shandong University of Science and Technology in 2004. He received Postdoctoral degree from Key Laboratory of Intelligent Information Processing (IIP), Institute of Computing Technology (ICT), Chinese Academy of Sciences (CAS), and his advisor was Professor Z.Z. Shi. And now, he works in China University of Minig and Technology as a professor and Ph.D supervisor. His research interests include Artificial Intelligence, Pattern Recognition, Machine Learning, Data Mining, and Granular Computing et al. He has published 5 books and more than 80 papers in international conferences and journals.

Study on Medical Image Processing Technologies Based on DICOM

Peijiang Chen

School of Automobile, Linyi University, Linyi, Shandong, China

Email: chenpeijiang@163.com

Abstract—DICOM is an international standard for the storage and transmission of medical image. With the popularity of pictorial and computerized medical equipments and the development of hospital management information system, the standard is widely used. The technologies of medical image display and processing based on DICOM standard are studied. On the basis of analyzing the DICOM standards and file formats, the general idea of converting between the DICOM format and BMP format is brought forward, and the medical images can be displayed in the windows platform. The grayscale processing technologies of medical image are focused on and implemented by programming. The main methods of edge detection are discussed and the implementation steps are given. The software of DICOM medical image processing is realized by Visual C++ which can convert the medical images to BMP formats and display these medical images. The grayscale processing, anti-color, strength testing and other basic functions of medial images can be come true. It can provide convenience for medical diagnostics.

Index Terms—DICOM, Medical image, Image Processing, Grayscale processing, Edge detection

I. INTRODUCTION

With the improvement of medical devices level, most hospitals have been already equipped with a variety of digital imaging equipments, and, PACS, picture archiving and transmission system, has been established. PACS is very important in modern hospital, the key technical problem needed to be solved is to unify the image data formats and data transmission standards of all various digital imaging devices. DICOM is generated for this purpose.

DICOM is the standard used for the storage and transmission of medical images which can provide the interface standards and protocols for the manufacturers and users of the medical imaging equipments[1]. With the extensive use of PACS in the hospital, the understanding of the DICOM standard has become increasingly important, and physicians make more and more needs for the post-processing of the medical images. Then, the interpretation of DICOM medical image files, the reading of medical image data, the display and processing of image processing are very important.

In this paper, display and processing of the medical image based on DICOM format are focused on, and they are realized by software programming. The system can

provide convenience and basis for medical diagnosis and remote consultation.

II. DICOM STANDARDS

DICOM is a standard name specifically for medical image storage and transmission established by American College of Radiology (ACR) and National Electrical Manufacturers Association (NEMA). After the development of many years, it has been widely accepted by medical device manufacturers and medical community and is very popular in medical equipments.

A. Target of DICOM Standard

DICOM standard target is to promote medical imaging equipment interoperate in the multi-vendor environment, which can be reflected in the following aspects specifically.

- (1) To promote network of digital imaging devices, regardless of the developers of equipment.
- (2) To help to develop and promote the PACS, and link to other medical information systems.
- (3) To establish valuable diagnostic information database, handle requests of different devices dispersed geographically[2].

B. Features of DICOM Standard

The nowadays widely used standard is DICOM 3.0, it has the following characteristics.

- (1) Be appropriate for network environment
The early DICOM versions were only used for point to point data transmission, but DICOM3.0 supports network environment based on the OSI, TCP/IP and other common industry standards, so as to create the conditions for telemedicine.
- (2) Respond to data exchange and related instructions
The older versions are confined only to data transfer, but DICOM3.0 specifies and defines the semantics of the instructions and data by using the concepts of service classes.
- (3) Define the standard level
The former DICOM can only provide the minimum requirements of the DICOM standard that the medical device should follow, and DICOM3.0 clearly describes some necessary conformance statements to achieve a particular level.
- (4) Scalability
DICOM3.0 supports the dilation of new features.

(5) Introduce the concept of generalized information object

Information objects include not only graphics and images, but also studying, reporting and other general information objects.

(6) Establish a unique method identifying a variety of information objects

It is very important to clearly define the relationship between information objects in the network environment.

C. Outline of DICOM Standard Contents

The contents of medical image communication standard are described in 15 chapters. The main contents of every chapter are as follows.

(1) Introduction and overview

It briefly introduces the scope of the DICOM standard relating, points out the background of standard drafting and its meaning and purpose, and lists the other standards which DICOM standard references.

(2) Conformance

The chapter 2 describes that DICOM standard is a multi-level specification and a professional standard that can be referred to follow. Any Implementation of the standard in engineering can be achieved in some part or all specifications in the product. The product statement must clarify the extent of compliance with the product standard. Conformance states the content requirements and general format of the product complying with the DICOM standard.

(3) Information defining object

This chapter detailedly gives the composition of information objects and concept connotation abstracted from the medical reality of the DICOM standard. It is an important part in studying the standard and in clinical diagnosis.

(4) Service specification class

In this chapter, the standard divides the communication types which may be involved into several detailed equivalent categories in image communication, regulates and defines them.

(5) Data structure and coding

DICOM standard provides dedicated medical image files. The formats are regulated specifically in this chapter, the data used and the character encoding requirements are provided.

(6) Data dictionary

It includes all the encoding and coding instructions of data elements in the DICOM standard. The standard uses unique identifier, which is unique in all international standards, makes DICOM more workable and reduces the conflict.

(7) Message exchange

The chapter 7 defines the method of operation information objects in network communication, encapsulates the data format of information, that is, DIMSE, DICOM Message Service Element.

(8) Network communication support for message exchange

It regulates the message exchange in communication network and defines the transmitted network data and the

state transition in the OSI reference model of ISO and TCP/IP-based protocol stack.

(9) Point-to-point communication support for message exchange

The details provided in this chapter are widely used in the era when the Internet is not widespread, but it is used very less now.

(10) Media storage and file format for data exchange

The chapter 10 specifies the model and functional division of medium storing DICOM files, defines the DICOM file format, and standardizes the logic relation and encoding format of DICOMDIR catalog files.

(11) Media storage application profile

This part provides the application selection mechanism and strategies in media storage, gives the general pattern of the mechanism and some examples.

(12) Storage functions and media Formats in data interchange

The main contain includes the data formants and directory and management specifications in storage process. The more key is the DICOMDIR file which plays an important role in the directory classification of the images.

(13) Support printing management of point-to-point communication

It detailedly describes the composition and conversion of protocol data in the communication process with the printer supporting the DICOM standard.

(14) Grayscale standard display function

It has a very important basis to show parameters, print film, and display image, and can directly affect the accuracy of clinical diagnosis.

(15) Security profiles

The DICOM standard contains very detailed personal information of patients. The data security in the processes of storage and transmission is critical, and the implementation of encrypting transmission of information is necessary. In this chapter, the DICOM standard working group cites the widely used information encryption standard to promote the transmission and storage more reliable and secure[3].

III. COMPOSITION OF DICOM FILE

A. File Structure

DICOM protocol allows that we store the transmission results of data as DICOM. The typical DICOM file structure is shown in Fig. 1.

DICOM file consists of the following components.

(1) Preamble

It has 128 bytes and the file instructions can be set in this part.

(2) Prefix

The four bytes define four characters, D, I, C, M.

(3) Data element

It generally has multiple sets of data element. Each data element corresponding to an IOD attribute has four domains, namely, Tag, Value Representation, Value Length and Value Field, in which, Value Representation is optional[4].

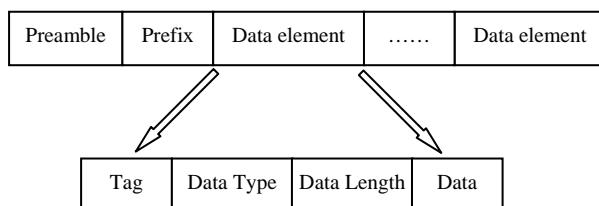


Figure 1. Typical DICOM file structure

B. Image Encoding

Pixel data element (7FE0, 0010) is the most important data unit of DICOM file, and it contains the necessary data for medical image display. The other data elements closely relating to pixel data element are as follows.

- (0028, 0008): number of image frames;
- (0028, 0010): number of image lines;
- (0028, 0011): number of image columns;
- (0028, 0100): distribution bit number;
- (0028, 0101): storage bit number;
- (0028, 0102): highest bit number.

The code of pixel data is determined by distribution bit number, storage bit number and the highest bit number, and the distribution bit number must be greater than the storage bit number. The pixel data of DICOM image is often 16-bit or 12-bit. If 16-bit format is selected, each pixel has two bytes; if 12-bit, the bytes distribution of per pixel is more complex, and we can determine the distribution bit number, storage bit number and the highest bit number by distinguishing the elements values of (0028,0100), (0028,0101) and (0028,0102).

When there are 16 distributed bits, 12 storage bits and the highest bit is 12 in each pixel, the pixel occupies only 2 bytes and uses the low 12 bits. When the highest bit is 12, each pixel uses the high 12 bits of the 2 bytes. When there are 12 distributed bits, 12 storage bits and the highest bit is 11 in each pixel, the pixel only occupies 3 bytes, the contents of the middle byte is divided into two parts, respectively, belonging to the front and back byte, then each pixel has 12 bits.

Pixel data can be compressed or not, when the data is transmitted as compressed format, the Value Representation is OB, on the contrary, the value is OW. For the uncompressed pixel data, the sequence is usually from top to bottom, from left to right and the data are encoded and stored as a continuous bit stream. For the compressed pixel data, they can be stored in segment and delimited by a series of bound items to support the image compression process with un-pre-known length.

IV. FORMAT CONVERSION AND DISPLAY OF DICOM IMAGE

DICOM standard image is a special format and it has complex types, various combination formats. Its display and processing need specially developed image processor, but the most current applications do not support it. Combining the clinical use, we find that physicians usually simply select the specific pieces of image as the

basis for conclusions in the disease diagnosis. The constituting characteristics of the image pixel data have many similarities with the common images, such as BMP, but also, the latter is suitable for image analysis, feature extraction and other image processing. On the other hand, there are still some medical images using BMP format to exchange information and diagnosis in the process of the digital hospital information construction. In order to facilitate the exchange of documents, we achieve the file conversion between the DICOM files and BMP files, the popular bitmap file in Windows system.

A. BMP File Format

BMP file consists of four parts, bitmap file header (Bitmap-file), bitmap information header (Bitmap-information), color table or palette, image data array.

Bitmap file header contains the file type, file size, storage location and other information, and it is defined by the structure BITMAPFILEHEADER in Windows, the length of this structure is fixed for 14 bytes.

Bitmap information header, the structure BITMAPINFOHEADER is a fixed length for 40 bytes.

Palette is optional. If the image has palette, it is actually an array and creates the corresponding relation between array and colors, the value is the number of colors used for the bitmap. The type of each element in the array is a RGBQUAD structure, accounting 4 bytes, it is defined as follows: 1 byte is used for the blue component, 1 byte for the green component, 1 byte for the red component, 1 byte for padding (set to 0). The image data are set after the palette[5].

If the image has no palette, the image data are after the BITMAPINFOHEADER structure.

B. Difference between the Two Image Formats

The file headers and data structures are very different in DICOM image and BMP image. DICOM image stores a lot of medical information of data elements in the data collection, such as patient name, age, hospital name, imaging time, checking site, and so on, besides the image size, height, width, number of bytes per pixel image, and other essential information.

The image data array of the two images are quite different, DICOM images are stored sequentially, the first byte of the array represents the upper left image pixels, and the last byte represents the lower right image pixels. But, BMP image is stored from bottom to top, that is, the first byte of the array indicates the lower left image pixels, and the last byte indicates the top right image pixels[6].

C. Conversion from BMP Image to DICOM Image

Comparing the BMP image with the standard DICOM image format, we can see that the BMP image includes only the characteristics of the object image corresponding to the image information in DICOM image, lacks of a range of information and other features which need be added to follow up. We can manually add them in the process of capture, can also obtain them from the patient information database. But, some characteristics of image IOD information are required to obtain from the imaging equipment, such as the current window width and

position, pixel space ratio[7]. Thus, the software programming process of converting the BMP file into DICOM file is shown in Fig. 2.

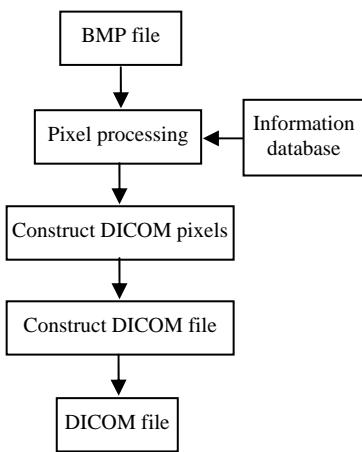


Figure 2. Process of converting BMP file to DICOM file

D. Conversion from DICOM Images to BMP Image

Firstly, we should define some parameter variables of image stored, and initialize the variables. Secondly, according to the requirements of transferring syntax, it traverses the relevant data elements of the DICOM files, then extracts the useful data elements from 0002, 0028, 7FE0 groups, respectively stores them into the previously defined variables, and closes DICOM files. It converses 12-bit or 16-bit image data into 8-bit grayscale data. Finally, it opens a blank document. In accordance with the requirements of DICOM files, the data extracted from DICOM files will be written to the new file.

The conversion process can be divided into three parts. Part one is document reading, including the basic parameters information of image and image data reading. Part two is the shift, capture, window level transformation of image data, etc.. Part three writes 8-bit grayscale image data, bitmap file header, information header and color table into the BMP bitmap file.

The process is shown in Fig. 3.

E. Problems Needed Attention

(1) To obtain image information is actually the traversal of all the data elements of DICOM file. When traversing the data elements, some information contained in data elements is irrelevant to the image, in order to improve the speed of program traversal, we can only read the useful data elements.

(2) When reading the data element (0002, xxxx), the Value Representations of all the data elements are represented, that is, Explicit VR LittleEndian. As the little-endian byte order, the byte orders should be first exchanged. We should pay special attention to the data element (0002, 0010), and the value of the data element can determine the file transfer syntax[8].

(3) The image data of (7FE0, 0010) in pixel data element is generally 16-bit or 12-bit, we need adjust the window width and window bit of the original data into 8-bit grayscale data. The so-called window width is the

range of the image data display, window bit is the center value of the image data. We can adjust them according to the following equation.

$$h = \begin{cases} 0, & b - w/2 > x \\ 255, & b + w/2 < x \\ [x - (b - w/2)] \times 255/w, & \text{else} \end{cases}$$

Where x represents the image data, y is the grayscale value of the bitmap, w is the window width, b is window bit.

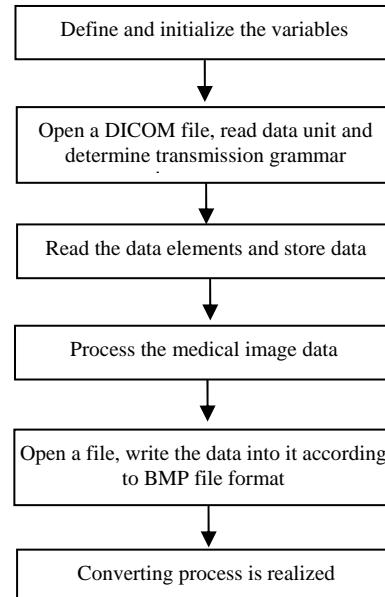


Figure 3. Process of converting DICOM file to BMP file

F. Display of DICOM Image

After the file format conversion, the medical image can be displayed. The software are programmed by using Visual C++, and the image display is implemented by the class CView[9]. The function OnDraw() can convert the image to a device independent bitmap (DIB) and the display it. The specific operations are related to two functions, BitBlt() and StretchBlt(). The former function can copy the bitmap pixels from a memory device context to display device environment or printer using logical coordinates. The latter one also copies a bitmap from one device to another scene.

The process of displaying a medical image is as follows.

- (1) Apply memory space to store bitmap files.
- (2) Read the bitmap file into the applied memory space.
- (3) Use the function CreateDIBitmap() to create and display a bitmap in the functions, such as, OnPaint(), and so on, use CreateCompatibleDC() to create compatible DC, and utilize SelectBitmap() to choose a bitmap.
- (4) Use BitBlt() or StretchBlt() function to display a bitmap.
- (5) Use DeleteObject() to delete the created bitmap.

V. GRayscale Processing of Image

A. Mathematical Morphology

Mathematical morphology makes the rigorous mathematical theory as the basis, focuses on studying the geometry structure and the relationship of the images. It uses a special tool, structural element, to measure the image shape and construct different structural elements to obtain different image processing results. Mathematical morphology has become a new kind of image processing theory and method, which basic ideas and methods have significant impacts on image processing theory and technology. Morphological operation is divided into binary and grayscale computing operations, respectively, corresponding to the binary image processing and grayscale image processing. Their basic algorithms include erosion, dilation, opening and closing. By using these basic algorithms and their combination, we can implement the analysis and processing of the shape and structure of image, including image segmentation, feature extraction, boundary detection, image filtering, image enhancement and restoration, and so on.

B. Window Processing of Grayscale

The grayscale images occupy most parts in medical imaging files, such as, CT images. This image only contains the information of brightness grayscale, without color information. Grayscale representation means to quantify the brightness values. Because the densities of the three-primary colors, red, green and blue, in the RGB representation, the grayscale level is usually divided into 0 to 255. 0 means the darkest (all black) and 255 is the brightest (all white). Point operation is a commonly used technology which can change the grayscale range that the image data occupy. An image can produce a new output image after point operation, and the grayscale range of the input pixel determines range of the output pixel.

If the input image is $A(x, y)$ and the output image is $B(x, y)$, then, the point operation can be expressed as:

$$B(x, y) = f[A(x, y)]$$

Where the function $f()$ is called grayscale transformation function, which is used to describe the conversion relationship between the input and output grayscale values. The point operation can be completely determined if the grayscale transformation function is made.

The window processing of grayscale is to limit a window range. The grayscale value remains unchanged in the range. The value is set to 0 if it is smaller than the lower limit, and it is 255 if greater than the maximum.

The expression of the function of the window processing of grayscale is as follows:

$$f(x) = \begin{cases} 0, & x < M \\ x, & M \leq x \leq N \\ 255, & x > N \end{cases}$$

Where M indicates the lower limit of the window and N represents the upper limit.

C. Grayscale Operation

The object of grayscale operation is image function and the structure element is a grayscale form of a small window, that is, the grayscale value of each point in the window need be determined when the window is

determined. The current point of the image corresponding to the form operation is usually made as a center in the structure elements.

In the following paper, we assume that the input image is f , the structural elements is g , D_f and D_g are the domains of f and g , s and x are the vectors of an integer space.

Erosion operation is one of the basic algorithms of the mathematical morphology. In the macroscopic view, the erosion operation can reduce the processed image. In other words, the function of erosion in mathematical morphology algorithm is to remove the image border. For example, we use the structural elements of 3×3 (in pixels) to corrupt an image, the border of the image will be reduced one pixel.

As a result, we can obtain an important conclusion that we can remove the image of these pixels by the erosion operation if the structural elements used for erosion are larger than some of the image pixels. Take advantage of this conclusion, we can choose different sizes of structural elements to remove the pixels with different sizes of the image. Moreover, we can separate the two pixels if there is small connectivity between two image pixels through the erosion operation.

The concrete realization process of erosion operation is the process of using structural elements to fill image. This process depends on shift which is a basic Euclidean space operation.

The erosion operation of g on f can be defined by the following equation.

$$(f \ominus g)(s) = \min\{f(s+x) - g(x) \mid_{x \in D_g, (s+x) \in D_f}\}$$

Erosion operation is carried out point by point and the computing involves the values of grayscale value and structural elements of the points around it. The operation result is the difference of grayscale values between a local point and the corresponding one in structure element, and we should select the minimum one.

There are two types of effects of the erosion to grayscale image if the values of all structural elements are non-negative.

(1) The image is darkened.

(2) The effect will be weakened if the size of brightened detail of the input image is smaller than the structure element and the weakened extent depends on the grayscale value surrounding these brighten details and the shape and amplitude of the structural elements. The grayscale value of points which grayscale value of edge part is relatively large by the erosion operation can reduce. Then, the edge will draw back to the areas which grayscale value is greater than adjacent areas.

The dilation is the dual operation of inverse operation of erosion, and it can be seen as the erosion operation to the complement set of the original image.

The dilation operation of g on f can be defined by the following equation.

$$(f \oplus g)(s) = \max\{f(s-x) + g(x) \mid_{x \in D_g, (s-x) \in D_f}\}$$

The computation of dilation operation is also carried out point by point. The operation result is the sum of grayscale values of a local point and the corresponding one in structure element, and we should select the

maximum one. Contrary to erosion, dilation can increase the grayscale of the whole image. The results of dilation operation of the parts which grayscale values change greater have larger difference contrary to the original image. For the parts which grayscale values change more gently, the grayscale of all points change very little in addition to increasing the size of core values by dilation operation. After dilation operation, the edge has been extended[10].

As same as erosion operations, we can get different dilation results to the given object image when taking different structural elements.

The role of dilation operations in mathematical morphology is to combine the background points around the image into the image. If the two objects are very close with each other, the dilation operation can cause two objects connect together. Dilation is useful to fill the holes after image segmentation.

Erosion and dilation are not reciprocal and their combination constitutes the other two basic mathematical morphology operations, opening and closing, the former is first erosion and last dilation, and the latter is first dilation and last erosion.

D. Program Implementation

The program implementation of erosion operation and dilation operation are not difficult[11]. The part code of erosion using Visual C++ is given as follow.

```

for (h = 1; h < hei - 1; h++)
{
    for (w = 0; w < wid; w++)
    {
        pSoue = ( char* ) pBit + wid * h + w ;
        pDest = ( char* ) lpNewDIB + wid * h + w ;
        p = ( unsigned char ) * pSou ;
        if ( p != 255 && *pSou != 0 )
            //grayscale detection
            break;
        *pDes = ( unsigned char ) 0 ;
        for ( m=0; m < 3; m++)
            //operate for 3×3 structure element
        {
            for ( n=0 ; n<3 ; n++)
            {
                if ( strct [m][n] == -1)
                    continue;
                p = *(pSou+((2-m)-1)*wid+(n-1));
                if(p==255)
                {
                    *pDes=(unsigned char)255;
                    break;
                }
            }
        }
    }
}

```

VI. EDGE DETECTION OF IMAGE

A. Edge Detection

Edge detection is a basic problem in computer vision and image processing which objective is to identify the points which brightness changes apparently in digital images or to use an algorithm to calibrate out the image edges.

The significant changes of image properties often reflect the important events and changes of properties.

- (1) The discontinuity of depth;
- (2) The discontinuity of surface orientation;
- (3) The changes of material properties;
- (4) The changes of scene illumination.

With the development of computer technology and artificial intelligence, the demand on edge detection should focus on the whole image, not be limited to the point of discontinuity of grayscale. The future trend of edge detection is the combination of doctor's experience and advanced technology and the practical and efficient methods of edge detection[12].

B. Methods of Edge Detection

There are many methods for edge detection which can be divided into two categories, search-based and zero-crossing-based.

The edge detection method based on search first calculates the edge strength, is usually expressed with the first derivative, and then, estimates the local edge direction by calculation, usually the direction of the gradient, and uses this direction to find the maximum value.

The edge detection method based on zero crossing can locate the edge by finding the zero-crossing point of the second derivative of the image. We usually use Laplacian operator or nonlinear differential equations to get it.

As the preprocessing of edge detection, filter is usually necessary and Gaussian filtering is often used.

C. Edge Detection of DICOM Image

In order to ensure the quality of edge detection, we do not directly binary the raw data of DICOM image in this method of edge detection, but first do opening and closing operations by making them as grayscale images, which requires to adjust the data range and variable types of grayscale operations by morphological, so as to meet the features of large scope of DICOM image data relatively.

In order to improve the operating speed, we should change the edge detection algorithms. Because the erosion operation scans the entire image according to the order from top to bottom, from left to right, while the difference operation also do scan the whole image. By analyzing the process of erosion operation, the subsequent erosion operation is only relevant to the original data and structural elements and is not affected by the result of erosion operation. We can modify the erosion algorithms and combine erosion and making difference. After changing the rules of erosion computing, the result is the edge image by scanning the whole image only once, and the efficiency is improved. In addition, before edge detection, by opening and closing the binary

data, we can smooth the edges and reduce operating amount of following operation[13].

The specific steps of edge detection to DICOM image are as follows.

(1) Preprocessing

The opening and closing operations of grayscale are carried out to DICOM image data. The operations can realize the functions of clipping and filling valley to local area of image, and can remove noise, smooth grayscale of the body target to facilitate the separation of the image and enhance edges.

(2) Binary

The data distribution of different tissues and organs of DICOM image has relatively fixed regional range, so we can do the threshold segmentation for the result of the first step according to the medical knowledge not to grayscale histogram, and get the binary image. Because only the edge detection is carried out, binary processing of image can simplify the calculation and increase the running speed.

(3) Smooth edges and do binary opening and closing operations.

Although the edge of human tissue is generally smooth, there are sharp corners, burrs, and so on, due to the noise effects and the limitations of imaging technologies. Binary opening and closing operations can smooth the edges.

(4) Edge detection

We do erosion operation, get the difference between the data before and after operation and detect the edge information. Although the morphological method can achieve better results, there are still many deficiencies. Such as, there are extra pseudo-boundary and needless inner boundary in the image. The edge image got by the above method is still dot-matrix data, and it is very difficult to exclude the inner boundary. This method has filtered a lot of noise information. If using larger structural elements, we can reduce the number of pseudo-boundary, but the computation will be larger and the quality of the real boundary will be impacted. To resolve these problems, the processing steps should be added. After the vectorization of the dot-matrix data, we will do specific treatment according to the features of various types of boundary and get satisfactory results.

(5) Contour extraction

Contour extraction means to vectorize the edge lattice vector. The method is to search along the image boundary, to store the coordinates records of points searched on the boundary in the points list. The result is that each contour line is represented as a point column. The boundary in the edge image got by this method is closed and single-pixel wide. We can use the more efficient eight-step tracking algorithm to vectorize them.

(6) Edge classification

The length of pseudo-boundary is generally shorter, and we should determine it is outer or inner boundary, which can be discriminated by points of the contour line. If choosing contour as starting point, we can make a ray to any horizontal direction and determine the crossing point of this horizontal ray and the other contours. If the

number of crossing point is odd, then the contour represents the inner boundary; or the number is even, then the contour line represents the outer boundary. In this method we can remove the needless inner boundaries.

VII. IMPLEMENTATION OF IMAGE PROCESSING SOFTWARE

The interface of the DICOM medical image processing software based on Visual C++ is shown as Fig. 4.

The result image showed by medical image conversion almost retains all important medical information of the original DICOM images. The software can process the images with a variety of forms, such as, flipping, inverting, grayscale, rotation, middle color, saturation, intensity detection, and so on. It offers a convenient and practical tool to learn how to parse the DICOM standard and complete medical diagnosis.



Figure 4. Interface of image processing software

VIII. CONCLUSION

As an international standard of medical image archiving and communications, DICOM is the basis for all medical imaging technologies. It is very important to study DICOM standard and file format. Based on the analysis of DICOM format and file composition, DICOM format is converted to BMP format and the medical image is displayed in Windows, and many processing operations, such as grayscale, edge detection, and so on, are realized to the medical images. The experiment proves that the system can achieve a better display and processing for DICOM medical images provide good conditions for the clinical diagnosis of doctors and digital storage, and communications of image, besides, it can lay the foundation for subsequent work. The system can run stably, has strong adaptability and can be connected well with PACS.

REFERENCES

- [1] Mildnerger P, Eichelberg M and Martin E, "Introduction to the DICOM standard," *European radiology*, vol. 12, no. 4, pp. 920-927, April 2002.
- [2] Chen Yuanxin, "Study on the Display and Processing of Medical Image based on DICOM 3.0," *Master Degree*

- Dissertation*, Chongqing, China: Chongqing University, April 2004.
- [3] Wu Ruiqing, "Transformation and Processing of Image based on DICOM standard," *Master Degree Dissertation*, Chengdu, Sichuan, China: University of Electronic Science and Technology, March 2002.
- [4] GAO Sheng and GE Yun, "The Display of DICOM Medical Image and Its Information," *Chinese Journal of Medical Physics*, vol. 23, no. 3, pp. 1885-1888, May 2010.
- [5] Huang Aiming, "The Conversion between DICOM Medical Image Format and Common Graphic Formats," *Master Degree Dissertation*, Chengdu, Sichuan, China: Sichuan University, 2006.
- [6] WANG Shigang, LI Yueqing and WANG Changyuan, "Translating from DICOM Image into BMP Image," *Journal of Taishan Medical College*, vol. 28, no. 4, pp. 269-271, April 2007.
- [7] SHI Xiaolei and WANG Mingquan, "Transformation of DICOM digital medical image format into BMP general image format," *Microcomputer Information*, vol. 26, pp. 195-197, September 2010.
- [8] PENG Chenglin, CHEN Cheng and CHEN Yuanyuan, "Conversion of Medical Image based on DICOM with VC++," *Journal of Chongqing University (Natural Science Edition)*, vol. 30, no. 10, pp. 126-129, October 2007.
- [9] Hou Qingfeng, LI Yueqing and WANG Changyuan, "A preliminary research on DICOM image reading and displaying under WINDOWS," *Journal of Medical Imaging*, vol. 15, no. 11, pp. 1013-1015, November, 2005.
- [10] Liu Sen and Chen Jiaxin, "An Edge Detection Method Based on Mathematics Morphology for DICOM Images," *13th Chinese Conference on CAD/CG*, Hefei, Anhui, China, pp. 463-468, 2004.
- [11] Sun Hongwei, "Development for Transmission, Display and Some Image Processing on DICOM Image," *Master Degree Dissertation*, Jilin, Changchun, China: Jilin University, April 2002.
- [12] Shi Xiaolei, "Research of Remote Transmission of DICOM Format Medical Image and Medical Image Processing," *Master Degree Dissertation*, Taiyuan, Shanxi, China: North University of China, June 2010.
- [13] Shen J. and Castan S., "An optimal linear operator for step edge detection," *CVGIP*, vol. 54, no. 2, pp. 112-133, March 1992.

Peijiang Chen

Peijiang Chen received his M.E. degree in Precision Instrument and Mechanics from Harbin University of Science and Technology in 2003. He is an associate professor of Linyi University in Shandong province, China.

His current research interests are focused on medical image processing and remote control.

A New Obstacle Avoidance Method Based on Biped Robot

Zetao Jiang

Nanchang Hangkong University/School of Information Engineering, Nanchang, China
Email: zetaojiang@126.com

Yanru Cui and Qiang Wang

Nanchang Hangkong University/School of Information Engineering, Nanchang, China
Email: cuiyanrumwyauavk@163.com, wangqiang000ws@126.com

Abstract—Currently, in order to avoid obstacle, most robots will choose a new path. This paper proposed a new obstacle avoidance method for biped robots. When lower obstacle exists in the road, we will calculate the obstacle's three-dimensional information instead of using the traditional method of re-select the path, so that the biped robot will avoid obstacles successfully by crossing the obstacle. In 3-d reconstruction, the stereo matching is complex, and has large time consumption. Therefore, the method in this paper has a combination of color segmentation and height detection, extracts obstacle's region accurately to reduce detection range of matching, then improves the real-time of reconstruction. Experiments show that the method proposed in this paper is feasible.

Index Terms—Biped robot, Color segmentation, Height detection, Stereo matching, Obstacle reconstruction

I. INTRODUCTION

In recent years, Robotics is a comprehensive new discipline, which focused on engineering technology such as machinery, computer, control, electronics and the latest research in artificial intelligence[1]. In robotics, machine vision which likes human eyes is one of the key technologies. Mobile robot obtains 3-d information about external environment, mainly through the following two ways: laser rangefinder and stereo vision technology[2]. Laser ranging obtains the 3-d distance images in front of robot through laser-scanning sensors. Therefore, we can get obstacle's information such as shape, height, width, etc[3-4]. However, this method requires relatively complex equipment, and has a bad ability to adapt to the environment, so the practicality is not strong. Stereo vision technology plays an important role in robot navigation applications for its low energy consumption, high reliability, etc[5]. In stereo vision, stereo matching is the most important, its main problem is to find matching points between images. According to different matching primitives, stereo matching can be divided into three types, region matching, phrase

matching and feature matching[6]. The essence of region matching algorithm is to use related degree of gray information between local windows, it can achieve high accuracy in places where have smooth diversifications and rich details. But it is difficult to choose the window size of matching, and has large compute-intensive and slow calculation speed. Phase matching reflects the signal's structure information, and has a good inhibitory effect on the high-frequency noise of images. But it exists major problems, such as phase singularities and phase winding. SIFT feature matching does not directly dependent on gray, and has a strong anti-interference, a small compute-intensive and fast calculation speed[7]. In this paper, we used this algorithm, although we can get only sparse disparity field, we can restore its dense disparity map by the existing dense matching method.

Currently, the robot can be divided into two categories: wheeled robots and biped robots. When the wheeled robot is moving and there is an obstacle in front of it, we only need to calculate some simple information, such as the distance between obstacle and the wheeled robot, the obstacle width, so that the robot has enough time for a new path planning to avoid a collision with the obstacle. In contrast, biped robot has a high degree of imitation of human characteristics, and can cross over the lower obstacle just like people instead of choosing a new path. We need to consider the biped robot's step height, step length and other issues, so we can't achieve its purpose only using traditional obstacle avoidance algorithm. To solve these problems, this paper proposed a new obstacle avoidance method of crossing over the obstacle for biped robot, calculated the obstacle's real 3-d information, so that we can determine the robot's step height, step length and other issues in order to avoid a collision with the obstacle.

II. GET A REAL 3-D OBSTACLE'S IDEAS AND STEPS

Calculating the obstacle's real 3-d information is to 3-d reconstruct the obstacle. It is that extract the 2-d information of obstacle images as rich as possible, then restore the 3-d information as much as possible, so that achieve the 3-d reconstruction of obstacle. Steps in this

Copyright credit, project number: Nature Science Foundation of China (60973096) and Science Foundation of aviation in China(2010ZC56007).

Corresponding author: zetaojiang@126.com

paper are as follows. Firstly, in order to extract the obstacle region quickly, reduce the computation of matching, and achieve a better real-time, we will search for a new method about color image segmentation in the analysis of existing methods[8-10]. Secondly, we will use existing methods based on stereo image pair for dense matching[11], restore a dense obstacle region, and calculate the 3-d coordinates of corresponding points by the use of camera matrix, then find the 3-d proportion of obstacle model. Finally, we will detect the obstacle's real height, and get its real 3-d information according to the 3-d proportion of obstacle model.

III. EXTRACT REGION OF OBSTACLE BY COLOR SEGMENTATION

In binocular vision, stereo matching is the most important task. It needs a lot of points to match, and has a huge computation, time consuming, and low real-time. To solve this problem, we should segment the collected images based on color, remove the complex background, and extract the obstacle region. Therefore, in order to reduce the computing time and improve the real-time, we can extract and match feature points only in the target area.

A. The Selection of Color Model

RGB, YUV, HIS are common color models[12-13]. RGB is the most common color space. After the digital signal processing, the robot's cameras get a value that refers to each pixel value in RGB space. But the RGB color space is uneven, the visual difference between two colors and two points' Euclidean distance in RGB color space are not linearly proportional. In different conditions, such as different types of light source and reflective properties of objects, the distribution of the same color's RGB value is dispersed, and the three components occupied almost the entire color space are interrelated, so it is difficult to find a exact color space threshold. YUV color model, Y component indicates the light intensity, U and V components indicate the color tone[14]. Compared to the RGB color model, the YUV color mode which is independent of light intensity is more suitable for the habit of human eyes, thereby reducing the light and reflection's impact on the color recognition. Another advantage of YUV color model is that it is a linear transformation relationship with RGB color model. When the robot is working online, it can convert the YUV color model to the RGB color model quickly. So some developers will use this color model as the basis for target recognition.

HIS color model is based on visual principles. H component represents the hue, I component represents the brightness, and S component represents the saturation. These three color properties are unrelated, especially the property of H, which has low sensitivity to changes in external light, can reflect the species of colors accurately. For the same color properties of objects, H which has a stable and narrow scope of transformation can be used as a prerequisite for color segmentation; S can be used as auxiliary to determine the conditions in order to

strengthen the image segmentation results; Typically, I component is not usually judged by the standard. Therefore, HIS color model will be used in this paper. HIS and RGB color models' transformation relationship is as follows.

$$\begin{aligned} H &= \cos^{-1} \left\{ \frac{1 / 2[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right\} \\ S &= 1 - \frac{\min(R, G, B)}{I} \\ I &= \frac{(R + G + B)}{3} \end{aligned} \quad (1)$$

B. Color Segmentation Algorithm and Steps

The color of ordinary road is very similar. With the light's change the color of road scene changes, but under the premise that the image is not saturated, the color difference between the road surface and obstacles is always there. Therefore, this paper uses HIS color model which has a smaller sensitivity to light for image segmentation. The purpose is to separate the road and obstacle's region. Specific steps are as follows:

Firstly, choose a $M \times N$ obstacle-free area in front of the robot, and calculate the road's average reference hue H_A and saturation S_A .

$$\begin{aligned} H_A &= \frac{\sum_{i=1}^M \sum_{j=1}^N H(i, j)}{M \times N} \\ S_A &= \frac{\sum_{i=1}^M \sum_{j=1}^N S(i, j)}{M \times N} \end{aligned} \quad (2)$$

Secondly, select each point (i, j) on the image, and calculate the 4×4 neighborhood W of average hue and saturation values.

$$\begin{aligned} h(i, j) &= \frac{\sum_{(u, v) \in W} f(u, v)}{N} \\ s(i, j) &= \frac{\sum_{(u, v) \in W} g(u, v)}{N} \end{aligned} \quad (3)$$

N denotes the number of pixels of the neighborhood W, $f(u, v)$ and $g(u, v)$ represent the hue and saturation of the pixel.

Thirdly, determine any pixel (i, j) is obstacle or not, the formula is as follows:

$$\begin{aligned} F(i, j) &= |h(i, j) - H_A| \\ G(i, j) &= |s(i, j) - S_A| \end{aligned} \quad (4)$$

When $F(i, j)$ is less than the set threshold H_T and $G(i, j)$ is less than the set threshold S_T , the pixel (i, j) represents the road, otherwise the obstacle. After such treatment, the obstacle region can be split out, and the whole background is set to black, leaving only the obstacle

region. This greatly reduces the workload of the feature points' extraction and matching in 3-d reconstruction and saves a lot of time.

IV. HEIGHT DETECTION AND RECONSTRUCTION OF OBSTACLE

A. Preparation Works

Binocular vision refers to that, at the same time, we shoot two images for the same object from different angles. In the world coordinate system, a space point on the target object can be mapped to these two images. Feature point matching is to search these two images, and identify each corresponding pixel points, the pixel points are called matching points. As the external environment, lighting and camera distortion and other negative factors will affect the images, it is difficult to find the matching points exactly. Therefore, feature point matching is an important part of binocular vision studies. Scale Invariant Feature Transform (SIFT) algorithm has more widely used and strong match ability in the strength texture feature areas, and remains invariance for image scaling, rotation, and affine transform[15]. So, in this paper, we use SIFT algorithm to extract and match the features.

The preparation works mainly include four steps. Firstly, extract features from a pair of images. Secondly, match the features by the traditional matching methods. Thirdly, Random Sample Consensus (RANSAC) algorithm is considered one of the most efficient robust estimation algorithm. Therefore, RANSAC will be used to eliminate the false matches further. Finally, compute the parallax that spatial points in two images.

B. The True Height Detection of Obstacle

The parallel stereo vision system based on the biped robots is similar to the whole process that the human binocular obtains the object's depth information. For any object point P in 3-d scene, we get the parallax d of this point first, and calculate the distance OD from this point in optical axis to the camera optical center O. The formula is as follows:

$$OD = \frac{Bf}{d} \quad (5)$$

f is the camera focal length, B is the distance between two camera optical centers, d is the parallax of the measured point mapping on the left and right images.

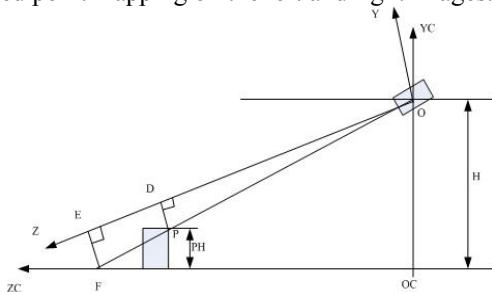


Figure 1. Obstacle height detection.

As shown in Figure1, we can calculate the depth information OD of the obstacle's top point P and OE of the point F on the ground. The pixel parallax of the point

P in images is d_p , and the pixel parallax of the point F in images is d_f , the relationship can be expressed as follows.

$$\Delta d = d_p - d_f = \frac{Bf}{OD} - \frac{Bf}{OE} = \frac{Bf}{OD} \left[\frac{DE}{OE} \right] \quad (6)$$

Seen by the 3-d geometry

$$\frac{DE}{OE} = \frac{P_H}{H} \quad (7)$$

We can know that

$$\Delta d = \frac{Bf}{OD} \left[\frac{P_H}{H} \right] \quad (8)$$

As Δd and OD are known, we can deduce the obstacle's height P_H .

C. The Obstacle Model's 3-d Scale and Reconstruction

In order to obtain the 3-d coordinates of the measured object's feature points, binocular vision sensor should at least obtain two images containing the object's feature points from different locations. Figure2 shows that two cross cameras observe the same object from different angles. In fact getting two images does not necessarily require two cameras. One camera through a certain movement can also obtain two images by observing a stationary object in different locations.

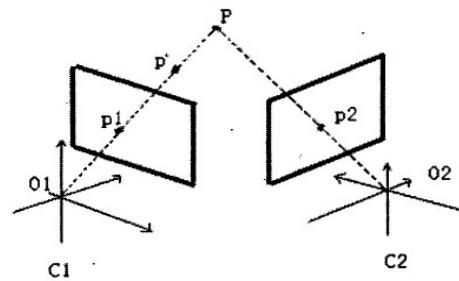


Figure 2. Dual camera observation of space point.

Figure2 shows, for any space point P, its image point in camera C1 is p_1 . Because of that the image point in camera C1 of any point on the connection line O_1P is p_1 , we could not get the 3-d position of P from p_1 . Therefore, from the location of image point p_1 , we only know that the space point P is located on the connection line O_1p_1 . While observing the space point P by camera C2 at the same time, we can get the depth information of point P because of that the space point P is not only on the connection line O_1p_1 , but also on the connection line O_2p_2 . So the point P is the intersection of two lines O_1p_1 and O_2p_2 , p_1 and p_2 is matching points with each other.

We must match the feature points on the images first, and find the matching points in order to determine the 3-d location of their corresponding space point. In this paper, the image retains only the obstacle area through the color segmentation. We can restore the dense matching points concentrated in the obstacle area through the existing algorithm based on image dense stereo matching. The space 3-d coordinates of points on the obstacle can be determined through each pair of matching points.

The cameras C1 and C2 have been calibrated, and their

projection matrices are M1 and M2, then there are the following formulas.

$$Z_{c1} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}^1 & m_{12}^1 & m_{13}^1 & m_{14}^1 \\ m_{21}^1 & m_{22}^1 & m_{23}^1 & m_{24}^1 \\ m_{31}^1 & m_{32}^1 & m_{33}^1 & m_{34}^1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (9)$$

$$Z_{c2} \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}^2 & m_{12}^2 & m_{13}^2 & m_{14}^2 \\ m_{21}^2 & m_{22}^2 & m_{23}^2 & m_{24}^2 \\ m_{31}^2 & m_{32}^2 & m_{33}^2 & m_{34}^2 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (10)$$

$(u_1, v_1, 1)$ and $(u_2, v_2, 1)$ are the homogeneous coordinates of p_1 and p_2 in their respective images. $(X, Y, Z, 1)$ is the homogeneous coordinates of the point P in world coordinate system. m_{ij}^k ($k=1, 2$; $i=1, 2, 3$; $j=1, 2, 3, 4$) is the element located row i column j in the projection matrix M_k . In the formula, Z_{c1} and Z_{c2} can be eliminated, then get four linear equations about X, Y, Z:

$$\begin{aligned} & (u_1 m_{31}^1 - m_{11}^1)X + (u_1 m_{32}^1 - m_{12}^1)Y + (u_1 m_{33}^1 - m_{13}^1)Z \\ &= m_{14}^1 - u_1 m_{34}^1 \\ & (v_1 m_{31}^1 - m_{21}^1)X + (v_1 m_{32}^1 - m_{22}^1)Y + (v_1 m_{33}^1 - m_{23}^1)Z \\ &= m_{24}^1 - v_1 m_{34}^1 \end{aligned} \quad (11)$$

$$\begin{aligned} & (u_2 m_{31}^2 - m_{11}^2)X + (u_2 m_{32}^2 - m_{12}^2)Y + (u_2 m_{33}^2 - m_{13}^2)Z \\ &= m_{14}^2 - u_2 m_{34}^2 \\ & (v_2 m_{31}^2 - m_{21}^2)X + (v_2 m_{32}^2 - m_{22}^2)Y + (v_2 m_{33}^2 - m_{23}^2)Z \\ &= m_{24}^2 - v_2 m_{34}^2 \end{aligned} \quad (12)$$

We can see from the analytic geometry, the geometric significance of formula 12 and 13 is that the line over O_1p_1 (or O_2p_2). Due to the space point P is the intersection of line O_1p_1 and line O_2p_2 , it must satisfy the formula 12 and 13 at the same time. Therefore, we can combine formula 12 with formula 13, then calculate the coordinates (X, Y, Z) of the point P.

After finding the 3-d coordinates of all points in the obstacle region, it is not difficult to find the maximum X_{max} and minimum X_{min} of the coordinate X. So that we can find the length of the obstacle model: $L_x = X_{max} - X_{min}$. Similarly, we can find the width L_y and height L_z , then determine the 3-d proportion of the obstacle model. Finally, the true obstacle's 3-d information can be restored according to the real obstacle height.

V. EXPERIMENTAL RESULTS ANALYSIS

In order to verify the feasibility of algorithm proposed in this paper. We did the experiments as follows. The experiments are realized in VC 6.0 environment.

Fig. 3 shows the original color image pairs of interior

ground scene. Fig. 4 shows the results segmented by general gray segmentation algorithm. Fig. 5 shows the result segmented by the color segmentation algorithm proposed in this paper. The ordinary gray segmentation algorithm can not satisfy our needs. However, through the color segmentation algorithm proposed in this paper, we can extract the obstacle region successfully. Compared to the ordinary gray segmentation algorithm, the segmentation results are much better.



Figure 3. Original color image pair.



Figure 4. Ordinary gray segmentation results.



Figure 5. Color segmentation results.

After segmenting the image pairs, this paper will use SIFT algorithm for feature extraction. Fig. 6 shows the results of SIFT feature extraction. The number of feature points is 128.



Figure 6. SIFT feature extraction.

The next step, we should match the feature points extracted by SIFT algorithm. Fig. 7 shows the results of feature matching. There are 31 pairs of matching points.

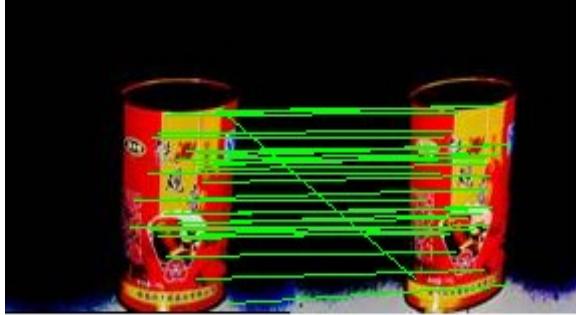


Figure 7. Feature matching.

Typically, the initial matching contains much false matches. Therefore, we should eliminate the false matches through RANSAC algorithm. Fig. 8 shows the results after eliminating the false matches. Finally, the number of matching points is 30.



Figure 8. RANSAC eliminate wrong matches.

As the SIFT algorithm extracts only a sparse image features, this paper will use the image dense stereo matching method based on the region growth. So we can obtain the dense depth map in obstacle region, and reconstruct the obstacle. The results shown in Fig. 9:

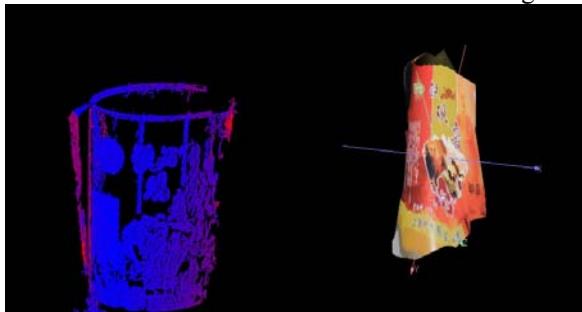


Figure 9. Obstacle depth map and 3-d Reconstruction results.

We can compare the experimental results through constructing a table. The experimental data consist of the obstacle's length, width, and height information. Specific experimental data shown in Table 1:

TABLE I.
COMPARISON OF THE EXPERIMENTAL DATA

Obstacles 3-d Information	Length	Width	Height
The method in this paper	9.93cm	4.93cm	12.23cm
Real obstacle 3-d information	9.85cm	4.75cm	11.98cm

Through measuring the real obstacle, we can know that the obstacle's 3-d information: the length is 9.85cm, the width is 4.75cm, and the height is 11.98cm. In this paper, we can calculate the obstacle model's 3-d ratio (0.812:0.403:1) by computing the 3-d coordinates of points in obstacle region, and then we can calculate the real obstacle's height (12.23cm) by the height detection method. Therefore, the real obstacle 3-d information can be obtained, the length, width, and height are: 9.93cm, 4.93cm, 12.23cm. Because of the color segmentation algorithm can not extract the obstacle region completely and accurately, the camera calibration errors are also exist, there are still errors between the obstacle 3-d information through the method in this paper and the real obstacle 3-d information. However, this small error does not affect the biped robot's decision. The biped robot can cross the obstacle successfully, avoiding collision with the obstacle.

VI. CONCLUSIONS

A new obstacle avoidance method which combines the color segmentation with the height detection and reconstructs the obstacle has been proposed in this paper. In order to determine some information (such as robot's step height, step length) and cross the obstacle successfully, we must calculate the real obstacle 3-d information. Through experimental verification, the proposed method in this paper is feasible. The detection range of match will be reduced by extracting the obstacle region based on the color segmentation, and the time-consuming in whole process of 3-d reconstruction will be also reduced.

ACKNOWLEDGMENT

This paper is sponsored by Nature Science Foundation of China (60973096) and Science Foundation of aviation in China(2010ZC56007).

REFERENCES

- [1] Pei-yu Jiang. "Study of dynamic walking of biped robot" *Master's thesis*. Northwestern Polytechnical University, 2004.
- [2] Wei Cui, Wen-yi Qiang, Xing-lin Chen. "Real time obstacle detection system for biped robot", *Control and Decision*, 2004, 19(1): 40-43.
- [3] Simmons R. "The curvature-velocity method for local obstacle avoidance", *International Conference on Robotics and Automation*, April, 1996.
- [4] Surmann H, Nüchter A, Hertzberg J. "An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitalization of indoor environments ", *Robotics and Autonomous Systems*, 2003, 45(3 — 4) : 181-198.
- [5] Chow Y H, Ronald C. "Obstacle avoidance of legged robot without 3D reconstruction of the surroundings", *Hong Kong: The Chinese University of Hong Kong*, 1999.
- [6] Yu Li. "Obstacle recognition research based on binocular vision", *Master's thesis*, Wuhan Technology University, 2007.

- [7] S. Allaire, J. Kim, S. Breen, D. Jaffray, and V. Pekar. "Full orientation invariance and improved feature selectivity of 3D SIFT with application to medical image analysis", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008*, pages 1-8, June 2008. Allaire S., Kim J., Breen S.
- [8] Qing-zhong Li, Xian-hua Chen, Wei-kang Gu. "Fast Method for Obstacle Detection Based on Color Stereo Vision", *Computer Science*, 2003, 30(9):72-75.
- [9] Kai-yan Lin, Jun-hui Wu, Li-hong Xu. "A Survey on Color Image Segmentation Techniques", *Journal of Image and Graphics*, 2005, 10(1):1-10.
- [10] G. Balasubramanian, "Unsupervised Color Image Segmentation Using a Dynamic Color Gradient Thresholding Algorithm", *Thesis (M.S.)-Rochester Institute of Technology*, 2008.
- [11] Bi-na Zheng, Ze-tao Jiang, Min Wu. "3-dimensional reconstruction based on image dense stereo matching", *Chinese Journal of Stereo Logy and Image Analysis*, 2009, 14(2):187-191.
- [12] Zheng Wang, Zeng-qi Sun. "Color Object Recognition Algorithm Based on RoboCup Biped Humanoid Robot", *Journal of University of Jinan*, 2007, 21: 18-22.
- [13] Asmare, M.H., Asirvadam, V.S., Iznita, L. "Color space selection for color image enhancement application," in *ICSAp*, 2009.
- [14] Kelson R.T. Aires, Andre M. Santana, Adelardo A.D. Medeiros "Optical flow using color information preliminary results". SAC'08 March 1620, 2008, Fortaleza, Cear'a, Brazil Copyright 2008 ACM 9781595937537/08/0003.
- [15] Lowe D G. "Distinctive Image Features from Scale-Invariant Points ". *International Journal of Computer Vision*, 2005, 60(2):91-100.

of Philosophy in Northwestern Polytechnical University, Xi'an, China, in 2006. Main research areas were image processing, computer vision, and network information security.

He works at School of Information Engineering at Nanchang Hangkong University in Jiangxi. More than 60 papers were published, nearly 30 articles were included in SCI, EI. He presided key technology research project "Control system based on multi-level security agent", which was awarded the second class prize in the 2009 annual science and technology progress of Jiangxi province. He presided the research of network and information security defense technology based on digital watermarking, which was awarded the first class prize in science and technology achievements of Jiangxi province. He presided and finished Nature Science Foundation of China (60673055), two subject Foundations of province, an open foundation of measurement and control center of province and more than 10 other subjects. Currently, he is presiding Nature Science Foundation of China (60673055) and Nature Science Foundation of Province, which is in the field of 3d reconstruction.

Dr. Jiang, Prof. Zetao is director of Chinese Society for Stereology.



Yanru Cui Female, born in 1987 in Shanxi Province, and received the B.S. degree in Changzhi College, Changzhi, China, in 2009. Currently, she is pursuing the M.S. degree in Nanchang Hangkong University, and the main research area is image processing and computer vision.



Zetao Jiang Male, born in 1961 in Jiangxi Province, Professor, Doctor. He is adjunct professor of Nanjing University of Aeronautics and Astronautics, PhD supervisor, academic leader in Jiangxi province. He received the B.S. degree in Beijing Normal University, Beijing, China, in 1986, and received the M.S. degree in Tongji University, Shanghai, China, in 1995, and received the Doctor



Qiang Wang Male, born in 1983 in Hubei Province, and received the B.S. degree in Huazhong University of Science and Technology Wuchang Branch, Wuhan, China, in 2009. Currently, he is pursuing the M.S. degree in Nanchang Hangkong University, and the main research area is image processing and computer vision.

The PID Controller Based on the Artificial Neural Network and the Differential Evolution Algorithm

Wei Lu

The Control Science and Engineering Department of Dalian University of Technology, Dalian, China
Email: luweimail@126.com

Jianhua Yang and Xiaodong Liu

The Control Science and Engineering Department of Dalian University of Technology, Dalian, China
Email: {jianhuay,xdliu}@dlut.edu.cn

Abstract—The conventional PID (proportional-integral-derivative) controller is widely applied to industrial automation and process control field because its structure is simple and its robust is well, but it do not work well for nonlinear system, time-delayed linear system and time-varying system. This paper provides a new style of PID controller that is based on artificial neural network and evolutionary algorithm according to the conventional one's mathematical formula. The artificial neural network (ANN) is used to approach PID formula and the differential evolution algorithm (DEA) is used to search weight of the artificial neural network. This new controller is proven better control effect in the simulation test. This new controller has more advantages than the conventional one, such as less calculated load, faster global convergence speed, better robust, more independence and adaptability on the plant and independent of human intervention and expert experiences etc.

Index Terms—the artificial neural network, the differential evolution algorithm, PID controller

I. INTRODUCTION

The conventional PID (proportional-integral-derivative) controller is widely applied to industrial automation and process control, for its control mode is direct, simple and robust. But, there are some disadvantages of PID control. Firstly, it is difficult to regulate the three parameters of PID controller: K_p , K_i and K_d in some control systems. Secondly, the conventional PID controllers generally do not work well for nonlinear system, time-delayed linear system, complex and vague system, time varying systems [1][2][3][4]. Various types of modified PID controller have been developed, such as self-turning PID controller, to overcome the one problem of the regulation conventional PID controller parameter. The fuzzy PID controllers and the neural network PID controllers are also designed for this purpose.

The natural representation of control knowledge make fuzzy controller easy to be understood [13]. But the most fuzzy controllers use two inputs, error, the change rate of

error to approximately behaves like a PD controller, and obviously there would exist steady-state error when industrial process systems are controlled by fuzzy controller. It can eliminate the steady-state error of the control system to consider the integration of error in input of the fuzzy controller. Of course this can be realized by designing ,a fuzzy controller with three inputs, error, the change rate of error and integration of error. However, it will be hard to implement in practice because of the difficulty in constructing control rules base. First, it is not the practice for expert to observe the integration of error. Second, adding one input variable in fuzzy controller will greatly increase the number of control rules [7][8].

The artificial neural network has the ability of learning and function approximation. In addition, the artificial neural network learning processes are independent of human intervention and expert experiences. For such situations, many studies use ANN to approximate PID formula to realize ANN-PID controller. But the learning method of ANN usually adopts some traditional algorithm, including the delta rule, the steepest descent methods, Boltzman's algorithm, the back-propagation learning algorithm, the standard version of genetic algorithm [9][10][11], etc. These traditional learning methods of ANN exists some deficiency including such as the problem of the slow speed of convergence, local minima, and the large amount of computation of network, etc, which lead to ANN-PID controller is difficult to use actually[15][16]. In this paper, a new ANN PID controller which is based on the differential evolution algorithm (DEA) is proposed. Here, artificial neural network is used to approximate PID formula and using DEA to train the weights of ANN. The simulation proves this controller can get better control effect, and it is easily realized and the less amount of computation.

The remainder of the paper is organized as follows. Section 2 briefly described the conventional ANN-PID controller. Section 3 briefly introduced the basic idea of the differential evolution algorithm. Section 4 presents proposed the framework and algorithm of the ANN-PID

controller based on the differential evolution algorithm (ANN-PID-DEA). Section 5 applies the proposed framework and algorithm to five examples with different complexity levels to demonstrate its control ability and learning capability of ANN-PID-DEA controller. Finally, Section 6 provides the conclusion.

II. THE CONVENTIONAL ANN-PID CONTROLLER

It is well known, there are two traditional PID controller modes, one is locational mode, and the other is incremental mode. The ANN realization of locational mode PID is shown below. It can be referenced to get the one of incremental mode.

$$\begin{aligned} u(k) &= K \left[e(k) + \frac{T}{T_i} \sum_{j=0}^k e_j(k) + \frac{T_d}{T} (e(k) - e(k-1)) \right] \\ &= K_p e(k) + K_I T \sum_{j=0}^k e_j(k) + K_D \frac{\Delta e(k)}{T} \end{aligned} \quad (1)$$

where $K_p = K$, $K_I = K/T_i$, $K_D = KT_d$, T is the sampling period, $u(k)$ is output of the PID controller, $e(k)$ is the deviation.

For equation (1), $u(k)$ is the linear combination of $\left\langle e(k), T \sum_{j=0}^k e_j(k), \Delta e(k)/T \right\rangle$, i.e.

$$u(k) = f \left(e(k), T \sum_{j=0}^k e_j(k), \frac{\Delta e(k)}{T} \right) \quad (2)$$

Feed-forward ANN is used to construct an ANN-PID controller. Generally, a three layered feed-forward ANN with appropriate network structure and weights can approach to any random continuous function. The ANN is designed with three layers in consideration of the control system real time requirement. Obviously, there are three nodes in input layer, the deviation $e(k)$, the cumulation of deviation $e_j(k)$ and the variety of deviation $\Delta e(k)$. Only one node in output layer, that is, the output of the controller $u(k)$. In order to simplify the structure of the ANN, hidden layer nodes, which can correctly reflect the relationship between the input and the output of the ANN, are designed as few as possible, and 8 nodes are assumed in this paper. In practice, hidden layer nodes can be acquired by experiment or experience [4][15][16]. The neuron activation function of input layer is assumed linear function $f_i(x) = x$; that of hidden layer is assumed the Sigmoid function $f_h(x) = 1/(1+e^{-x})$ and that of output layer is assumed linear function $f_o(x) = x$. So a 3-8-1 network is constructed and is shown in Fig. 1, which can take the place of traditional PID controller.

In Fig.1, the ANN-PID three inputs of input layer are $\left\langle e(k), T \sum_{j=0}^k e_j(k), \Delta e(k)/T \right\rangle$, $u(k)$ is the output of ANN-PID.

ANN-PID controller generally adopts steepest descent method to learn weights based on the objective function J.

$$J = \frac{1}{2} [y_p(k+1) - y(k+1)] \quad (3)$$

Where, y_p is the desired output of the controlled object at step $k+1$, y is the actual output of the controlled object at step $k+1$.

As we all know, using gradient descent method to train

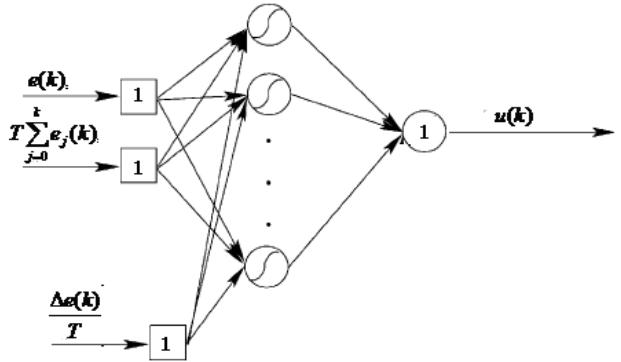


Figure 1. 3-8-1 ANN-PID controller.

ANN leads to falling into local minimum and slowing speed of converge, therefore in order to get ride of this defect, in next section, the different evolution algorithm (DEA) is adopted to train weights of the ANN-PID.

III. THE DIFFERENTIAL EVOLUTION ALGORITHM

The different evolution algorithm (DEA) is a branch of the evolution algorithms, which has been developed by Storn and Price [17]. DEA is a simple evolutionary algorithm that creates new candidate solutions by combining the parent individual and several other individuals of the same population. A candidate replaces the parent only if it has better fitness. This is a rather greedy selection scheme that often outperforms traditional EAs. Like the other EA, utilizes N, n-dimensional parameter weight vectors $w_{i,G}$, $i = 1, \dots, N$, as a population for each iteration, called generation, of the DEA algorithm. The initial population is taken to be uniformly distributed in the search space. At each iteration, the mutation and crossover operators are applied on the individuals, and a new population arises. Then, the selection phase starts, where the N best points from both populations are selected to comprise the next generation.

According to the mutation operator, for each weight vector, $w_{i,G}$, $i = 1, \dots, N$, a mutant vector is determined through the equation:

$$V_{i,G+1} = W_{r1,G} + F \cdot (W_{r2,G} - W_{r3,G}) \quad (4)$$

where $r1, r2, r3 \in \{1, \dots, N\}$, are mutually different random indexes and also different from the current index i . $F \in (0, 2]$ is a real constant parameter that affects the differential variation between two vectors, and N must be

greater than or equal to 4, in order to apply mutation. Following the mutation phase, the crossover operator is applied on the population, combining the previously mutated vector, $v_{i,G+1} = [v_{1i,G+1}, v_{2i,G+1}, \dots, v_{Di,G+1}]$ with a so-called target vector, $w_{i,G+1} = [w_{1i,G+1}, w_{2i,G+1}, \dots, w_{Di,G+1}]$. Thus a so-called trial vector, $u_{i,G+1} = [u_{1i,G+1}, u_{2i,G+1}, \dots, u_{Di,G+1}]$ is generated, according to

$$u_{ji,G+1} = V_{ji,G+1}, \text{ if } (\text{Randb}(j) \leq CR) \text{ or } j = \text{rnbr}(i) \quad (5)$$

or

$$u_{ji,G+1} = V_{ji,G+1}, \text{ if } (\text{Randb}(j) > CR) \text{ or } j \neq \text{rnbr}(i) \quad (6)$$

where $i = 1, \dots, N$, $\text{randb}(j) \in [0, 1]$ is the j th evaluation of a uniform random number generator, for $j \in 1, 2, \dots, D$, and $\text{rnbr}(i) \in 1, 2, \dots, d$ is a randomly chosen index. $CR \in [0, 1]$ is the crossover constant (user defined), a parameter that increases the diversity of the individuals in the population. The three algorithm parameters that steer the search of the algorithm are the population size (N), the crossover constant (CR) and the differential variation factor (m). They remain constant during an optimization. To decide whether or not the vector $u_{i,G+1}$ should be a member of the population comprising the next generation, it is compared to the initial vector $w_{i,G}$. Thus,

$$w_{i,G+1} = \begin{cases} u_{i,G+1}, f(u_{i,G+1}) < f(w_{i,G}) \\ w_{i,G}, \text{ otherwise} \end{cases} \quad (7)$$

The procedure described above is considered as the standard variant of the DE algorithm. Different mutation and crossover operators have been applied with promising results. In addition, DE algorithms have a property that Price has called a universal global mutation mechanism or globally correlated mutation mechanism, which seems to be the main property responsible for the appealing performance of DE as global optimizers.

To apply DEA algorithms to ANN training weights as an initial weight population, and evolve them over time; N is fixed throughout the training process, and the weight population is initialized randomly following a uniform probability distribution.

At each iteration, called generation, new weight vectors are generated by the combination of the weight vectors randomly chosen from the population. This operation is called mutation. The derived weight vectors are then mixed with another predetermined weight vector, the “target” vector, through the crossover operation. This operation yields the so called trial vector. The trial vector is accepted for the next generation if and only if it reduces the error value of the objective function J ((3)). This last operation is called selection. The above-mentioned operations introduce diversity in the population and are used to help the algorithm escape the local minima in the weight space. The combined action of mutation and crossover is responsible for much of the effectiveness of DEA’s search, and allows them to act as parallel, noise-tolerant, hill-climbing algorithms, which efficiently search the whole weight space [22][23].

IV. THE ANN-PID LEARNING ALGORITHM BASED ON THE DEA

As shown in the following Fig. 2, ANN-DEA-PID is controller based on the artificial network and the different evolve algorithm, r is system referenced input, y is system output, u is controlling output of ANN-DEA-PID controller. Deviation e , cumulation of deviation Σe and variety of deviation Δe are applied as the inputs of ANN-DEA-PID controller. In this control system, we adopt the hybrid method combined of differential evolution algorithm and steepest descent algorithm for ANN-DEA-PID. The DE algorithm works on the termination point of steepest descent algorithm. Thus the method consists of a steepest descent strategy-based ANN-DEA-PID training stage and a differential evolutionary strategy-based ANN-DEA-PID retraining stage.

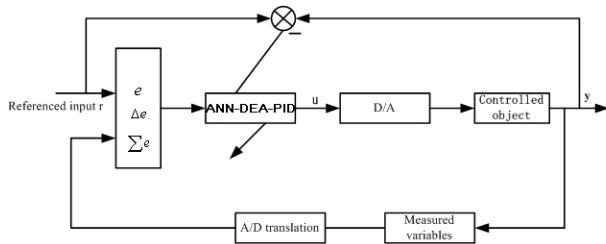


Figure 2. The ANN-PID control system based on DEA.

A. First Stage: the Steepest Descent Algorithm for ANN-DEA-PID

The proposed learning procedure is based on the gradient descent learning rule for ANNs learning. In Fig.1, Input layer’s weights are $\langle 1, T, 1/T \rangle$. Input of hidden layer’s j th neuron can be written as follows:

$$I_{hj} = \sum_{i=1}^3 \omega_{ij} X_i \quad (8)$$

Where ω_{ij} is the weight which connect j th ($j=1,2,\dots,8$) neuron in hidden layer with i th ($i=1,2,3$) neuron in input layer; X_i ($i=1,2,3$) are the outputs of input layer, that is:

$$\left\langle X_1 = e(k), X_2 = T \sum_{j=0}^k e_j(k), X_3 = \frac{\Delta e(k)}{T} \right\rangle \quad (9)$$

Output of the j th neuron in hidden layer is:

$$O_{hj} = f_h(I_{hj}) \quad (10)$$

Input of the neuron in output layer is

$$I_{o1} = \sum_{j=1}^8 \omega_{j1} O_{hj} \quad (11)$$

ω_{j1} ($j=1,2,\dots,8$) is the weight which connects the only neuron in output layer with the j th neuron in hidden layer. Output of the neuron in output layer is

$$O_{o1} = f_o(I_{o1}) = I_{o1} \quad (12)$$

Output of the neutral network PID controller is

$$u(k) = O_{o1} = f_o(I_{o1}) = I_{o1} \quad (13)$$

1) Adjusting of Output Layer's Weights

Weight learning rule is

$$\Delta\omega_{j1} = -\eta \frac{\partial J}{\partial \omega_{j1}} \quad (14)$$

Where η is learning speed, it commonly is (0, 1).

$$\begin{aligned} \frac{\partial J}{\partial \omega_{j1}} &= \frac{\partial J}{\partial y(k+1)} \frac{\partial y(k+1)}{\partial O_{o1}} \frac{\partial O_{o1}}{\partial \omega_{j1}} \\ &= -(y_p(k+1) - y(k+1)) O_{hj} \frac{\partial y(k+1)}{\partial u(k)} \end{aligned} \quad (15)$$

where $\partial y(k+1)/\partial u(k)$ is unknown which denotes system's input-output relationship. For most systems, its signal is definite. $\partial y(k+1)/\partial u(k)$ is replaced with $\text{sgn}(\partial y(k+1)/\partial u(k))$, and learning speed η is used to equalize the calculate error. Adjusting rule of output layer's weights ω_{ji} is:

$$\begin{aligned} \omega_{j1}(k+1) &= \omega_{j1}(k) + \\ &\eta(y_p(k+1) - y(k+1)) O_{kj} \text{sgn}(\frac{\partial y(k+1)}{\partial u(k)}) \end{aligned} \quad (16)$$

2) Adjusting of Hidden Layer's Weights

Weight learning rule is

$$\begin{aligned} \Delta\omega_{ij} &= -\eta \frac{\partial J}{\partial \omega_{ij}} = -\eta \frac{\partial J}{\partial I_{hj}} \frac{\partial I_{hj}}{\partial \omega_{ij}} = -\eta \frac{\partial J}{\partial I_{hj}} X_i \\ &= -\eta \left(\frac{\partial J}{\partial O_{hj}} \frac{\partial O_{hj}}{\partial I_{hj}} \right) X_i \\ &= -\eta \left(\frac{\partial J}{\partial O_{hj}} \right) O_{hj} (1 - O_{hj}) X_i \end{aligned} \quad (17)$$

$$\begin{aligned} \frac{\partial J}{\partial O_{hj}} &= \frac{\partial J}{\partial I_{o1}} \frac{\partial I_{o1}}{\partial O_{hj}} \\ &= \frac{\partial J}{\partial I_{o1}} \frac{\partial J}{\partial O_{hj}} \sum_{i=1}^8 \omega_{j1} O_{hj} = \frac{\partial J}{\partial I_{o1}} \omega_{j1} \\ &= \frac{\partial J}{\partial y(k+1)} \frac{\partial y(k+1)}{\partial O_{o1}} \omega_{j1} \\ &= -(y_p(k+1) - y(k+1)) \omega_{j1} \cdot \frac{\partial y(k+1)}{\partial u(k)} \end{aligned} \quad (18)$$

where $\partial y(k+1)/\partial u(k)$ is replaced by $\text{sgn}(\partial y(k+1)/\partial u(k))$, learning speed η is used to equalize the calculate error. Corresponding adjusting rule of hidden layer's weights ω_{ij} is:

$$\begin{aligned} \omega_{ij}(k+1) &= \omega_{ij}(k) + \eta(y_p(k+1) - y(k+1)) \\ &\omega_{j1} O_{hj} (1 - O_{hj}) X_i \text{sgn}(\frac{\partial y(k+1)}{\partial u(k)}) \end{aligned} \quad (19)$$

A generic description of the proposed hybrid algorithm is given in Algorithm 1.

Algorithm 1

Stage 1: the steepest descent learning

Step 1a: Decide network structure at first. Because the nodes of network input layers and output layers are known, only the nodes of hidden layers remained undecided.

Step 2a: Initialize the weights of hidden layers ω_{ij} and the ones of output layers ω_{j1} with less random number, select the speed of learning η ;

Step 3a: Repeat for each input concept state (k).

Step 4a: Sample the system,

get $\left\langle e(k), T \sum_{j=0}^k e_j(k), \frac{\Delta e(k)}{T} \right\rangle$, which are the network inputs;

Step 5a: According to formula (9) and (11) calculate the outputs of hidden layer and output layer, get the controlling amount u ;

Step 6a: Calculate the system output and get $y(k+1)$;

Step 7a: According to the weights adapting rule (15) and (18) of output layer and hidden layer, regulate each connection weight of output layer and hidden layer.

Step 8a: Calculate objective function J

Step 9a: Until the termination conditions are met.

Step 10a: Return the final weights $W_{SDA}(k+1)$ to the Stage 2.

Stage 2: the differential evolution learning

Step 2b: Initialize the DE population in the neighborhood of $W_{SDA}^{(k+1)}$ and within the suggested weight constraints (ranges).

Step 2b: Repeat for each input concept state (k).

Step 3b: For $i = 1$ to NP .

Step 4b: MUTATION ($w_i^{(k)}$) → Mutant_Vector.

Step 5b: CROSSOVER (Mutant_Vector) → Trial_Vector.

Step 6b: If $J(\text{Trial_Vector}) \leq J(w_i^{(k)})$, accept Trial_Vector for the next generation.

Step 7b: End For.

Step 8b: Until the termination condition is met.

First, the steepest descent algorithm is outlined in the Stage 1 of Algorithm 1, and provides convergence of concepts' values in a desired state. The key features of the steepest descent algorithm method are the low storage requirements and the inexpensive computations. In Stage 2 of Algorithm 1, the differential evolution (DE) algorithm, responsible for the ANN-DEA-PID retraining is outlined.

B. Second Stage: the Differential Evolution Algorithm

To apply DEA algorithms to ANN-PID retraining weights starting with a specific number (N) of n-

dimensional weight vectors, as initial population, and evolve them over time; N is fixed throughout the training process, and the weight population is initialized by perturbing the appropriate solution provided by the steepest descent algorithm. Also, the appropriate fitness function is determined. In this case, the steepest descent algorithm seeds the DE, i.e. a preliminary solution is available by the steepest descent algorithm; so, the initial population might be generated by adding normally distributed random deviations to the nominal solution. However, in the experiments reported in the next section, we have also used the constraints on weights defined initially by experts to perturb the approximated solution provided by steepest descent algorithm.

Let us now give some details about the version of DE algorithm used here. Steps 4b and 5b implement the mutation and crossover operators, respectively, while Step 6b is the selection operator. For each weight vector $w_i^{(k)}$ the new vector called mutant vector is generated according to the following relation:

Mutant vector

$$\begin{aligned} &= v_i^{(k+1)} \\ &= w_i^{(k)} + \mu \cdot (w_{best}^{(k)} - w_i^{(k)} + w_{r1} - w_{r2}) \\ &= 1, 2, \dots, NP \end{aligned} \quad (20)$$

Where $w_{best}^{(k)}$ is the best population member of the previous iteration, $\mu > 0$ is a real parameter (mutation constant) which regulates the contribution of the difference between weight vectors, and w_{r1} , w_{r2} are weight vectors randomly chosen from the population with $r_1, r_2 \in \{1, 2, \dots, i-1, i+1, \dots, N\}$, i.e. r_1, r_2 are random integers mutually different from the running index i . Aiming at decreasing the diversity of the weight vectors further, the crossover-type operation yields the so-called trial vector, $u_i^{(k+1)}$, $i = 1, \dots, N$. This operation works as follows: the mutant weight vectors ($v_i^{(k+1)}$, $i = 1, \dots, N$) are mixed with the “target” vectors, $w_i^{(k+1)}$, $i = 1, \dots, N$. Specifically, we randomly choose a real number r in the interval $[0, 1]$ for each component j , $j = 1, 2, \dots, n$, of the $v_i^{(k+1)}$. This number is compared with $CR \in [0, 1]$ (crossover constant), and if $r \leq CR$; then, the j th component of the trial vector $u_i^{(k+1)}$, gets the value of the j th component of the mutant vector $v_i^{(k+1)}$; otherwise, it gets the value of the j th component of the target vector, $w_i^{(k+1)}$. The trial vector is accepted for the next generation if and only if it reduces the value of the following proposed fitness function J ; otherwise the old value, $w_i^{(k+1)}$, is retained. This last operation is the selection, and due to the moving “optimum” nature of the differential evolution task, it ensures that the fitness J starts steadily decreasing at some iteration. Here, clearly, the fitness function J is formula (3).

The purpose is to determine the values of the weights of the ANN-DEA-PID that produce a desired behavior of the system. The determination of the weights is of major significance and it contributes towards the establishment of ANN-DEA-PID as a robust methodology, and improves the performance of ANN-DEA-PID.

V. SIMULATION ANALYSIS

The neural network PID controller with the differential evolution algorithm, which is proposed in this paper, is constructed by neural network PID controller and the differential evolution algorithm. In this new PID controller, the artificial neural network is used to approach the conventional PID formula and the differential evolution algorithm (DEA) is used to search weight of the artificial neural network.

Five examples with different complexity levels are described in this section and are used for the simulations.

A. First Order System

The first example is the first order systems model as follow

$$G(s) = \frac{1}{95s + 1} \quad (21)$$

For this delay time system, the parameters are chosen as follow: sample period $T_s = 1s$, reference value $r = 100$, the traditional PID controller parameters are $K_p = 23$, $K_i = 1.375$, $K_d = 0.002$. The output response obtained is shown in Fig. 3 as curve 1 for step input signal.

Curve 2 in Fig.3 shows the result of ANN-DEA-PID controller with the 3-8-1 neural network and ANN's learning speed is 0.1.

Curve 3 in Fig.3 also depicts the output performance of the ANN-PID controller with the 3-8-1 neural network in which the parameters after learning are: $T_s = 1s$, $r = 100$, and ANN's learning speed is 0.1.

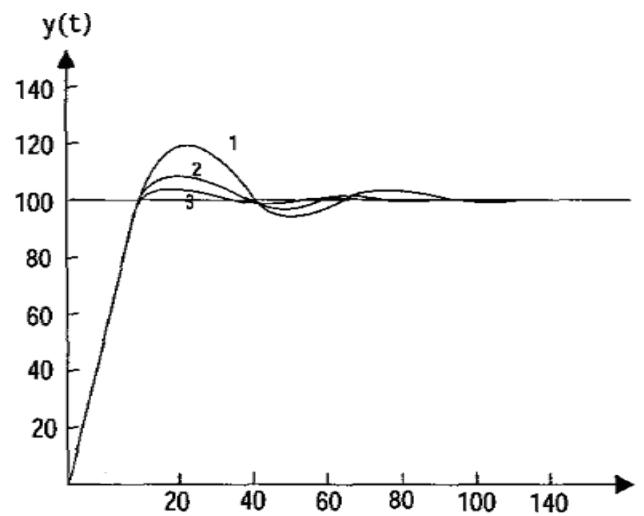


Figure 3. The ANN-PID control system based on DEA.

B. First Order System with Time Delay

The second example is the first order systems with the time delay model as follow

$$G(s) = \frac{e^{-10s}}{95s + 1} \quad (22)$$

For this delay time system, the parameters are chosen as follow: sample period $T_s = 1s$, reference value $r = 100$, the traditional PID controller parameters are $K_p = 13, K_i = 0.325, K_d = 0.008$. The output response obtained is shown in Fig. 4 as curve 1 for step input signal.

Curve 2 in Fig.4 shows the result of ANN-DEA-PID controller with the 3-8-1 neural network and ANN's learning speed is 0.1.

Curve 3 in Fig.4 also depicts the output performance of the ANN-PID controller with the 3-8-1 neural network in which the parameters after learning are: $T_{s_1} = 1s, r = 100$, and ANN's learning speed is 0.1.

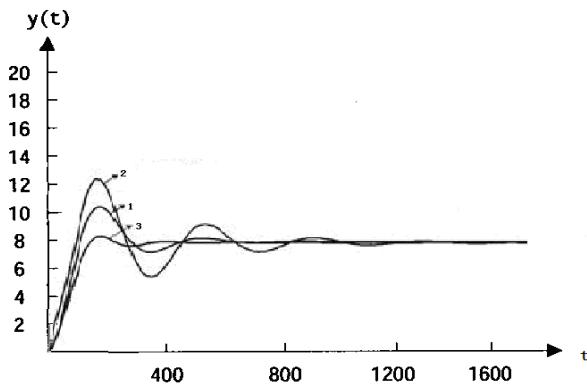


Figure 5. The control effect of PID, ANN-PID and ANN-DEA-PID controller for second order system.

C. Second Order System

The second example is the second order systems with the simple model as follow

$$G(s) = \frac{1}{s(s+1)} \quad (23)$$

For this second order system, the parameters are chosen as follow: sample period $T_s = 0.5s$, reference value $r = 8$, the traditional PID controller parameters are $K_p = 3.57, K_i = 0.125, K_d = 0.012$.

The output response obtained is shown in Fig. 5 as curve 1 for step input signal.

Curve 2 in Fig. 5 shows the result of ANN-DEA-PID controller with the 3-8-1 neural network and ANN's learning speed is 0.1.

Curve 3 in Fig. 5 also depicts the output performance of the ANN-PID controller with the 3-8-1 neural network in which the parameters after learning are: $T_{s_1} = 1s, r = 100$, and ANN's learning speed is 0.1.

D. Second Order System with Time Delay

The third example is the second order system with time delay. The model of it is obtained as following

$$G(s) = \frac{e^{-2s}}{(10s+1)(23s+1)} \quad (24)$$

For this process, the suitable parameters of conventional PID controller are considered as: sample period $T_s = 0.5s$, reference value $r = 10$; $K_p = 0.1902, K_i = 0.279, K_d = 0.26628$.

The curve of output response that marked curve 1 is shown in Fig 6.

For ANN-PID controller with the 3-8-1 neural network and ANN's learning speed is 0.1, the output response, called curve 2, is given in the same time.

Curve 3 in Fig.6 gives the result of ANN-DEA-PID controller with the 3-8-1 neural network and ANN's learning speed $\eta = 0.1$.

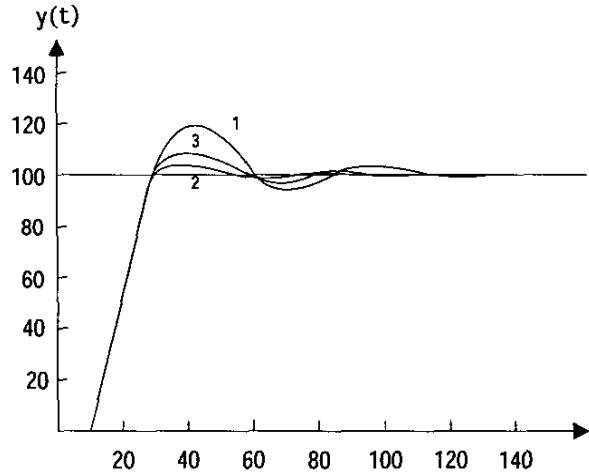


Figure 6. The control effect of PID, ANN-PID and ANN-DEA-PID controller for first order system with time delay.

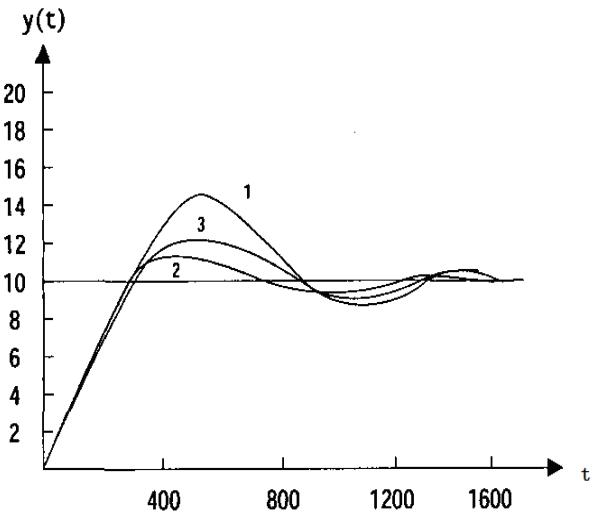


Figure 4. The control effect of PID, ANN-PID and ANN-DEA-PID controller for second order system with time delay.

E. the Time Variable Nonlinear System

During the simulation, a typical time variable nonlinear system, which model is

$$y(k) = \frac{a_0(k)y(k-1) + u(k-1)}{1 + y^2(k-1)},$$

$$a_0(k) = 1 + 0.1 \sin\left(\frac{k\pi}{25}\right) \quad (25)$$

It is taken as the controlled object. Obviously, the conventional PID controller can not be applied to time variable nonlinear system such as (24).

In here, The ANN-DEA-PID controller with the 3-8-1 neural network structure and the traditional ANN-PID controller are applied respectively. When the input signal is square wave with amplitude is [-1, 1] and period is 10s. The control system sample period is $T_s=0.05$ s, the corresponding response curves are shown in Fig.6 and Fig.7.

From Fig.7 and Fig.8, it is obvious that the control effect of ANN-DEA-PID controller better than ANN-PID controller and the ANN-DEA-PID controller is more independent and adaptable on the model of the controlled object. This also exactly reflects the outstanding learning and retaining ability and self-adapting ability of neural network. In addition, the different evolve algorithm is adopted to train weights of network, which promotes the performance of the control system and decrease calculated load.

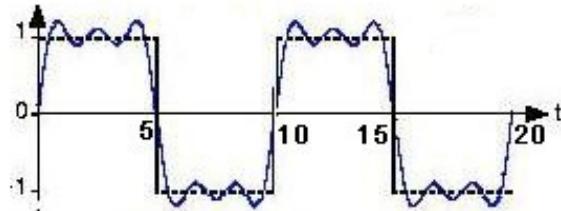


Figure 7. The control effect of ANN-PID controller for the time variable nonlinear system.

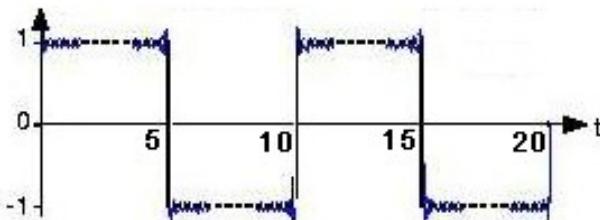


Figure 8. The control effect of ANN-DEA-PID controller for the time variable nonlinear system.

VI. CONCLUSION

The advantages of the ANN-PID controller based DEA are summarized as following:

1. The mathematics model for the controlled object is not necessary;
2. The structure of the controller and the algorithm is simple, and it is convenient to apply them to online real-time control;
3. It is more robust. The method can be applied to industrial process control, and take full advantage of both neural network and traditional PID.

4. ANN-PID controller based on DEA convergence speed is faster than ANN-PID controller and ANN-DEA-PID controller can reach global minimum, due to adopt the different evolutional algorithm to train weight of neural network.

REFERENCES

- [1] Yu Yongquan, Huang Ying and Zeng Bi, "A PID Neural Network Controller," Proceeding of the International Joint Conference on Neural Networks, IEEE Computer Society Press, California, vol. 3, pp. 1933-1938, 2003.
- [2] Yue Pan,Ping Song and Kejie Li, "PID Control of Miniature Unmanned Helicopter Yaw System based on RBF Neural Network," Intelligent computing and Information Science, vol. 135, pp. 308-313, 2011.
- [3] Haiquan Wang, Dongyun Wang and Guotao Zhang, "Research of Neural network PID Control of Aero-engine," Advances Automation and Robotics, vol. 1, pp. 337-343, 2010.
- [4] Iwasa,T.,Morizumi,N.andOrmatu,S. "Temperature Control in a Batch Process by Neural Networks," Proceeding of IEEE World Congress on Computational Intelligence, IEEE Press, New York, vol. 2, pp. 992-995, 1992.
- [5] Li qi, Li shihua. "Analysis and Improvement of a Kind of neural Networks Intelligent PID Control Algorithm," Control and Decision Editorial Department, ShenYang China, vol. 13, pp. 311-316, 1998.
- [6] Moradi,M.H. "New Techniques for PID Controller Design," Proceeding of 2003 IEEE Conference on Control Applications, IEEE Press, New York, vol. 2, pp. 903-908, 2003.
- [7] Indranil Pan,Saptarshi Das,Amitava Gupta, "Tuning of an optimal fuzzy PID controller with stochastic algorithms for networked control systems with random time delay," ISA Transaction, vol. 50, pp. 28-36, 2011.
- [8] G.jahedi, M.M.Ardehali, "Genetic algorithm-based fuzzy-PID control methodologies for enhancement of energy efficiency of a dynamic energy system," Energy Conversion and management, vol. 52, pp. 725-732, 2011.
- [9] T. Back,H.P. Schwefel. "An overview of evolutionary algorithms for parameter optimization," Evol. Comput., pp. 1-23, 1993.
- [10] N. Dodd, "Optimization of network structure using genetic techniques," in Proceedings of the 1990 International Joint Conference on Neural Networks, 1990.
- [11] P. Bartlett, T. Downs, "Training a Neural Network Using a Genetic Algorithm," Tech. Report, Department of Electrical Engineering, University of Queensland, 1990.
- [12] H.G. Beyer, H.P. Schwefel, "Evolution strategies: a comprehensive introduction," Nat. Comput., vol. 1, pp. 3-52, 2002.
- [13] Saptarshi Das,Indranil Pan,Shantanu Das,etc, "A novel fractional order fuzzy PID controller and its optimal time domain tuning based on integral performance indices," Engineering Applications of Artificial Intelligence, online, 2010.
- [14] I.L. Lopez Cruz, L.G. Van Willigenburg, G. Van Straten, "Efficient differential evolution algorithms for multimodal optimal control problems," Application Soft Computing, vol. 3, pp. 97-122, 2003.
- [15] G.D. Magoulas, V.P. Plagiannakos, M.N. Vrahatis, "Neural network-based colonoscopic diagnosis using online learning and differential evolution," Appl. Soft Comput., vol. 4, pp. 357-367, 2004.
- [16] Mohannad Anwar Hosen,Mohd Azlan Hussain,Farouq S.Mjalli, "Control of polystyrene batch reactors using

- neural network based model predictive control (NNMPC): An experimental investigation," Control Engineering Practice, vol. 19, pp. 454-467, 2011.
- [17] R. Storn, K. Price, "Differential evolution: a simple and efficient heuristic for global optimization over continuous spaces," J. Global Optimization, vol.11, pp. 341–359, 1997.
- [18] G. Weiss, "Towards the synthesis of neural and evolutionary learning," in: O. Omidvar (Ed.), Progress in Neural Networks, 5, Ablex Pub, 1993 Chapter 5.
- [19] G. Weiss, "Neural networks and evolutionary computation, partI: hybrid approaches in artificial intelligence," in: Proceedings of ICEC, 1993.
- [20] G. Weiss, "Neural networks and evolutionary computation, part II: hybrid approaches in neurosciences," in: Proceedings of the IEEE World Congress on Computational Intelligence, 2009.
- [21] J. Zurada, "Introduction to Artificial Neural Systems," West Publishing Company, 1992.
- [22] I.L. Lopez Cruz, L.G. Van Willigenburg, G. Van Straten, "Efficient differential evolution algorithms for multimodal optimal control problems," Appl. Soft Comput., vol. 3, pp. 97–122, 2008.
- [23] Elpiniki I. Papageorgiou, Peter P. Groumpos, "A new hybrid method using evolutionary algorithms to train Fuzzy Cognitive Maps," Applied Soft Computing, vol.5, pp. 409–431, 2005.
- [24] R. Storn, K. Price, "Differential evolution: a simple and efficient heuristic for global optimization over continuous spaces," J.Globaloptimization, vol.11, pp. 341–359, 1997.
- [25] Shu Huailin, "PID Newral Network Control forComplex Systems," Processdings of InternationalConference on Computational Intelligence for Modelling, Control and Automation CCIMCA 99'2,10s Press, 1999, pp. 166-171.
- [26] Shu Huailin, "Analysis of PID Neural Network Multivariable Control Systems," ACTA Automatica Sinica, vol.25, pp. 105-111, 2008.
- [27] K.V. Price, "An introduction to differential evolution," in: D.Corne, M. Dorigo, F. Glover (Eds.), New Ideas in Optimization, McGraw-Hill, New York, 1999.
- [28] J.C.F. Pujol, R. Poli, "Evolving the topology and the weights ofneural networks using a dual representation," Appl. Intell., vol.8, pp. 73–84, 2010.
- [29] H.P. Schwefel, "Numerical Optimization of Computer models," Wiley, Chichester, 1981.
- [30] H.P. Schwefel, Evolution and Optimum Seeking, Wiley, NewYork, 1995.
- [31] R. Storn, "On the Usage of Differential Evolution for Function Optimization," NAFIPS, Berkely, pp. 519–523, 1996.
- [32] R. Storn, K. Price, "Minimizing the real functions of the ICEC'96 contest by differential evolution," in: Proceedingsof the IEEE Conference on Evolutionary Computation, Nagoya, pp. 842–844,2009.
- [33] A. Tiwari, R. Roy, G. Jared, O. Munaux, "Evolutionary-based techniques for real-life optimization: development and testing," Appl. Soft Comput., vol.1, 2002, pp. 301–329.
- [34] J. H. Kim, K.K.Choi, "Self-turning discrete PID Controller," IEEE Trans. Indst. Electron., vol.34, pp. 298–300. 2007.
- [35] Larsson,Hagglund, "Control signal constraints and filter order selection for PI and PID controllers," American Control Conference, pp. 4994-4999,2011.

A Measuring Method for Angular Displacement Based on Correlation Algorithm

Ping Yan

State Key Laboratory of Mechanical Transmission/Chongqing University, Chongqing, China
Email: 13983038527@139.com

Yan He

State Key Laboratory of Mechanical Transmission/Chongqing University, Chongqing, China
Email: heyans@cqu.edu.cn

Runzhong Yi

State Key Laboratory of Mechanical Transmission/Chongqing University, Chongqing, China
Email: yrz111@gmail.com

Fei Liu

State Key Laboratory of Mechanical Transmission/Chongqing University, Chongqing, China
Email: fliu@cqu.edu.cn

Guorong Chen

School of Electronic & Information Engineering/Chongqing University of Science and Technology, Chongqing, China
Email: cqgcr@189.cn

Abstract—The existing angular displacement measurement methods rely on the manufacturing precision of the fixed-plate and moving-plate with high manufacturing cost, and it is difficult to overcome a static error and drift for the static measurement method. A kind of angular displacement measuring method based on correlation algorithm is presented with the characteristics of a low manufacturing cost, high precision, anti-noise and anti-partially damaged properties and so on. The method is that firstly full circle broadband or white noise, random data is a pre-prepared coaxially in a rotating body, then the periodic random signal is continuously formed using pickup head reads this data in the basic uniform rotation process of the rotating body. The instantaneous angular displacement of the pickup head relative to the rotating body is obtained by means of the correlation operations between the periodic random signal and the signal sequence of pre-stored data. The angular displacement among the different pickup heads is gained by subtraction for the instantaneous angular displacement of different pickup heads relative to the rotating body at the same instant. The functional relationship between the relative angular displacement of the different pickup heads and the measured angular displacement is determined by linkage equations, which is used to calculate the measured

angular displacement. Furthermore, the schematics of detective devices and the principle of the signal processing are developed to implement the method.

Index Terms—angular disposition, signal processing, correlation algorithm, measurement

I. INTRODUCTION

The existing angular displacement measurement methods are essentially classified into two categories: static measuring method and counting method. Give two examples for the former, one is that the angular displacement is read directly from mechanical indexing plate in which there are marked lines in the sub-degree angle plate, and the pointer on the base[1]. Another example is the capacitive angular position sensor, which transforms angular displacement into the variance in capacitance value[2-5]. For the latter, radial grating angular displacement sensor is most typical one. This kind of sensor can turn angular displacement into grating rotating angular, and the angular displacement value is obtained by measuring the number of the counting pulse that the grating outputs in the course of the relative rotation between the fixed plate and the moving plate on the radial grating[6-7]. Some literature has reported the current angular displacement measurement methods. Reference [8] introduced the determination of angular

Manuscript received July 10, 2011; revised December 20, 2011; accepted January, 2012.

Copyright credit, project number: National Natural Science Foundation of China (51105394), "Five-second" National Science and Technology Support Program of China (No. 2011BAF11B10), and Doctoral Fund of Ministry of Education of China (20100191120004), corresponding author: Ping Yan

position of rotary device by means of plate freely placed on the rotating part and the laser scanning instrument is presented together with some experimental results of such layout application. Reference [9] proposed a displacement and angular drift simultaneous measurement technique based on a defocus grating. The displacement and angular drift of the incident beam can be detected by monitoring the movements of +/- 1 diffraction order spots of the defocus grating. The relationship between drift of the incident beam and movements of +/- 1 diffraction order spots was studied in detail. Reference [10] presented an optical approach of attenuated total reflection. As a laser beam is incident upon a planar optical waveguide, an m line is obtained by scanning the incident angle. Theoretical analysis showed that the m line sharply shifts with a tiny variation of the thickness of the wave guided layer. And the specific schemes for ADM are analyzed, which are based on the angular interrogation and the intensity measurement. The calculated result of sensitivity demonstrates that the intensity measurement is more efficient than the angular interrogation. Reference [11] proposed a high-accuracy angular displacement measurement using sinusoidal phase-modulating Fabry-Perot interferometer. A CCD image sensor is used for measuring the distance between the transmitted beams from the two faces of the Fabry-Perot plate. From the distance, the initial angle of incidence is calculated. The sinusoidal phase-modulating interferometry is used for improving the measurement accuracy. It is insensitive to the external disturbance. Numerical calculation and experimental results make it clear that the interferometer allows high-accuracy measurements of angular displacements. Reference [12] provided an optical pickup head applied in angular measurement, which uses a 45 DEG plate glass to generate an astigmatic optical pickup head that is used in combination with an atomic microscope to precede angular measurement. The device of optical pickup head includes a laser diode, a 45 DEG plate glass, a collimating lens, a reflector, an object lens, a tracking starter and a quadrant detector (PDIC), wherein the quadrant detector is employed to detect signals, and the offset relation between the signals and the surface under test is determined to confirm the relation between the displacement/angle and the change of the signals. Reference [13] presented a small in-plane angular displacement by adopting two matched moire fringes and frequency analysis method. Reference [14] showed the use of path-entangled states of photons, having nonzero orbital angular momentum (OAM), increased the resolution and sensitivity of angular-displacement measurements performed using an interferometer.

However, for the above existing method, the accuracy deeply depends on the manufacture accuracy of fixed plate and the moving plate, and the manufacturing cost of instrument is relatively high. Additionally, it is difficult to avoid static errors and drift for the static measuring method. Therefore, a new method of angular displacement measurement based on correlation

algorithm is proposed to resolve these problems in the paper.

II. PRINCIPLE OF ANGULAR DISPLACEMENT MEASUREMENT BASED ON CORRELATION ALGORITHM

Full circle sequence is etched or recorded in advance by the optical, electrical, magnetic or mechanical methods on a plate or a drum body rotating around its central axis (which is called as rotation body). The sequence is located in one or more of the circular track around the same central axis, so that it forms one or more circular information orbits. The etched or recorded sequence is broadband, or white noise random, or pseudorandom sequence which has sufficient broadband random signal characteristics. The pickup head aims at the circular information orbits and continuously reads the pre-recorded random signal while the plate or a drum body rotating around its central axis. The output of the pickup head is the continuous loop random signal which has the period of the body rotational cycle along with its rotation.

The relationship between the measured angular and the angular of two pickup heads relative to the rotation axis is determined using arbitrary linkage methods. Thus the measured angular is calculated by the following formula based on the angular of two pickup heads relative to the rotation axis.

$$\omega = f(\theta). \quad (1)$$

Where, ω is the measured angular, θ is the angular of two pickup heads relative to the rotation axis, $f(\theta)$ is the function according to the linking method.

The function is determined by the linking method. The measurement of the measured angular displacement ω is transformed into the measurement of the angular of two pickup heads relative to the rotation axis θ with respect to (1).

There are two processing methods after the output of the pickup head is sent into Signal Processing Unit (simplified as SPU):

1) SPU stores previously the same random sequence as the etched or recorded one on the plate or a drum body. The sequence of correlation function value is obtained, which is the result of cross-correlation operation between the output sequence of pickup head and the pre-stored sequence in SPU at any time. Let τ be the time among two near maximums in the sequence of correlation function value. And also, τ is the time lapse a between the output sequence of pickup head and the pre-stored sequence. Then the instantaneous angular displacement α between the start position of the body data track and the position of the pickup head is calculated by the following the formula:

$$\alpha = \frac{a}{T_0} \times 360^\circ. \quad (2)$$

Where, T_0 is time corresponds to the body rotating one cycle.

The angular value θ between two pickup heads is calculated by subtraction operation between the instantaneous angular displacement of the two pickup heads and by eliminating the zero error. θ is calculated as follows:

$$\theta = \alpha_1 - \alpha_2 - \varphi. \quad (3)$$

Where, α_1 , α_2 are individual instantaneous angular dispositions of the two pickup heads, φ is the zero error.

At last, the measured angular displacement can be calculated by (1).

2) If the two pickup heads read the same information orbits, the time lapse is obtained directly by correlation operation between their output sequences of the two pickup head. Their relative angular displacement a is calculated by (2), then the measured angular displacement calculated by (1).

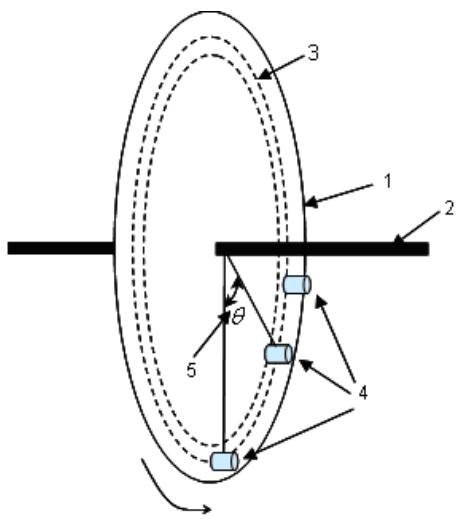
Furthermore, the random data sequences recorded in different tracks may be the same or completely different. In addition, these sequences may also include clock information.

The number of pickup heads is optional according to the number of measured angular displacements. The orbit which the pickup heads read may be the same, or not. If some of the pickup heads is fixed on the base as benchmark, the absolute angular displacement for the other pickup heads relative to the benchmark can be obtained. On the other hand, the relative angular displacement between any two pickup heads can be directly measured without a benchmark.

III. IMPLEMENTATION OF METHOD

The apparatus implementing the above method consists of Detector Unit (simplified as DU) and Signal Processing Unit (simplified as SPU).

A. Detector Unit



1. rotation body 2. centre axis 3. information orbits 4. pickup head(s)
5. angular between two pickup heads

Figure 1. Note how the caption is centered in the column.

DU is composed of one centre axis, one rotation body and one or more pickup heads (see as Figure 1). The centre line of the axis must be coincided with that of measured object rotating. The rotation body is a plate or a drum body as 2. There are many implementing methods for the rotation body and the pickup head, as follow:

(1) Using optical disk as a rotation body, pickup head is laser pickup head. In this case, the information orbits of rotation body are the circular ones which are made by pressing from a stamper or by recording using CD/DVD writer. The optical disk rotates in uniform speed driven by the spindle motor which is controlled by servo system in order to make sure the reliability of the output of the pickup heads. The optical disk manufacture, the optical disk recording, the laser pickup heads, and the servo-controlling systems is implemented by the existing components of CD, DVD, blue ray DVD, so the cost is low.

(2) Using hard magnetic disk as a rotation body, the pickup head is magnetic head. The information orbits of rotation body are circular magnetic ones. The relative parts are implemented by the existing components of magnetic disk.

(3) Using metal plate as a rotation body, the pickup head is metal electrode. The data is gained by the capacitance value between the plate and the electrode.

The recorded information sequence is analog signals or digital signals as mentioned above. If it is digital signal, the recovered data is processed directly by high-speed correlation operation and the angular displacement is calculated after being sent into SPU. If it is analog signals, the output of the pickup heads should be converted into digital signal before being sent to SPU, then post-processed signal is processed by high-speed correlation operation and the angular displacement is calculated. In order to facilitate signal recovery, recorded signal can be modulated, and also include clock information.

The pickup head aims at the circular information orbits and continuously reads the pre-recorded random signal while the rotation body is rotating. The output of the pickup heads is continuous periodic random data sequences. The autocorrelation function of the output satisfies the property shown in Figure 2.

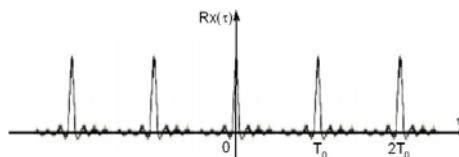


Figure 2. The autocorrelation function of wideband or white noise with the period of T_0

B. Signal Processing Unit (SPU)

The Flow of Signal Processing: One flow of signal processing in SPU is shown in Figure 3. The signal should be pre-processed through filtering, amplification and sampling, then recovered the data sequence and clock when the output of the read pickup is analog signal. After pre-processing and recovering, continuous data flow is formed. The data flow repeats continuously pre-record

random sequence when the rotation body rotating. The same sequence is also pre-stored in SPU. The sequences of the correlation function value are obtained by cross-correlation calculation between the data flow and the sequence stored in SPU at any time point. The τ value corresponding to the first maximum value in the sequences of the correlation function value is time delay between the data flow and the sequence stored in SPU. The a value is corresponding to instantaneous angular displacement between the current data start position on tracks in the rotation body and the pickup head, according to (2).

The angular displacement between the two pickup heads is obtained by subtraction operation between the instantaneous angular displacements of two pickup heads

respectively corresponding to the data start position on tracks in the rotation body. If the data tracks on which the two pickup heads read data and their start position are different, the measured value of the angular displacement between the two pickup heads is obtained by the value of the subtraction operation subtracting a zero-error (the included angle between the start positions of the two tracks). Then the measured angular displacement is gotten by (3).

The signal processing also adopts the flow shown in Figure 4 instead of Figure 3, if the two pickup heads read data on the same track. The time delay a is obtained by the direct correlation operation between the output of the two pickup heads according to (2). Then measured angular displacement is calculated by (1).

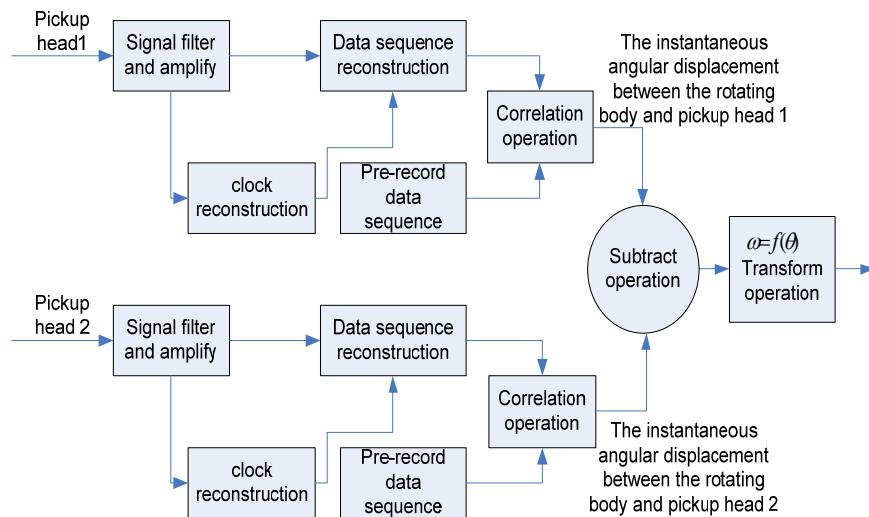


Figure 3. One Flow Chart of Signal Processing in SPU

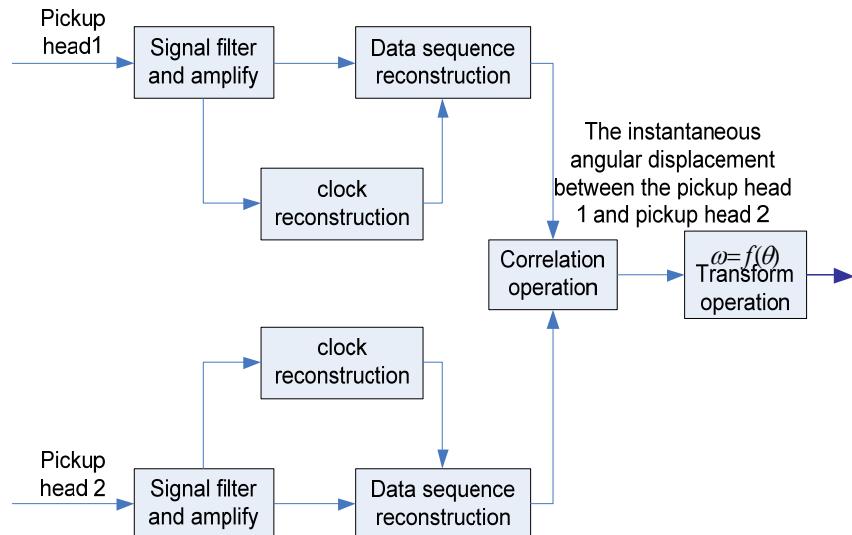


Figure 4. Another Flow Chart of Signal Processing in SPU

The function relationship between the relative angular displacement of pickup heads and the measured angle displacement: The conversion formula (see in (1)) between the relative angular displacement of the pickup heads and the measured angle only depends on the inking geometrical relationship of the pickup heads and the measured angle.

For example, the relationship is $\omega = \theta$ when the pickup heads are directly fixed at the two sides of the measured angle, and the rotary axis of the rotation body is vertical to the measured angle and through the angular vertex.

For another example, the relationship is $\omega = k\theta$ when the measured angle links with the pickup head by a speed reduction set for gear transmission. Here, k is the transmission ratio of gear.

The calculation method of time delay between two signal based on correlation algorithm: If $x(t)$ is a sample function of ergodic random processes, and $x(t+\tau)$ is a sample that is time shifted by τ from $x(t)$, the $R_x(\tau)$ which represents a self-correlation function of $x(t)$ is defined as follows [15-18] :

$$R_x(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)x(t+\tau)dt. \quad (4)$$

For different properties $x(t)$, the self-correlation functions have following characteristics: For signal that has the periodic elements, the self-correlation function does not even attenuate when the τ is very large and has obvious periodicity. For the signal that has not periodic elements attenuates quickly and tends to zero with the τ increasing. Especially, the value of the self-correlation functions for broadband random noises is maximum at $\tau=0$ and quickly attenuate to zero shown in Figure 5, while that for narrowband random noise has a slow attenuation property. The self-correlation function for white noises (whose frequency band is limitless) is a δ function (a pulse function) at $\tau=0$ shown in Figure 6.

In this method, long enough information sequence (T_0) is intercepted from the broadband random noise or the white noise as pre-recorded information sequence. Repeating this information sequence forms a pseudo random signal with period T_0 . The figure of self-correlation function of this signal is shown in Figure 2. This figure has sharp features with period T_0 . It reaches to its maximum (sharp peak value) at $\tau=0, \pm T_0, \pm 2T_0, \dots, \pm nT_0$. If the signal is digital signal, the maximums only appear at $\tau=0, \pm T_0, \pm 2T_0, \dots, \pm nT_0$.

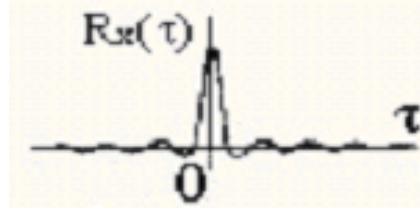


Figure 5. The self-correlation function of the broadband random signal

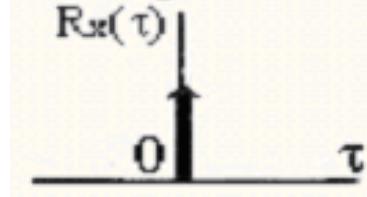


Figure 6. The self-correlation function of the narrowband random signal

For ergodicity random process, the cross-correlation function $R_{xy}(\tau)$ of two random signal $x(t)$ and $y(t)$ is:

$$R_{xy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)y(t+\tau)dt. \quad (5)$$

Especially, if $y(t) = x(t-a)$, in other words $y(t)$ is a time delay signal of $x(t)$, according to the definition above about the self-correlation function and the cross-correlation function, it is easy to derive out:

$$R_{xy}(\tau) = R_x(\tau-a). \quad (6)$$

So the figure of cross-correlation function for $x(t)$ and $y(t)$ is the same as that of self-correlation function for $x(t)$, only with a time shift which is equal to the time delay from $y(t)$ to $x(t)$.

Consequently, if $x(t)$ and $y(t)$ both are broadband random noise or the white noise with period T_0 , and $y(t) = x(t-a)$, the cross-correlation function $R_{xy}(\tau)$ is shown as Figure 7. It reaches its maximum at $\tau=0, \pm T_0, \pm 2T_0, \dots, \pm nT_0$. Otherwise, it attenuates very quickly at other positions.

If a few noises or errors are introduced to the signal during the transmission or reading, it means that $x(t)$ and $y(t)$ are not only different in time, but also in shape. It can be proved by simple derivation that if these noises and errors are not very serious, the figure of correlation function is similar to Figure 7, but its peak is lower than that in Figure 7. The connotation of a is shown in Figure 7. This result can ensure that this measurement method possesses the strong anti-noise and anti-destruction properties.

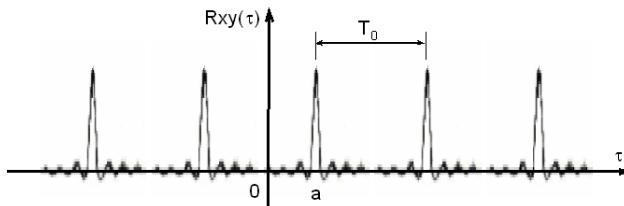


Figure 7. The correlation function between $y(t)$ and $x(t)$ if $y(t) = x(t - a)$ and $x(t)$ is the broadband random signal of the repeating period T_0

IV. CASE STUDY

An experimental setup for angular displacement measurement is designed to apply the method as shown in Figure 8. A 2.5-inch micro disc of hard disk is used as rotation body in this setup. The disc is fixed on the rotation axis through its center and driven at constant rotation speed by a micro-DC servo motor. Two magnetic heads are installed in one end of two branch arms as read-write heads, and the other end of the arms is installed on the rotation axis through the bearing. The measured angular displacement is the angle between the two lines connecting two heads and the center of the rotation axis. The distance between the two heads and the center of the rotation axis is the same. The two heads read the same information. So the measured angular displacement is calculated only by related operation between the signals of the two heads.

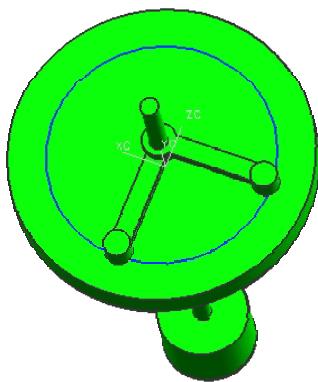


Figure 8. The experimental setup

The first step of experiment is recording random data on the disc of the magnetic disk. After the disc achieves uniform rotation driven by motor, the signals, which is modulated by the completely random binary white noise data generated by a white noise signal generator, is added to one magnetic head. After a period of time, the signal is revoked. So the random data is recorded on the disc. While the disc is rotating in uniform manner, the output of the two magnetic heads is amplified and demodulated to regain the binary data stream. Then the two signals are processed by synchronization and related operation. The method is described as the following steps:

Firstly, completely record long enough sequence (two or more laps complete data) according to the output data of any head. This sequence is processed by

autocorrelation operation. So a full circle of data sequence can be cut accurately. And calculate the size of this data sequence and store the size. Each data represented the value of angular displacement is gained by 3600 divided by the size.

Next, the sequence delay (number of binary bit) between the magnetic head is gained by correlation operation between the output of the heads. And the angular displacement between the two heads related to the rotation axis is obtained by multiplying the number of binary with the angular displacement per bit.

In the experiment, the disc rotates uniformly at the speed of 50 laps per second, the white noise data is recorded on the disc at the speed of 100Mbps, so that 2Mbit information orbit per lap is gained and per bit represents 0.648 arc-sec. The experiment shows that the measurement resolution can reach 0.648 arc-sec. The experiment also proves that the measurement accuracy is slightly affected when the data on the disc is damaged partly (less than 20% of the non-continuous trauma).

V. CONCLUSIONS

A new method of measuring angular displacement based on correlation algorithm is put forward in the paper. Comparing with the existing methods, the proposed method has strong anti-noise and anti-part-data-destruction properties. And also, the measurement set based on this method has many advantages such as high measuring accuracy, low manufacturing cost, etc. Since this method makes full use of the existing record methods and parts in the optical disk or magnetic disk field, it can widely apply in many measuring angular displacement fields.

ACKNOWLEDGMENT

This work was supported in part by a grant from National Natural Science Foundation of China (51105394), "Five-second" National Science and Technology Support Program of China (No, 2011BAF11B10), and Doctoral Fund of Ministry of Education of China (20100191120004)

REFERENCES

- [1] T. W. Eom, C. S. Kang, J. A. Kim, J. W. Kim, G. J. Sik, K. J. Wan, et al. "System for simultaneous measurement of linear and angular displacement," KR Patent 2009121885-A. Dec. 26, 2009.
- [2] G. Brasseur, "A Robust Capacitive Angular Position Sensor," *Proceedings of IEEE Instrumentation and Measurement Technology Conference*, Brussels, Belgium, June 04-06, 1996.
- [3] C. Gerard, M. Meijer, MeijerM, "A contactless capacitive angular disposition sensor," *IEEE Sensors Journal*, vol. 3, no. 5(2003), pp.607 – 614.
- [4] C. Liang, K. Yan, "Capacitive angular-disposition detective system based on single chip processor," *Transducer and Microsystem*, vol. 8, no. 25(2006), pp.52-54.
- [5] A. H Falkner, "The Use of Capacitance in the Measurement of Angular and Linear Displacement," *IEEE*

- Transactions on Instrumentation and Measurement*, vol. 43, no. 6(December1994), pp.939-942.
- [6] Ouyang, Wangli, "The principle and setting of the cylindrical grating angular disposition sensors," *Chinese Journal of Scientific Instrument*, vol. 2, no. 11(1990), pp.102-106.
- [7] S.Qin, M. Zhang, D. Luo, "The Principle and Technology of Mechanical Engineering," Chongqing University Press, Chongqing , 2002.
- [8] D. Brucas, V. Giniotis, "Measurement of angular displacement by means of laser scanner," *Journal of Vibroengineering*, vol. 12, iss. 2 (2010), pp. 177-184.
- [9] H. T. Ma, F. J.Xi, X. L.Wang, L. A. Liu, Y. X. Ma, X. J. Xu, et al. *Applied Optics*, 49, 4420(2010).
- [10] F. Chen, Z. Q. Cao, Q. S. Shen, Y. J. Feng, "Optical approach to angular displacement measurement based on attenuated total reflection," *Applied Optics*, vol. 44, Issue 26(2005), pp. 5393-5397.
- [11] C. Zhang, X. Wang, *Acta Optica Sinica*. 24, 1141 (2004).
- [12] D. Hsieh, W. Sun, J. Yang, "Optical pickup head applied in angular measurement," TW Patent 200905675-A, Feb. 1 2009.
- [13] J Y Shao, H Z Liu, Y C Ding, L Wang, "Determination of small in-plane angular displacement in machine tool alignment using two matched moire fringes". *Proceedings of the Institution of Mechanical Engineers Part B-Journal Engineering Manufacture*, vol. 225, no. B4(2011),pp.495-504
- [14] A. K. Jha, G. S. Agarwal, R. W. Boyd. "Supersensitive measurement of angular displacements using entangled photons". *Phys. Rev.*, Vol. 83, no 5(2011),pp. 053829-1-5
- [15] V. O. Alan, S. Alan, S. Willsky, N. Hamid, "Signals & Systems (2nd)," Prentice Hall.USA ,1997.
- [16] P. H. Hwei (L. Luo, J. Hu, Z. Li), "The signal and system," Science Press, China ,2002.
- [17] G. Hu, "Digital signal processing theory algorithm realization," Tsinghua Publishing House, China ,2003.
- [18] P. Q. Cheng, "Course of Digital Signal Processing," Tsinghua University Press, China ,2001.

Ping Yan, Borned in Sichuan Province of China, earned Ph.D degree in electrical engineering and automation of Chongqing university of China in 1997.

She is a professor in College of Mechanical Engineering in Chongqing university, China. She has published more than 100 papers. Her research area focuses on Mechatronics and Intelligence manufacturing.

Prof. Yan is member of Standing Committee in Chinese Mechanical Engineering Society.

Yan He, Borned in Chongqing of China, earned Ph.D degree in mechanical engineering of Chongqing university of China in 2007.

She is an associated professor in College of Mechanical Engineering in Chongqing university, China. She has published more than 20 papers. Her research area focuses on sustainable manufacturing and Intelligence manufacturing.

Runzhong Yi, Borned in Sichuan Province of China, earned master degree in mechanical engineering of Chongqing university of China in 1997.

He is a senior engineer in HTK system integrated Co. LTD in Chongqing, China. He has published more than 10 papers. His research area focuses on embedded system and measurement technology.

Fei Liu, Borned in Sichuan Province of China, earned Ph.D degree in mechanical engineering of Chongqing university of China in 1987

He is a professor in College of Mechanical Engineering in Chongqing university, China. He has published more than 100 papers. His research area focuses on systems engineering and Intelligence manufacturing.

Prof. Liu is member of ASME , member of editorial board of robotics and computer integrated manufacturing and concurrent engineering: research and applications

Chen Guorong, Borned in Sichuan Province of China, earned Ph.D degree in mechanical engineering of Chongqing university of China in 2010

He is an associated professor in College of Mechanical Engineering in Chongqing university of technology and science, He has published more than 10 papers. His research area focuses on systems engineering.

Research on Low Power Sigma-Delta Interface Circuit used in Capacitive Micro-accelerometers

Yue Ruan, Ying Tang and Wenji Yao
 Zhejiang Shuren University, Hangzhou, China
 Email: andyruan729@gmail.com; {635766060, 25428878}@qq.com

Abstract—Accelerometers have a wide range application in many fields, such as airbag deployment system and electronic stability control system in vehicles, and inertial navigation system in aircrafts and rockets. As Micro-Electro-Mechanical System (MEMS) technology advances, accelerometers are made smaller and smaller, often integrated in a single chip, with lower implementation costs and power consumptions. Modern MEMS accelerometers can be used in portable devices such as cell phones, RFID tags and even integrated in other sensors. This paper presents chip-level design and implementation of a second order sigma-delta interface circuit used in capacitive micro-accelerometers. The interface circuit provides 1-bit data stream and operates at a sampling frequency of 2.5MHz. The system model considers non-ideal factors in the circuit such as nonlinear distortion and noises. These non-ideal factors have been discussed through system level simulation in MATLAB. The micro-accelerometer, which is a highly integrated MEMS device, is then designed and implemented in silicon-on-insulator (SOI) substrate. Finally, the chip-level layout of interface circuit is implemented. Results have shown the chip with area of 1.32mm² and power consumption of about 5mW.

Index Terms—micro-accelerometer, sigma-delta, capacitive, interface circuit, layout

I. INTRODUCTION

Micro-Electro-Mechanical Systems (MEMS) are small integrated devices which are usually fabricated on a silicon substrate and able to combine electrical and mechanical elements for sensing or actuating purposes [1]. Examples of MEMS components include accelerometers, RF MEMS switches, microphones, and micro-resonators. With the fast development of modern MEMS technology, the micro-accelerometers, as the most mature MEMS-based inertial sensor application, have seen significant progress over the past decades. The advantages such as low-cost, low-power, small size, batch fabrication make micro-accelerometers have a wide range of applications, such as automotive safety and stability, biomedical applications, oil and gas exploration, and computer accessories [2]. Take the automotive as example: the micro-accelerometer is used in airbag deployment system to make sure the airbag bounces out when collision occurs (High-G acceleration). In modern vehicles, electronic stability control system (ESC) is often installed. One of the most important sensors in ESC is the accelerometer. A group of accelerometers measure

vehicle accelerations in each direction and send data to the control system to increase safety. Besides, modern consumer electronic devices usually use micro-accelerometers in game control. The hard disks in PCs also have micro-accelerometers to detect free-fall and help prevent data loss. Current micro-accelerometers based on MEMS technology have the highest degree of integration, with sensing elements and electronic interface circuitry integrated on a single chip together [3]. The high integration makes micro-accelerometer tiny and consumes little power, which can satisfy requirements in many applications.

Capacitive MEMS accelerometers have been implemented using various surface and bulk micromachining technologies. Unlike bulk micromachining, which defines structures by selectively etching inside a substrate (wafer), surface micromachining creates structures on top of a substrate by using a succession of thin film deposition and selective etching [4]. In MEMS devices, the thickness of the deposited layer and hence the proof mass is small, result in limitations on the performance of the accelerometers. Typically, the resolution of the commercial MEMS accelerometers is in the milli-gravity (mG) range [5]. On the other hand, bulk micromachining features larger proof mass and larger capacitive area that leads to higher sensitivity and higher resolution approaching micro-gravity (μG).

Currently, high performance mixed-signal interface circuits have received growing attention towards high-level of integration, power reduction and noise cancellation (improved resolution). The new generation of accelerometer interface architecture should have the versatility of interfacing with sensors of various sensitivities while maintaining low power consumption, small drift, increased functionality, and large dynamic range.

Silicon-on-insulator (SOI) technology has been widely applied to automotive industry. One of the most obvious advantages of SOI is the reduction of parasitic capacitances compared with conventional silicon (bulk CMOS) due to isolation from the bulk silicon, which improves power consumption at matched performance. In addition, SOI provides resistance to latch-up due to complete isolation of the n and p well structures [6]. Therefore, power consumption is reduced, and faster operation can be achieved.

The objective of this work is to design and implement a second order - modulator to readout the MEMS SOI

accelerometer. Sigma-delta modulators are often used as interface circuit due to its wide dynamic range, inherent linearity and relaxed accuracy requirements on the analog circuits. The accelerometer described in this paper is designed with the SOI technology. The fully differential 2nd order switched capacitor - modulator as interface circuit is implemented using 1 μ m, 5V SOI-CMOS process. The modulator performs well in simulation. The peak Signal-to-noise and distortion ratio (SNDR) is about 70 dB (minimum capacitance resolution of 15 aF) in a 5 kHz signal bandwidth from -40°C to +150°C. The chip area is 1.32 mm² with power consumption of about 4.8 mW. The system level simulation is done in Matlab, and Cadence software environment is used for circuit design and layout.

II. SYSTEM MODELLING

In common accelerometers, a mechanical sensing element converts the unknown quantity of acceleration into a displacement that is then detected and converted to an electrical signal output. The simplified schematic of a fully-differential capacitive MEMS accelerometer is depicted in Figure 1 below [7]. The central part of the accelerometer is a micromechanical proof mass M suspended to a supporting frame by mechanical springs with effective spring constant K , which acts as the sensing element. The squeezed film damping D is imposed by the surrounding air on the structure. The accelerometer has a fully differential sense topology; it means that four sense electrodes with one common node at the proof mass are devised in the fabricated MEMS accelerometers. C_{S1} , C_{S2} , C_{S3} , and C_{S4} are sensing capacitors between the proof mass fingers and the four sense electrodes, respectively.

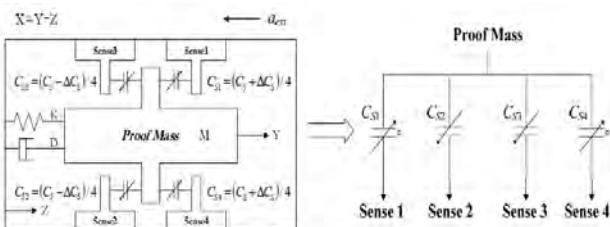


Fig. 1 Schematic of a fully-differential capacitive accelerometer

From Figure 1, C_{S1} , C_{S2} , C_{S3} and C_{S4} has the following relationships:

$$\begin{aligned} C_{S1,4} &= (C_S \pm \Delta C_s)/4 \\ C_{S2,3} &= (C_S \mp \Delta C_s)/4 \end{aligned} \quad (1)$$

Where C_S is the rest capacitance at zero acceleration and ΔC_s is the capacitance variation of the sensor at acceleration a_{ext} .

When an external acceleration a_{ext} is applied, the proof mass will move along with the sensing axis with respect to the moving frame of reference ($X = Y - Z$), causing a change in distance between it and the adjacent fixed electrodes. The displacement of the proof mass can be measured as a very small change in capacitance between

it and the fixed electrodes. Figure 2 shows a scanning electron microscope (SEM) image of how this is implemented in silicon. In figure 2, the movable proof mass fingers and fixed sense fingers generate capacitances. When no acceleration detected, ΔC is zero. As external acceleration increases, ΔC follows. By integrating the interface circuit on the same chip with the sensor, extremely small changes in capacitance (ΔC_s , in aF level) can be detected, which means small accelerations can be detected.

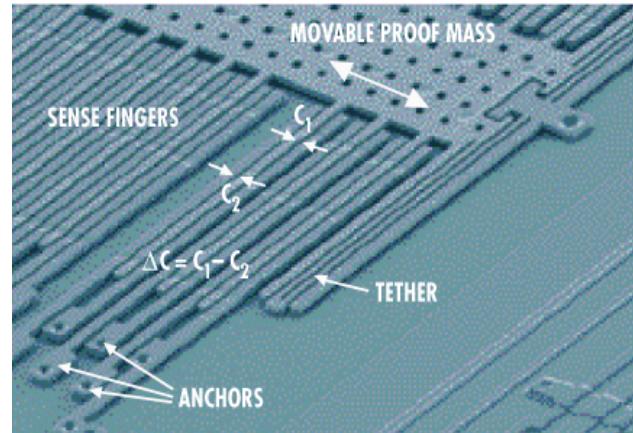


Fig. 2 SEM image of the sensor element^[16]

The schematic as shown in Fig. 1 shows the mechanical parameters for the sensing element. According to Newton's law, the differential equation for the displacement x as a function of external acceleration a_{ext} is that of a second-order mass-spring-damper system [1]:

$$M \frac{d^2x}{dt^2} + D \frac{dx}{dt} + K_{eff} x = F_{ext} = Ma_{ext} \quad (2)$$

Where, D and K_{eff} are the damping coefficient and spring constant, respectively, and linear relations are assumed. Taking Laplace transform of equation 2 yields the second order equation:

$$(Ms^2 + Ds + K_{eff}) \cdot X(s) = M \cdot A(s) \quad (3)$$

Solving for $X(s)$ gives the transfer function:

$$\frac{X(s)}{A(s)} = \frac{1}{s^2 + \frac{D}{M}s + \frac{K_{eff}}{M}} = \frac{1}{s^2 + \frac{\omega_r}{M}s + \frac{\omega_r^2}{M}} \quad (4)$$

With the resonant angular frequency $\omega_r = \sqrt{K_{eff}/M} = 2\pi f_r$ and quality factor $Q = \sqrt{K_{eff}M/D}$. When the system frequency is well below resonance frequency, that is $\omega \ll \omega_r$, the displacement value $x \approx a / \omega_r^2$. The system sensitivity is expressed as:

$$S = x / a \approx 1 / \omega_r^2 \quad (5)$$

This relationship on sensitivity states that there is a tradeoff between the system bandwidth and sensitivity of sensor. As the bandwidth of system increases, ω_r also increase, which results in low sensitivity of the system. On the other hand, low resonant frequency results in large displacements and high sensitivity, but restricts the bandwidth of the sensor.

In this design, an open-loop system is used, as it is simple in hardware, consumes less power than a closed-loop system and fits into automotive stability system application, because such application does not impose requirements on linearity and bandwidth. The system is depicted in Figure 3, which consists of a sensor functioning as the acceleration-to-displacement (capacitance) converter, and a position readout circuit generating the output voltage.

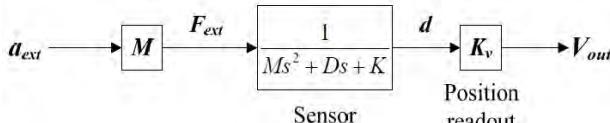


Fig. 3 Schematic of an open loop system

The architecture of a 2nd order Σ-Δ modulator as an interface circuit for accelerometers is shown in Figure 4. The 1st integrator also acts as a capacitance-to-voltage (C/V) converter for the accelerometer.

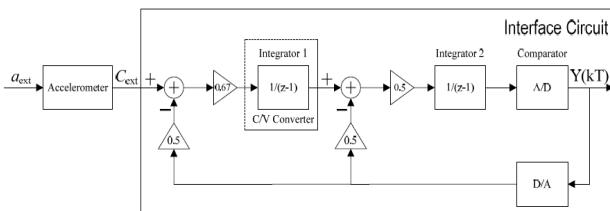


Fig. 4 A 2nd order Σ-Δ modulator as the interface circuit

III. DETAILED INTERFACE CIRCUIT DESIGN

This part of the paper illustrates the detailed interface circuit design of the Σ-Δ ADC using 1 μm 5V SOI-CMOS technology. In the detailed interface circuit design, the schematics of integrators, with 1st and 2nd order are presented first, and other required modules including op-amps, comparator, band-gap reference and clock generator are designed and simulated. Finally, the layout is carried out.

The Σ-Δ modulator as interface circuit is shown in Figure 5 [8].

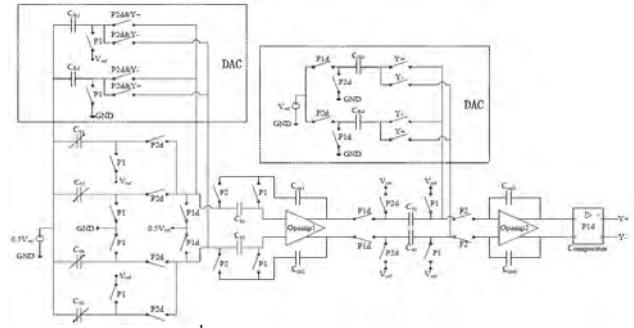


Fig. 5 The 2nd order Σ-Δ ADC as interface circuit

A. Integrators

The integrators have a significant impact on the performance of the Σ-Δ modulator. Some safety margins are taken in the implementation of Σ-Δ modulator to assure integrator performance. Schematic of the 1st and 2nd integrators are shown in Figure 6 and 7 below.

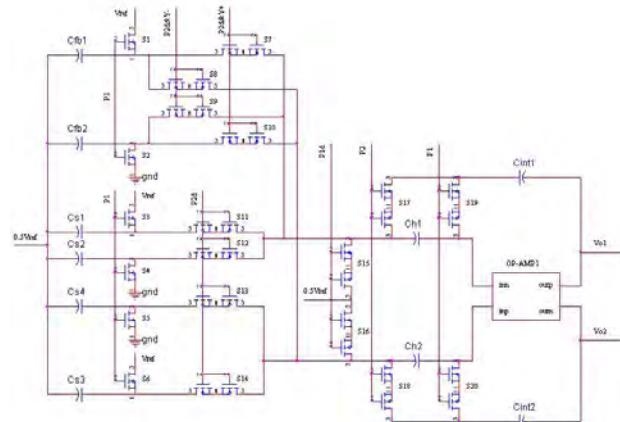


Fig. 6: Schematic of the 1st integrator

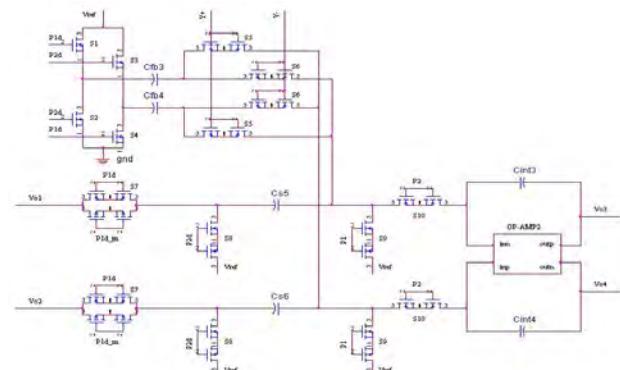


Fig. 7 Schematic of the 2nd integrator

B. Operational Amplifiers

The operational amplifier used in integrators is the most important component of the modulator. In order to suppress harmonic distortion, the op-amps should have enough DC gain. Besides, they should have sufficient slew-rate and large bandwidth to allow fast settling response within the available period. The need for high speed, large bandwidth, coupled with a relatively modest gain requirement of 70 dB to suppress harmonic

distortion, encouraged the use of the fully-differential folded-cascode op-amp as shown in Figure 8^[9].

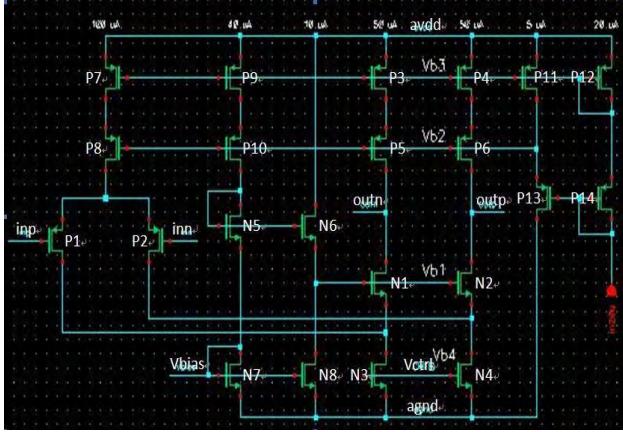


Fig. 8 Folded-cascode op-amp with biasing circuits

C. Comparators

The second major component of the modulator is the quantizer. The one-bit quantizer is realized with a dynamic comparator and a SR latch as shown in Figure 9^[10] and Figure 10 respectively. The function of the comparator in a sigma-delta modulator is to quantize a signal in the loop and provide the digital output of the modulator. The structure and operation are explained as follows.

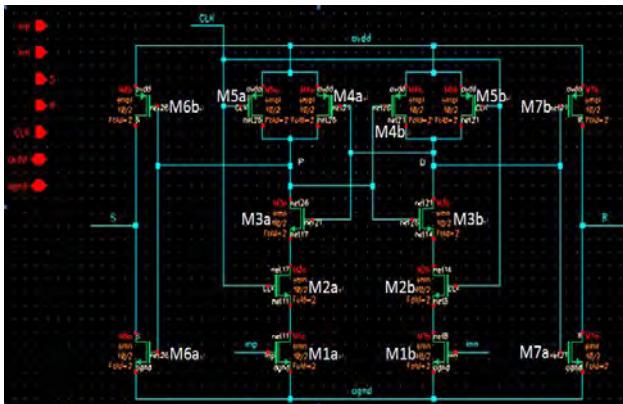


Fig. 9 Schematic of the dynamic comparator

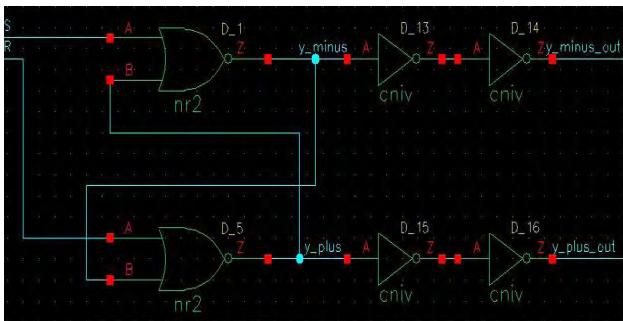


Fig. 10 Schematic of the SR latch and buffers

For a single-bit $\Sigma\Delta$ modulator, the requirement for the quantizer is quite relaxed as non-idealities such as the comparator offset and hysteresis in this stage can be

largely suppressed in the baseband by the second-order noise shaping. The comparator outputs are buffered by digital inverters (CNIV) and then recorded by off-chip data acquisition system for testing. All the transistors are of the size $10 \mu\text{m}/2\mu\text{m}$.

IV. SIMULATION RESULTS

The modulator depicted in Figure 3 is simulated using transistor-level models to evaluate the performance on signal transfer function, quantization noise shaping and distortion, etc. The transistor level is the bottom level in IC design, with the most complicated models and the most precise simulation results. The input signal in the proposed model is $\Delta C_{\text{smax}} = 0.039 \text{ pF}$ at 1.6785 kHz . The output signal swings of the 1st integrator and the 2nd integrator are depicted in Figure 11. Their envelopes follow ΔC_{smax} . They are both within $\pm 2\text{V}$, which fits the op-amp output range ($\pm 3\text{V}$) without significant harmonic distortions.

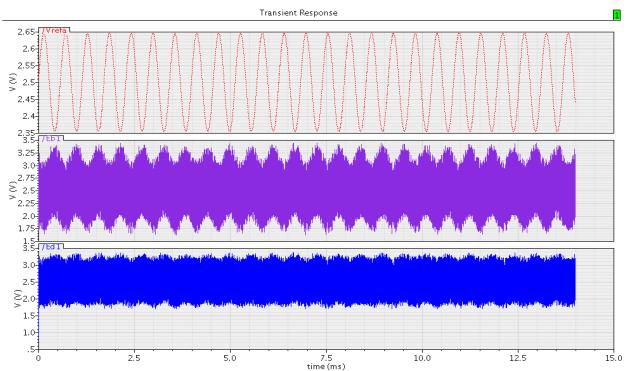


Fig. 11 Output swings of the 1st integrator (middle) and the 2nd integrator (bottom).

Figure 12 shows the two-level digital single-bit stream at the output of the modulator in time domain. The duty cycle of the output pulses follows the input signal ΔCs .



Fig. 12: Output single-bit stream of the modulator and input sinusoidal signal

In order to increase the spectral resolution of a Fast Fourier Transform (FFT), coherent sampling technology is used which is one of the most useful techniques for evaluating the dynamic performance of high-speed ADCs. Coherent sampling describes the sampling of a periodic signal, where an integer number of cycles fit into a

predefined sampling window^[11]. Figure 13 and Figure 14 illustrate the output spectrum of the modulator in frequency domain without (Figure 13) or with (Figure 14) DC offset in the 1st op-amp respectively.

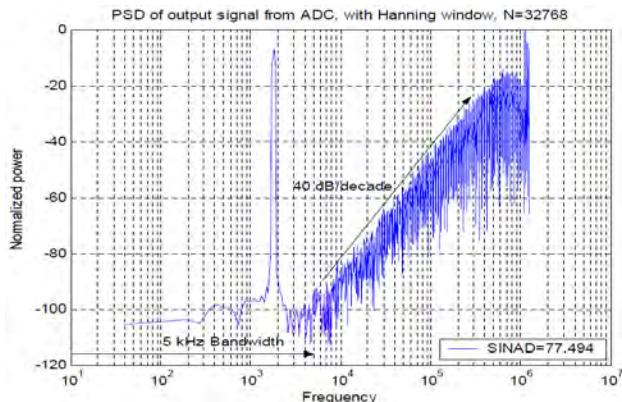


Fig. 13 Output spectrum of the modulator (without DC offset), SNDR = 77.49 dB

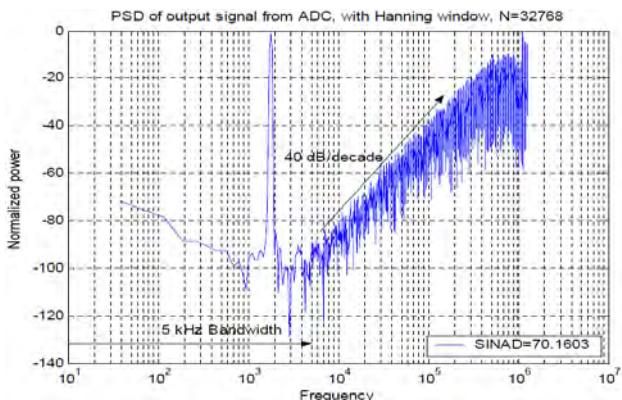


Fig. 14 Output spectrum of the modulator (with DC offset), SNDR = 70.16 dB

From the simulation results in figures above, we can get some comments and conclusions:

1) Since the device models used in transient analysis are noiseless, the output power spectrums in Figure 13 and Figure 14 do not include the contributions of switch noise, op-amp thermal noise and flicker noise, but only contain quantization noise, harmonic distortions and numerical errors due to accuracy of finite in-band FFT bins, simulator algorithm and device models, etc.

2) Both figures show the 2nd order quantization noise shaping of 40 dB / decade, and most of the quantization noise moves to higher frequencies outside the signal bandwidth.

3) In both figures, significant harmonic distortion is not observed within the signal bandwidth of 5 kHz. Therefore, it can be inferred that the implementation of the op-amps, the capacitors, the switches and the DACs have sufficient linearity.

4) In Figure 13, the signal to noise-and-distortion ratio (SNDR) is 77.49 dB when DC offset is not introduced, and it is 70.16 dB in Fig 11 when DC offset is added. Compared to the result from system-level simulation, which is 82.1 dB for the ideal modulator, one reason is

the finite DC gain of op-amps degrades the attenuation of quantization noise in the signal bandwidth. The other reason is the limited accuracy of the simulator.

5) In Figure 14, the FFT bin at the lowest frequency is corresponding to frequency smearing of DC offset by Hanning window, and its power is -73dBr. The auto-zeroing technology effectively reduces the op-amp DC offset together with the flicker noise.

V. HARDWARE IMPLEMENTATION AND PERFORMANCE TEST

The layout of the interface circuit is derived via Cadence tools, which is depicted in Figure 15 below. The total chip area is 1.2mm×1.1mm. The blank region is filled by decoupling capacitors for supply voltage and reference voltages, which is not shown for clarity. Separate bond-pads for analog and digital supply voltage are added. On-chip sampling and feedback capacitors C_{S1-4} and C_{fb1-2} (0.17 pF each) in the 1st integrator, together with their enable/disable pins, are implemented to make it possible to test the modulator without or with an accelerometer.

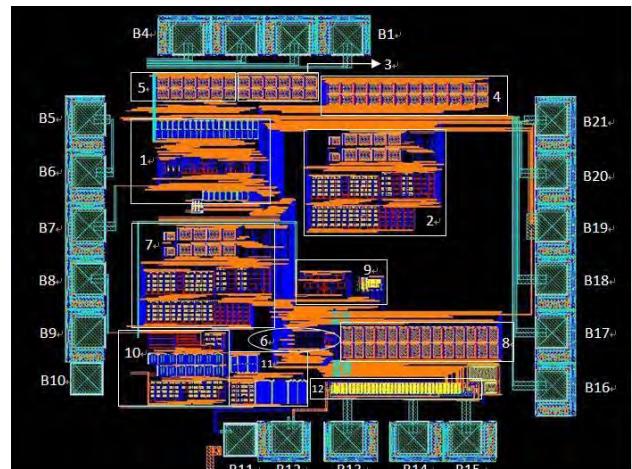


Fig. 15 Layout of the interface circuit

Table 1 below lists the system specification of the proposed MEMS accelerometer. The accelerometer has a full scale measurement acceleration range of 5G. When in 5G acceleration, the capacitance variation of the sensor reaches its full scale value of 0.039pF (39fF), which results the 7.8fF/G capacitive sensitivity. Besides, the rest capacitance C_S is 0.67pF.

Table 1: Specifications of the proposed MEMS accelerometer

Proof mass dimensions ($L \times W \times T$)	0.5 mm × 0.6 mm × 1.5 μ m
Proof mass M	9.011×10^{-10} kg
Rest capacitance C_s	0.67 pF
Full scale ΔC_{\max}	0.039 pF
Full scale a_{\max}	5 G
Capacitive sensitivity S	7.8 fF/G
Mechanical noise floor	77.53 μ G/ $\sqrt{\text{Hz}}$
Capacitive gaps $d_{\text{narrow}}/d_{\text{wide}}$	0.5 μ m/1 μ m
Quality factor Q	0.92
Resonance frequency f_r	5 kHz
Mechanical bandwidth $f_{-3\text{dB}}$	4.6 kHz

Table 2 below summarizes the performance of the accelerometer implemented, which is based on $1\mu\text{m}$ SOI-CMOS technology.

Table 2: Summary of the accelerometer performance

Rest capacitance C_s	0.67 pF
Full scale ΔC_{\max}	0.039 pF
Signal bandwidth	DC to 5 kHz
Oversampling ratio OSR	250
Sampling frequency f_s	2.5 MHz
Peak SNDR	70 dB
Resolution ΔC_{\min}	15 aF
Supply voltage	5 V
Power consumption	5 mW
Chip area	1.32 mm ²
Temperature range	-40°C to +150°C
Technology process	1 μm SOI-CMOS

VI. CONCLUSION AND FUTURE WORKS

MEMS devices especially sensors and micro motors are widely used in applications ranging from automotive safety and stability systems to biomedical applications, oil and gas exploration, computer accessories, etc. The objective of this work is to design and implement an interface circuit for SOI accelerometers in automotive stability systems. This work has focused on the analysis of accelerometer system and the design of a low-power high-resolution interface circuit in top-down design methodology.

The principle of the accelerometer model indicates that sensor capacitance changes in response to external acceleration. Accelerometer systems are classified into open-loop and closed-loop systems, depending on whether a force feedback loop is applied to the sensor. In this design, open-loop structure is used as it is simple and effective for the application. The specifications of SOI accelerometers and its readout circuitry have been proposed. A fully differential sigma-delta modulator is adopted in this project due to its wide dynamic range, inherent linearity and relaxed accuracy requirements on the analog circuit.

The 1st integrator is a crucial component in the design of a 2nd order sigma-delta modulator. It acts as a capacitance-to-voltage converter and determines the noise and distortion performance of the modulator. Based on noise analysis, auto-zeroing technique is applied to reduce flicker noise and DC offset, achieving high resolution.

The specifications on circuit blocks are easily established by modeling and simulating non-idealities of the $\Sigma-\Delta$ modulator, such as sampling jitter, kT/C noise and operational amplifier parameters (noise, finite DC gain, finite bandwidth and slew rate, and saturation voltages) on the system-level in MATLAB.

An experimental $\Sigma-\Delta$ modulator that fulfills these specifications is implemented using $1\mu\text{m}$ 5V SOI-CMOS technology. This includes implementation of the integrators, op-amps, comparator, band-gap reference, clock generator and layout. In the simulation, the

modulator performs well in a 5 kHz bandwidth from -40 °C to +150 °C, and has a peak SNDR of 70 dB which corresponds to 11-bit resolution and minimum capacitance resolution of 15 aF. The chip area is 1.32 mm² with power consumption of approximately 5 mW.

Regarding to future works, this design can be extended to further lower voltage operation (i.e., 3.3 V) despite the large threshold voltage of transistors (approximately 1 V). The reference voltage can be tied to supply voltage for better resolution. Some trade-offs can be made between the power dissipation and resolution, to allow the use of other ADCs such as successive approximation registers (SAR) as interface circuit for accelerometers.

This work has been verified by CADENCE, which is a mixed-mode simulation tool. The transistor-level simulation is very time consuming. In recent years, the mixed-signal circuit design is more and more popular. In order to speed up the period of design flow path, using other modeling level simulation tools such as Verilog-A is an important method. Besides, demonstrating the functional operation and performance metrics of the designed interface circuit by taping out and measuring is another important procedure in the future.

The future research interest would be to employ a closed-loop system. Force-feedback may improve many characteristics of a sensor including bandwidth, dynamic range and linearity. The most popular approach is to pulse-modulate the force-feedback signal [12]. High-order closed-loop system (i.e., 2nd order sensor element and 2nd order electronic filter [13]) may result in low quantization noise.

REFERENCES

- [1] B. V. Amini, "A Mixed-Signal Low-Noise Sigma-Delta Interface IC for Integrated Sub-Micro-Gravity Capacitive SOI Accelerometers", *Doctoral Thesis*, Georgia Institute of Technology, May 2006.
- [2] G. Zhang, "Design and Simulation of A CMOS-MEMS Accelerometer", *Master Thesis*, Carnegie Mellon University, May 1998.
- [3] W. Yun, "A surface micromachined accelerometer with integrated CMOS detection circuitry", *Doctoral Thesis*, U.C. Berkeley, 1992.
- [4] http://en.wikipedia.org/wiki/Bulk_micromachining.
- [5] B. V. Amini and F. Ayazi, "A 2.5-V 14-bit $\Sigma\Delta$ CMOS SOI Capacitive Accelerometer", *IEEE J. Solid-State Circuits*, vol. 39, no. 12, pp. 2467–2476, Dec. 2004.
- [6] http://en.wikipedia.org/wiki/Silicon_on_insulator.
- [7] B. V. Amini and F. Ayazi, "Micro-gravity Capacitive Silicon-on-Insulator Accelerometers", *J. Micromech. Microeng.*, vol. 15, pp. 2113-2120, Sep. 2005.
- [8] Y. Yu, A Sigma-delta ADC as Interface Circuit for SOI Accelerometers, *Master Thesis*, TU Delft, 2005.
- [9] B. Razavi, "Design of Analog CMOS Integrated Circuits", *McGraw-Hill*, 2001.
- [10] T. B. Cho and P. R. Gray, "A 10 b, 20 M sample/s, 35 mW pipeline A/D converter", *IEEE J. Solid-State Circuits*, vol. 30, no. 3, pp. 166–172, Mar. 1995
- [11] http://www.maxim-ic.com/appnotes.cfm/an_pk/1040/, "Coherent Sampling vs. Window Sampling", Application Note 1040, *Maxim-IC*, Mar. 2002.

- [12] A. M. Lemkin, B. E. Boser, "A three-axis micromachined accelerometer with a CMOS position-sense interface and digital offset-trim electronics", *IEEE J. Solid-State Circuits*, vol. 34, no. 4, Apr. 1999.
- [13] V. P. Petrov and B. E. Boser, "A fourth-order sigma-delta interface for micromachined inertial sensors", ISSCC 2004, pp. 320-329, 15-19 Feb. 2004.
- [14] Kulah. H., Chae. J., Yazdi. N. and Najafi. K, "Noise analysis and characterization of a sigma-delta capacitive microaccelerometer", *Solid-State Circuits, IEEE Journal of*, vol.41, no.2, pp. 352-361. Feb.2006.
- [15] Jeongjin Roh, etc. "A 0.9-V 60- μ W 1-Bit Fourth-Order Delta-Sigma Modulator With 83-dB Dynamic Range", *Solid-State Circuits, IEEE Journal of*, vol.43, no.2, pp. 361-370. Feb. 2008
- [16] <http://www.micromagazine.com>



Yue Ruan received the B.S degree in electronic engineering from Zhejiang University, Hangzhou, China, in 2006, and the M.Sc degree in electrical engineering from Royal Institute of Technology (KTH), Stockholm, Sweden, in 2008.

He is currently a Faculty Member with the college of Information Science and Technology, Zhejiang Shuren University, Hangzhou, Zhejiang, China. From 2008 to 2009, he worked as a research engineer in Motorola Inc. He has published more than 10 technical papers and a text book entitled *Microcontroller Theory and Applications – Based on sample driven and Proteus simulation* (China, Science Press, 2011), of which he is the co-author. His current research involves wireless sensor networks, MEMS sensors, EDA based VLSI design and RF integrated circuits.

Non-linear Multi-attribute Based Online Auction Bidding Model and Platform

Jie Lin, Jinghua Zhao

(School of Economics and Management, Tongji University, Shanghai 200092, China)

Email: sunnyjh.love@yahoo.com.cn, jielinf@163.com

Ling Xue

Department of Operations and Information Management, University of Scranton,

Scranton, PA 18510, USA

Email:xuel2@scranton.edu

Abstract—To build a new bidding model of multi-attribute online auctions, this paper introduces the concepts of multi-type attribute and nonlinear function. The model can measure the effect of the attributes in auctions by adopting the normalized method of weighting and it will set the stage for the purchasers to evaluate suppliers in the procurement management. Furthermore, this paper developed an application platform system of nonlinear multi-attribute online auction (which is based on B/S) using the Struts framework developed J2EE. Finally, the related experiments and practices were conducted.

Index Terms—auctions/bidding; multi-attribute; non-linear; auction model

I. INTRODUCTION

Multi-attribute auction is different from traditional single-attribute auction. In traditional auction, the competition focus is mainly on the single-attribute—price, though the modes and the relationships between the counterparts of auctions vary differently. However, as values of many substitutes do not solely depend on price, it is unilateral to consider the price only. Especially in the procurement auction between enterprises, attributes other than price always yields the same importance. Scholars usually use *multiple-attribute auction* to refer to the auction in which the buyers consider more attributes besides price.

Traditional online auction has been studied for a long time. However, the importance of multi-attribute online auction has attracted more attention in recent years. Many scholars have studied the models of multi-attribute online auctions and mainly focused on the profits generated for

buyers and sellers, the bidding strategies and the set of multi-attribute weight. For example, Vulkan and Jeenings focused on the specific problem of service allocation among autonomous, automated agents[1]; Bitchler conducted a practical analysis on various kinds of auctions and found that more utilities can be generated from multi-attribute auction than from single-attribute auction[2]; Sairamesh studied the problem of service allocation among agents using multi-attribute English auction protocol[3]; C.C.Lee and C.OU-Yang investigated the negotiation efficiency for the various bidding strategies the demander employed in different order conditions[4]; Lee put forward an algorithm of multi-attribute computer negotiation; one of their contributions is the two-phases process which solves the most important attributes first and then the rest[5]. Recently, researchers in China begin to study this area and the majority of them are still concentrating in the analytical studies instead of practical applications of multi-attribute auctions.

In the practice of enterprise procurement, it is common that a single buyer deals with multiple suppliers. By introducing multi-attribute auction in procurement, we can achieve benefits such as full competition, optimized resource allocation, supply chain improvement, full development of suppliers' potential as well as avoiding collusions among suppliers to some extent. However, due to the requirements on both technology and personnel, few applications can implement multi-attribute online procurement auction.

To improve the feasibility and adaptability of the multi-attribute procurement auction, we introduce non-linear function into the auction, put forward a non-linear multi-attribute-based online procurement auction bidding model and design the implementation of the auction platform based on the model.

II. NON-LINEAR ULTI-ATTRIBUTE BASED ONLINE PROCUREMENT AUCTION MODEL

National Natural Science Foundation, China (No. 71071114), Doctoral Fund of Ministry of Education of China (No. 200802470009), Shanghai Leading Academic Discipline Project (No. B310).

Corresponding author address: School of Economics and Management, Tongji University, Shanghai, 20092, P. R. China. Tel: +86 13918288246. E-mail: sunnyjh.love@yahoo.com.cn

This model is developed from Bichler's weighted sum model, and then we introduce the concept of multi-type attributes and non-linear function into the model. In this model, different methods are used to calculate the values of different types of attributes. Finally, the weight sum of the normalized attribute value is used to calculate a comprehensive score, which is used to determine the auction winner.

A. Assumed Condition

Prior to the establishment of the model, we have the following assumptions:

- 1) A buyer has a number of suppliers to choose, each supplier's ability to provide products varies.
- 2) When bidding, suppliers compete with each other on a limited number of product attributes Y_i ($i > 1$) designated by the buyer. Each bid contains the required values of all attributes (compared to the traditional auctions each bid contains only a single attribute value, the "price") and regards the comprehensive score of multi-attribute as the final evaluation of the bid.
- 3) Suppliers do not know each other's bids but their own current ranks.
- 4) Evaluate each bid from suppliers and offer a comprehensive score, at the end of the auction, the supplier with the highest score wins.

B. Non-linear Multi-attribute Based Online Procurement Auction Model

(1) The handling of multi-type attributes

Suppose that there are j suppliers participating in this auction, and they need to bid on i attributes of the product, then

$$Y_{ij} = f_i(x_{ij}) \quad (1)$$

where Y_{ij} represents the i^{th} attribute score obtained by the j^{th} supplier; x_{ij} represents the bidding value of the i^{th} attribute submitted by the j^{th} supplier; f_i represents the scoring rule applied to the i^{th} attribute, which is determined by the buyer.

In the traditional single-attribute auction, we only compare prices among suppliers and determine the ranking (namely $Y_j = x_j$, rank Y_j). In multi-attribute auction, since the meanings of the attribute values are different (such as the values of price and accounts payable period: the lower the price and the longer the accounts payable period, the better for the buyer. In addition, the orders of magnitude of these two attributes are also different), we cannot simply compare them based on the values. Instead, we use functions to convert the bidding into proper values.

The attributes of a product can be various, but we can divide the attribute types into three categories, namely,

the numeric, interval and text. This paper deals with the three types respectively as follows:

① The numeric type

The numeric type attribute means that the attributes can be directly described by numbers, such as price, delivery time (days) and quantity.

In practice, since the order-of-magnitude of each number type attribute is different, it is necessary to normalize the value of each attribute to facilitate the set of functions, thus

$$X_{ij} = \frac{x_{ij} - x_{i\min}}{x_{i\max} - x_{i\min}} \quad (2)$$

where x_{ij} represents the actual bid value on the i^{th} attribute submitted by the j^{th} supplier; $x_{i\min}$ is the minimum value of the i^{th} attribute; $x_{i\max}$ represents the maximum value of the i^{th} attribute; X_{ij} represents the normalized value of i^{th} attribute. Thus, we can limit the value of each numeric type attribute in the range of 0 to 1.

Then we substitute X_{ij} into the function f_i to control the final comprehensive score, as well as normalize the value of the function to make the weight set effective, thus

$$F_i(X_{ij}) = \frac{f_i(X_{ij}) - f_{i\min}}{f_{i\max} - f_{i\min}} \quad (3)$$

where $f_i(X_{ij})$ represents the function value after the substitution of X_{ij} ; $f_{i\min}$ represents the minimum value of the function; $f_{i\max}$ represents the maximum value of the function (When the function is an increasing function, then $f_{i\min} = f_i(0)$, $f_{i\max} = f_i(1)$ and vice versa).

In this model, we can use simple linear function, such as $f(x) = ax + b$ and $f(x)$, or more widely used non-linear function such as S shape function $f(x) = \frac{1}{a + be^x}$ and so forth. In linear multi-attribute auction, since the change range of the function trend is fixed and not related to the bid value, a supplier may submit an extreme value to take the advantage (or disadvantage) in specific area to enhance the comprehensive score abnormally, thereby undermining the principles of scientific evaluation. Non-linearizing the multi-attribute auctions can well introduce the economics concept of "marginal utility", and then an extremely high (low) value will lead to even smaller (greater) income (loss). Therefore, it can balance the various attribute values (For example, when the delivery time of a product

part ranges from 1 to 4 days, and the supply of the other parts has not kept pace with it, there is no additional income for the buyer no matter how fast this kind of part can be supplied. Thus, when the value of this attribute ranges from 1 to 4, the evaluation score of this attribute will not change a lot; when the value ranges from 5 to 9, it will accelerate the decline of score; and when the value is bigger than 10, maybe due to the best sales time has passed, the score will be directly 0). In addition, the introduction of non-linear function also brings the buyers greater flexibility - their benefit can be maximized by getting guidance on suppliers' bids and adjusting the function slope according to their own needs or preferences.

②The interval type

To make the model more applicable, we divide the value range of an attribute into intervals described by using different functions. The precondition is that the sub-functions must have the same monotonicity, called interval type attribute.

Suppose the sub-function is a monotone increasing function, then

$$F_i(X_{ij}) = \frac{f_i(X_{ij}) - f_{\text{first min}}}{f_{\text{last max}} - f_{\text{first min}}} \quad (4)$$

where $f_{\text{last max}}$ represents the maximum value of the last interval of the value range (namely, the maximum value of the function), $f_{\text{first min}}$ represents the minimum value of the first interval of the value range (namely, the minimum value of the function), and when the sub-function is a monotone decreasing function, then just exchange f_{last} and f_{first} . It is apparent that the number type attribute is a special form of the interval type attribute (when the attribute has only one interval).

Integrating numeric and interval type attribute, formula (1) can be converted to

$$Y_{ij} = F_i(X_{ij}) \quad (5)$$

③The text type

In practical, not every attribute can be described by numbers, such that buyers have different preferences on some attributes such as quality. For these attributes, we can make a number of text options for suppliers to select and give each option a value based on the following calculation, this is the so-called text type attribute. Suppose the r^{th} attribute is a text type attribute and has totally m options, then

$$Y_{rj} = \frac{x_{rj}}{\max\{x_{r1}, x_{r2}, \dots, x_{rm}\}} \quad (1 < m) \quad (6)$$

Where x_{rj} represents the value of the option of the r^{th} attribute selected by the j^{th} supplier; Y_{rj}

represents the score of the r^{th} attribute obtained by the j^{th} supplier; it is apparent that the score has been limited in the range from 0 to 1, so the normalization is not needed.

(2) Weight settings

After calculating the score of all the attributes, we assign each attribute a weight to reflect the preferences of buyers, then

$$W_i = \frac{w_i}{\sum_i w_i} \quad (7)$$

Where w_i represents the weight assigned to the i^{th} attribute by the buyer; W_i represents w_i 's proportion of the total weight, namely the real weight in the calculation of the comprehensive score.

(3) Determination of the winner

To sum up, we can draw

$$S_j = \sum_i W_i Y_{ij} \quad (8)$$

Where S_j represents the comprehensive score of one bid of the j^{th} supplier. By doing so, we can rank the suppliers based on the scores and determine the final winner.

According to condition (3) "Suppliers do not know each other's bidding but their own current ranking position," we can avoid to some extent of the vicious competition on one attribute among the suppliers. The final comprehensive score of each supplier reflects the comprehensive effect of all the values of attributes submitted. If the suppliers want to make their own score the highest, after full competition, there must be

$$S_j = S_{j \max} = \phi(x_{1j \max}, x_{2j \max}, x_{3j \max}, \dots) \quad (9)$$

Where $S_{j \max}$ represents the highest comprehensive score of the j^{th} supplier, $x_{1j \max}, x_{2j \max}, x_{3j \max}$ respectively represents the highest value of each attribute submitted by the j^{th} supplier, ϕ represents the scoring rule.

Finally, based on the ranking among the comprehensive scores, the supplier with the highest score wins, that is

$$S_n = S_{\max} = \max\{S_1, S_2, \dots, S_j\} \quad (1 \leq n \leq j) \quad (10)$$

where S_n represents the comprehensive score of the winner (suppose the n^{th} supplier is the winner), S_{\max} represents the highest comprehensive score, formula (9) and (10) mean that only the supplier who gives full play to its advantages on each field can be the final winner.

III. APPLICATION PLATFORM DESIGN

Since that the online procurement auction may involve many suppliers from different regions, we choose a more adaptable framework- B/S (Browser/Server) framework (Figure 1). In Windows operating system, we use the design mode, which is widely used in Web system development, MVC (Model, View and Controller). We use Struts of J2EE, which is a mature classic open-source application framework based on MVC, the application framework. By integrating Servlet, JSP, JavaBean, custom labels and information resources into a unified framework, Struts offers an MVC development mode which can be highly configured in Web-based system development.

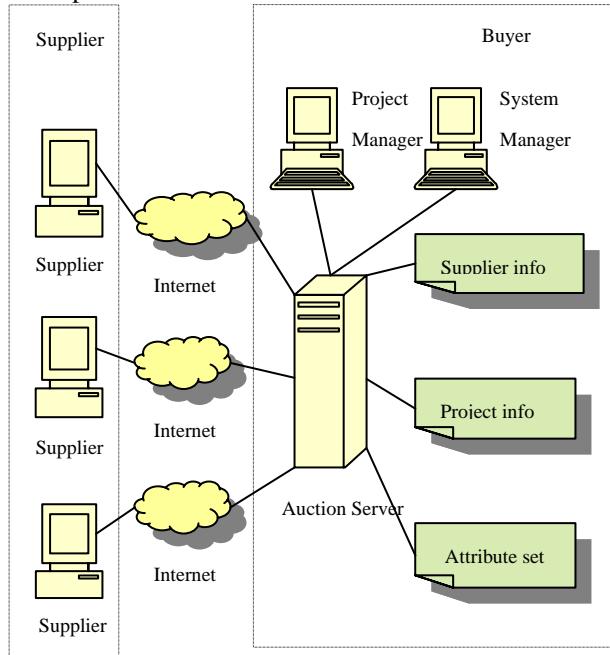


Figure 1 System platform architecture

A. System Function Design

Based on the bidding model, non-linear multi-attribute online procurement auction system should involve three groups of users, including suppliers, project manager (the person who creates, manages and controls the auction project and is assigned by the buyer) and system administrator. We use unified modeling language (UML) to establish the use-case diagram (Figure 2).

The system uses role-based access control: users from three different groups use different URL address to login, and the system identifies them by user names and passwords and provides the appropriate interface. Among different group of users, the system administrator is responsible to audit the registration of new suppliers and other registrants and at the same time maintain the new project manager. Only the supplier or the project manager who has been audited by the system administrator can login to the system. The project manager is the host of the auction, who is mainly in charge of managing the procurement project, including the establishment of the project (set up the name of the project, the time to begin

and end and all the attributes, etc.), inviting the suppliers to participate in the auction (only the invited suppliers can see the project when login), amending the project (including the time to begin and end, attribute parameters, etc.), starting the project and reviewing the auction results. The suppliers can review the invited available procurement projects, select the project to participate in and amend their own registration information to get in touch with buyers.

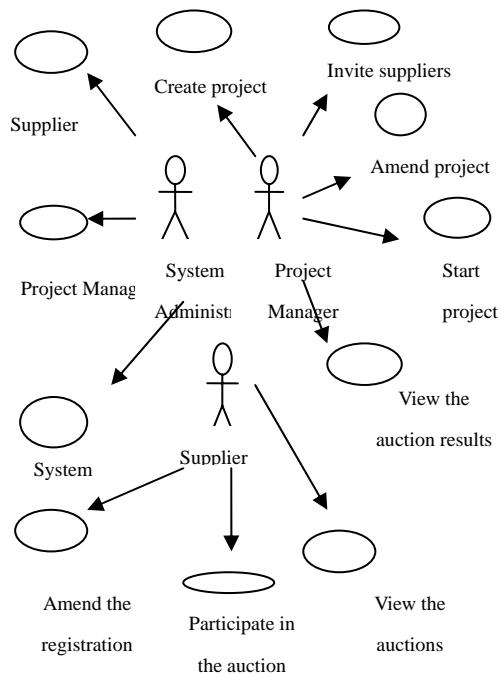


Figure 2 the UML of Multi-attribute online auction system

B. Bidding Process Design

The core function of the multi-attribute online procurement auction system is to allow multiple suppliers to bid online at the same time and the buyers to monitor the process of the bidding. So we put forward the following suppliers bidding process based on the bidding model (Figure 3).

After the auction begins, suppliers can submit their bids through web controls such as text boxes and drop-down lists. The system then provides comprehensive scores and the ranking position for suppliers. The position is updated every second (via AJAX, Asynchronous JavaScript and XML, ranking update will not cause the refresh of the whole web explorer and greatly enhance the user experience), so that the suppliers can bid in real time and response timely. If the supplier's bid ranks first, there is no need to respond until the end, unless the position dropped; if its ranking position fell behind, then the supplier has to decide whether to bid again or not. If yes, then the supplier should consider how to adjust the attributes values to enhance the comprehensive score, leading us to the first step in this cycle.

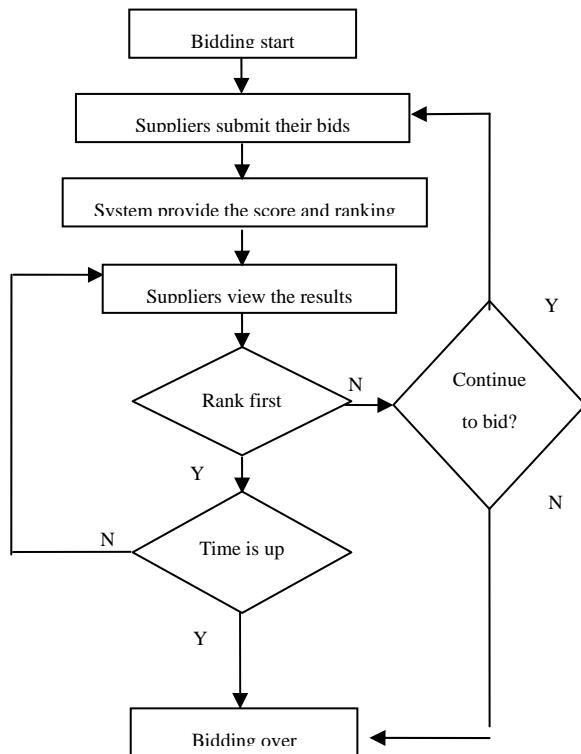


Figure 3 Multi-attribute online auction bidding process

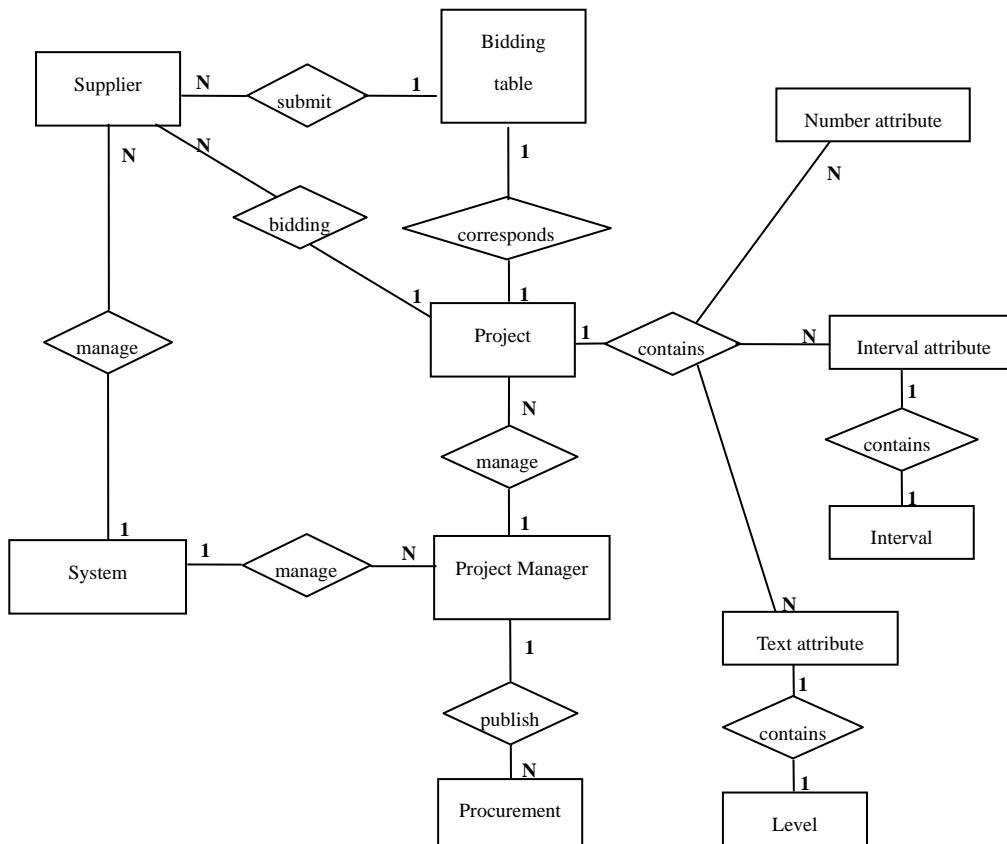


Figure 4 ER diagram of the multi-attribute online auction

C. Database Design

One of the difficulties of the system implementation is the database design. A project manager can create a number of procurement projects, each project contains a number of attributes, each attribute has a number of functions to use, and each function contains different parameters (we provide two kinds of functions, line function and S shape function). Database tables involved and the basic ER diagram is shown in Figure 4 (only show the relationship between the main entities).

The following is a brief description of the ER diagram, in which PK represents the primary key and FK represents the foreign key. The supplier table contains supplier ID (PK), user name, password, industry, region, company name, address, telephone, etc.; The system administrator table contains system administrator ID (PK), user name, password, name, telephone, etc.; The project manager table contains project manager ID (PK), user name, password, name, telephone, department, etc.; The procurement project table contains project ID (PK), project name, auction mode, project manager ID (FK), telephone, starting time, end time, project status, creating time, number of attributes, etc.

The relationship between the projects and the attributes is the main point in the database design. The attributes and projects relationship table (project ID (FK), attribute ID) is to indicate the relationship between the projects and the attributes, showing which attributes a project is related to. In addition, the number type attribute table—number type attribute ID (PK), attribute name, weight, minimum value, maximum value, used function and function parameter. In which the “number type attribute ID” column is the column of auto-generated code, and the same as the following attribute ID. The “minimum value” and the “maximum value” identify the value range of this attribute. In addition, the interval type attribute table (Interval attribute ID (PK), attribute name, weight, number of intervals) is used to save the basic information of the attributes. Since the number of intervals of each interval attribute is different, the relationship between the attribute and the intervals is a one-to-many relationship. We need another table to save the value of each intervals—the interval table, which includes interval type attribute ID (FK), interval ID, minimum value, maximum value, used function and function parameter. For each interval, value range is defined by a pair of “minimum value” and “maximum value”, and each interval has a corresponding used function. Same as the interval type attribute table, the text type attribute table (text attribute ID (PK), attribute name, weight, number of levels) is also used to save the basic information of the attributes. . Moreover, the level table (text type attribute ID (FK), level ID, level name, description, corresponding value) is to store the detailed information of levels. For each text level, we need to set the name and description to tell the demand of the buyers in detail. At last we need to set the corresponding values which are used to calculate the scores. The text type attributes do not have any functions, and the corresponding values equal the values that are calculated from the functions of the other attributes.

In addition, we can see that from the above, a project corresponds to a number of attributes and the interval or text type attribute corresponds to a number of intervals or levels. Therefore, it is impossible to record all the bids from all the projects in one table. So we use the bidding record table corresponding to one project and each table is generated automatically by the system. The table includes: supplier ID (FK), attribute ID 1, attribute ID 2, ……, comprehensive score, submit time.

IV. APPLICATIONS

Based on the system design, we complete the multi-attribute online procurement auction system development. We use the B/S structure and the Struts framework of J2EE in Windows operating system. We choose the HyperText Mark-up Language (HTML), Java Server Pages (JSP), Java and JavaScript as the programming languages and supplemented by AJAX, partial page update technology, use Eclipse3.1 as the

development environment and SQL server as the database. If the servers of the system are deployed by the buyer, the suppliers can just login the system and participate in the auction project at the scheduled time.

The main interface of the system consists of a title on the top, a menu on the left and the main part on the right. When the users from different groups login to the system, the menu bar will show different options accordingly. Figure 5 shows the menu when a project manager login: the main part shows the core process to create a new project, which is the setting of each attribute. In this example, the manager chooses price, delivery time and quality as the attributes, for the project, and number, interval and text as the attribute type. For number type attributes, the manager need to type in the value range and choose the corresponding function and the function parameters; for the interval type attributes, he need to select the interval number first and then set each interval as number type attributes; and for text type attributes, he also need to select the level number first, then name and describe each level and set the corresponding values respectively. When the attribute setting is completed, then the project manager can invite suppliers and start the project.

When the project starts, suppliers can login to the system and participate in the auction. Through textboxes (number type and interval type) and drop-down list boxes (text type), they can type in the bids and submit to the server. They also can see their rankings and bidding records in the page as reference for the next bid.

While suppliers are bidding, the project manager can monitor the competition and see the ranking via the system. When the bidding is over, the manager can contact the winner soon.

Since the remaining functions and the functions of system administrator are external functions of the system, we will not discuss them in this paper.

V. CONCLUSIONS

The multi-attribute online procurement auction can bring full competition, optimize resource allocation, improve supply chain process and develop suppliers' full potential. It is also an opportunity for the buyers to re-select suppliers. The competition of non-price attributes exposes suppliers' ability completely, so it is a good chance for buyers to transform the supply chain and an opportunity for the suppliers to adjust the production structure. By developing their own potential, transforming the industrial structure and improving the concept of cost, the suppliers can also pave the way for the future development.

By introducing the multi-attribute and non-linear functions, we have improved the bidding model and successfully completed the application platform, enhanced the feasibility and adaptability of the multi-attribute online procurement auction.

Non-linear multi-attribute auction theory can also be integrated into group decision support theory. Integrating

this application platform with

The S function: $y = \frac{1}{(a + be^{-x})}$		Linear function: $y = ax + b$	
Name	Value mode	Function used	Param a
Price	Number Type Range: 1 to 10000	Line Shape	a=1 b=2
Delivery Time	Interval Type The1th: >= 1 < 5 The2th: >= 5 < 10	Line Shape	a=0.6 b=1
Quality	Text Type The1th: good The2th: medium	Divide into 2 levels Description: **param>5 **param<=5	Value: 4 Value: 2
<input type="button" value="Next"/> <input type="button" value="Back"/>			

Figure 5 Core processes to create a project

group decision support system (GDSS) to build a new procurement platform and studying the procurement process, supplier management and incentive mechanism is also a new idea to solve the modern procurement problem. It also can aid the business decision-making and promote the supply chain coordination.

ACKNOWLEDGEMENTS

This work was supported by National Natural Science Foundation, China (No. 71071114), Doctoral Fund of Ministry of Education of China (No. 200802470009), Shanghai Leading Academic Discipline Project (No. B310).

REFERENCES

- [1] Vulkan N., & Jennings N. R.. Efficient Mechanisms for the supply of services in multi-agent environments. Decision support systems, 2000, 28, 5-19
 - [2] M.Bichler. An Experimental Analysis of Multi-attributes Auction. Decision Support System, 2000, 29(3):249-268
 - [3] J.Sairamesh, R.Mohan. A Platform for Business-To-Business Sell-Side, private Exchanges and Marketplaces. IBM Systems Journal, 2000, 41(2):242-252
 - [4] C.C.Lee & C.OU-Yang. Development and evaluation of the interactive bidding strategies for a demander and its suppliers in supplier selection auction market. International Journal of Production Research, 2008, 17: 4827-4848
 - [5] Lee.H.G. Electronic Brokerage and Electronic Auction: the impact of IT on market structure. Proceedings of the 29 th Hawaii International Conference on System Sciences, Maui, HI,1999, 4:397-406
 - [6] Y.K.Che. Design Competition through Multidimensional Auction. RAND Journal of Economics, 1993, 24(4):668-680
 - [7] R.P.McAfee, J.M. Auctions and bidding. Journal of Economic Literature, 1987, 7:699-738
- Jie Lin**, a Sichuan Quxian native (1967-), a Ph. D. graduated from Fudan University and a professor in School of Economics and Management at Tongji University, his research interests include management information system, electronic commerce and supply chain management, etc.
Email: jielinf@163.com
- Jinghua Zhao**, a Shandong Guanxian native (1984-), a Ph. D. candidate in School of Economics and Management at Tongji University, her research interests include management information system, electronic commerce and supply chain management, etc.
Email: sunnyjh.love@yahoo.com.cn
- Ling Xue**, a shanghai native (1975-), a Ph.D. graduated from University of Texas at Austin, and is a assistant professor in the university of Scranton. His Research and Teaching Interests include: E-Commerce, IT governance, IT value, IT security, E-Commerce, Database, IT management.
Email: xuel2@scranton.edu

Virtual Machine-based Intrusion Detection System Framework in Cloud Computing Environment

Huaibin Wang

Key Laboratory of Computer Vision and System, Ministry of Education Tianjin University of Technology,
Tianjin, China
Email: hbwang@tjut.edu.cn

Haiyun Zhou and Chundong Wang

Key Laboratory of Computer Vision and System, Ministry of Education Tianjin University of Technology,
Tianjin, China
Email:haiyun313@163.com

Abstract—Cloud computing an emerging approach by sharing infrastructure is an overwhelming trend. While in the process of cloud deployment, the security issues can not be underestimated. Traditional Intrusion Detection System (IDS) because of lower detection rate and higher false rate couldn't be suitable the cloud here. Extensibility is the main requirement for IDS framework of cloud environment in this paper as follows. First the cross-platform and strong isolation properties of virtualization have been fully reflected here, that is to say, an extensible VM-based multiple IDSs are deployed in each layer to monitor specific virtual component. Moreover, during the process, we also propose the cloud alliance concept by the communication agents exchanging the mutual alerts mainly to resist Denial-of-Service (DoS) and Distributed Denial-of-Service (DDoS) - the single point attack of failure. On this basis, we have the identity certification of the communication agents to improve the reliability of the alerts. Through the comparison of simulation results, the proposed system framework has a great advantage for monitoring VMs on the detection rate.

Index Terms—cloud computing, VM-based IDS, cloud alliance, communication agent, detection rate

I. INTRODUCTION

Cloud computing the new IT concept is being spread rapidly around the world. Cloud computing is named by the model of service delivering and using. It is regarded as on-demand scalable way to obtain the necessary service which can be Internet-based software services, bandwidth services or any other services. This is our so called cloud. The services provided by cloud computing could be divided into three types:

1) Cloud Computing Software as a Service (SaaS) provided by the cloud service provider could be accessed by the interfaces of a variety of clients. The underlying infrastructure of the cloud including networks ,servers, operating systems, storage or even a single application functionality need not to be managed by the user. All

customers share a single instance of the hosted application, with the virtualization system managing access. Its typical case is Google's app engine [1], and so on.

2) Cloud Computing Platform as a Service (PaaS), this application development environment created by the tools (such as Java, python, .Net) provided by provider are automatically developed to the cloud computing infrastructure. The underlying cloud infrastructure including networks, servers, operating systems and storage needn't to be managed and controlled by the user. The consumer could control and deploy the application and environment. For example, this type of service could be provided by Windows Azure [2].

3) Cloud Computing Infrastructure as a Service (IaaS) computing power, storage capacity, network rent provided by provider is available to users. Any software including operating system and application configuration could be deployed by users. The underlying infrastructure are not be controlled or managed by users. Amazon [3] is a typical IaaS service provider.

Intrusion detection system has been developed for a long time as a security mechanism for monitoring, and resisting the intrusion. The IDS can be divided into host-based intrusion detection system (HIDS), which is used to monitor the software application of the single host by the means of verifying the operation system and checking the log file, the file system message and the connections of network as in [4]. Network-based intrusion detection system (NIDS) is often used as the non destructive way. The flow of information in the LAN area could be captured by the system and compared with the known attack signatures as in [5]. Distributed intrusion detection system (DIDS), according to the scope of intrusion, is a kind IDS designed to discover attacks on single host as well as the network which is used to aggregate data generated by individual intrusion detection systems [6]. The IDS technology could also be divided into signature-based detection and anomalous-based detection. The

former is used to describe the known attack and intrusion model and form the corresponding event model. When the audited events match the known attack, then the alert is generated as in [7]. The measurement parameters including the number of audit events, time interval, resources consumption, are often used by the latter method as in [8]. The average of measuring property will be used to compare with the behavior of network and system. Any observation outside the normal range is considered to invasion by NIDS. Owing to different deployment mechanisms, IDS can be divided into software-based IDS, hardware-based IDS and VM-based IDS [9]. Strong isolation, fast recovery, and cross-platform are the strengths of the virtualization technology. So the newly emerged VM-based IDS implementations are usually more strong, practical and convenient. In the proposed virtual cloud infrastructure, due to the highly heterogeneous architecture, the VM-based IDS new structure is the core of the paper.

Owing to the combination of the means of cloud service and the different deployment of cloud computing, it brings about the security problem generated by each layer during executing the system. In addition, new security challenges emerge such as how to resolve the deployment of the virtual infrastructure in cloud platform when virtual technology provides the flexible deployment of resource for cloud computing platform. In order to ensure the above issues are more reasonably controlled, it's necessary to deploy IDS sensors to monitor the separated VM at each layer which is controlled by the VM management unit. In order to integrate and analyze the alerts generated by multiple distributed sensors deployed in cloud, a plug-in-concept is proposed in core management unit including collector component, database, threshold compare. They are mainly used to integrate, unify and analyze a large number of security-related events generated by sensors. Furthermore, in the paper, the single point attack of failure must be considered, that is to say, to realize cloud alliance concept by the communication agents. The alert information must be mutual exchanged by communication agents reducing the impact of Denial-of-Service (DoS) and Distributed Denial-of-Service (DDoS). While in order to ensure the communicate agents' robustness so that they could not be disguised by intruders, we introduce a third party certificate module here. All above are proposed in proof-of-concept to realize the architecture.

The rest paper is arranged as follows. II describes the cloud computing involved related work. III introduces the proposed infrastructure. IV simulates about the infrastructure. V lists the future work. VI gives a brief summary.

II. THE RELATED WORK

The simple cloud model and its associated threats, virtualization technology, DoS and DDoS attack are mainly introduced here.

A. The Cloud Model, Technologies-related and Threats Vulnerably Suffered at Each Layer.

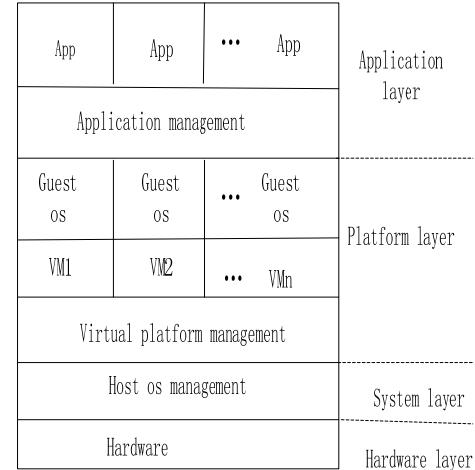


Figure 1. Architecture of cloud computing

As shown in Fig. 1, the cloud architecture can be divided into four layers: the hardware layer, the system layer, the platform layer and the application layer.

1) Hardware layer: include a variety of storage facilities, network facilities, computing facilities, software, host and server. Cloud computing abstracts the infrastructure by using virtualization technology, and formats the corresponding resource pool for the user to call and all this is transparent to users.

2) System layer: the second layer in the cloud architecture includes virtualized hosts and networks. One example for this layer is the Amazon EC2 service [2], which provides virtual hosts and network to the customers.

3) Platform layer: the third layer of the cloud architecture includes virtualized operating systems as well as runtimes and Application Program Interfaces (APIs). A famous example is the Platform Windows Azure [3], which provides the users with several APIs to storage and management.

4) Application layer: the top layer of the architecture provides virtual applications. Google App Engine [1] is a known infrastructure on this layer.

So as [10] described, the specific service which could be directly used by user deployed in each layer is provided by provider with the computing technique described specifically as follows: Grid Computing which is the virtualized combination of computing power from multiple domain getting high computing resource; Utility Computing that consumers pay for computing resources as much as they use without buying them, Server based Computing that any applications and data exist in server. Clients access the server and utilize them using server's computing power and so on.

The hardware layer which is not directly provided to users is not considered here. The system layer threats suffered easily are affected by the traditional

security methods. The user can simultaneously run on several virtual hosts which provide Linux-based and Windows-based web page and ftp file-sharing [11]. Take advantage of this opportunity, the activities of Internet users and pages viewed may be spied through cookies by some commercial companies. In addition, other computers could be controlled by using corpse programs to master what service could be used in cloud. The threats are suffered easily by platform layer especially the hacker attacks. It is also possible that the reliability of the system itself may not fully be guaranteed. The example of application layer suffering the attacks is as follows. Google Gmail service interrupted for up to four hours in February 2009 [12]. The failure is due to the data center in Europe makes another data center overload and spreads the data to other data centers when routine maintenance.

Visibly each layer of the cloud structure may be subjected to some different degrees of damage and attacks.

Another important secure threat of cloud computing is the concept of multi tenancy in VMs. Multi tenancy can be viewed as a hierarchical model, where appropriate policies are enforced on the VMs at every level leading to better governance and segmentation of the consumers. Enforcing different policies at different levels of hierarchy also leads to a secure environment for the consumers to store and access their confidential data. When the VMs are deployed on the physical server security threats always play a major role. Even during the everyday routine utilization there is always a way for the attackers to consolidate their VMs and gain control over the OS. VMs could be moved over from one host to another and has a major threat of being collapsed. While the VM is copied over the network, the state of the VM could be On, Off or suspended. In this research, different VMs need to be assigned to different users because of the possible security threats in a server virtualized environment.

In a word, the following aspects cloud-related securities need to be considered thoroughly such as credibility, reliability, confidentiality and privacy.

B. Virtualization

Virtualization technology has developed rapidly because of the rapid decrease in hardware cost and concurrent increase in hardware computing power. Several features of the virtualization are used in deploying multiple sensors in cloud environment process. In order to avoid a promised IDS sensor to be used to attack other sensors, the virtualization of IDS sensor is described in [13]. A variety of snort-based sensors in the format of plug could be used to effectively monitor virtualization components. The cloud provider wants to identify the components running in a virtual host as directly as running in their underlying architecture. It's necessary that the virtualization technology including VM state, VM work-platform and the monitoring of the configuration information of IDS in VM need to be joined together to integrate the cloud infrastructure with the virtual context e.g. [14] virtualized the system which

makes the system hard to be promised, as well as with the self-reflection of memory .

C. DoS and DDoS Attack

In order to make the computer and network deny to serve the normal service by consuming bandwidth and hosting system resources, mining program defects and providing false DNS information. For example, network communication is blocked and access to service is denied, server crashes or service have been damaged. The denial service capability of DDoS is increased through depending on client and server technology with multiple computers together as an attack platform to launch DoS attack to one or more targets and generate more attacks traffic than DoS [15].

Two primary aspects about a typical DDoS attack are as follows:

1) Initial mass-intrusion: look for the puppets. That is to say, make the internet systems that easy to damage compromised, and then install attack tools in these vulnerable systems.

2) Denial of service attack: the target system will be paralyzed, once the puppets receive an attacking command issued by the attacker through a secure channel.

Many research papers have give taxonomy on DoS and DDoS [16-20]. All of them have analyzed DoS and DDoS attack and defense in different perspectives.

III. PROPOSED ARCHITECTURE

A. VM Management Unit

The security of IDS sensors which are endowed with specific VM component of different layer must be ensured by provider [21]. Network-based or host-based sensors of different layers in virtual environment are managed by VM management unit which is a part of the core management unit. The state of VM such as start, shutdown, stop, continue, reset or update and weather the VM is running, how its platform are involved in virtual environment. The attacks related to the virtual component could be recognized by the provider with the VM management unit. The attack also could be interrupted immediately by automatic counter-measures. Multiple IDS sensors are deployed in specific layer as shown in Fig. 2 by which each virtual component security is insured.

B. Collector

Alerts generated from multiple sensors are collected by the component. Then the alert with the format IDMEF [22] has been proposed as a standard to enable interoperability among different IDS approaches. IDMEF defines a unified format for communication between IDS sensors, response systems, and the core management unit. IDMEF-Message is the top-level class for all IDMEF-Messages. Every type of message belongs to the subclass of the top-level. Alerts and HeartBeats are mainly two types of message. Within each message, the detailed information is provided by the subclass of the message class. It specifies an Extensible Markup

Language (XML)-based data model to represent the exchanging data between IDSs [23]. The Extensible Markup Language (XML) is a simplified version of the Standard Generalized Markup Language (SGML), the syntax for specifying text markup defined by the ISO 8879 standard. XML is gaining widespread attention as a language for representing and exchanging documents and data on the Internet, and as the solution to most of the problems inherent in HyperText Markup Language (HTML). XML is met language-a language for describing other languages – that enables an application to define its own markup. XML allows the definition of customized markup languages for different types of documents and different applications. A message with the type of IDMEF message is the part of IDMEF library which is based IDMEF XML scheme by RFC [23]. However due to the highly heterogeneous architecture, especially VM-based new structure IDMEF is inevitable.

C. Analyze & Compare and Threshold Computing

The alerts unified by the front component are passed to database to compare with the signatures. If the type of packet is correspondence with the one listed in the database, then the alert is considered to be anomaly and to be dropped. Otherwise the alert is continued passed to

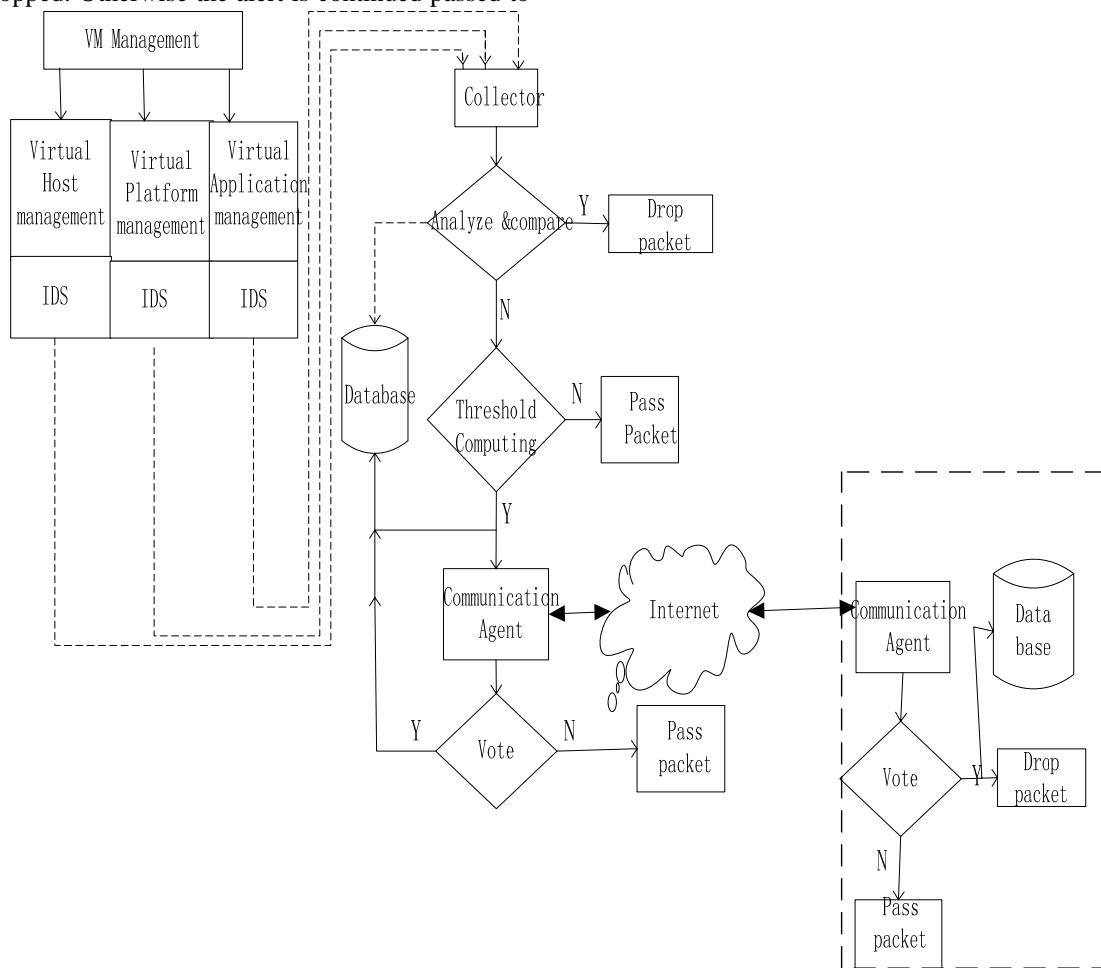


Figure 2. The core management unit

threshold computing. Thus, the time comparing is saved and the detection efficiency is promoted.

The threshold computing formula is described as follows:

$$\text{Threshold} = W + U * V. \quad (1)$$

W means the average of the same type alerts received during a specific time interval. U means the standard deviation. V is a constant dynamically determined by administrator. If the result computed is larger than the specific threshold, then the packet is considered to be anomaly. Last the alert forwards two directions: firstly, the rule of the alert is added into the database shown in Figure 2. Second in order to avoid the single point attack of failure, the alert is continued to be transmitted to communication agent communicating with other cloud region's communication agent. Reducing the communication overheads among IDSs and improving the accuracy of the alerts are the main priorities of the component. The reliability of the system will be improved owing to that false alerts and communication overheads caused by message exchanging are reduced remarkably.

D. Certification Module

Before the alerts are exchanged among communication agents of different cloud regions through the internet, we introduce an improved Kerberos protocol certification [24] module as the third party in figure.3 to ensure the communication agents' true identities.

First of all, the key components of the module are described specifically. Followed by is the introduction of the entire workflow.

Certificate Authority (CA) which is the third party trusted organization provides information security service for the network agents and equipments. Authentication Server (AS) could confirm the agent's identity when the agent logs by sharing a key with each user. Ticket Granting Service (TGS) distributes the ticket for the communications among agents. So that the application server believes that the holder of TGS is as same as it claims.

The specific workflow of the module is as follows (the detailed description of (1), (2)..... in Fig. 3 is accordance with the followed 1), 2)):

1) The agent requests the key certificate of TGS for CA by sending its public key, agent's server name and the server name of TGS.

2) After checking the agent's legitimate identity, CA issues the public key certificate attached with the key certificate of the information server of agent's TGS and issuance time for agent and TGS, through the public key.

3) Each time when agent needs to use the service in application server, generate an authentication code containing the user name, time stamp and additional session key from the AS.

4) After receiving the request, TGS decrypts the license ticket with its own private key for attaining session key, decrypted Authentication Code (AC) and verifying the validity of the timestamp. While, TGS confirms the legitimacy of the agent, the license ticket containing the session key to agent whose responsibility is decrypting the new public key certificate with own private key to get the public key and session key of TGS.

5) Agent encrypts the AC by use of the session key, then re-encrypts with private key. At last, after agent using the server's public key to encrypt its own public key and ticket, it is sent to TGS server. So here, we achieve a two-factor authentication, that is to say, Kerberos's session key certification and CA certificate authentication.

6) After the TGS server decrypts the got information with private key to obtain the agent's public key, other information is also decrypted to verify the agent's identity. What's more, the TGS server return a ticket and time stamp encrypted with private key to agents. So that after the more secure recognition based on the two-factor authentication, the two sides can communicate through the secure channel.

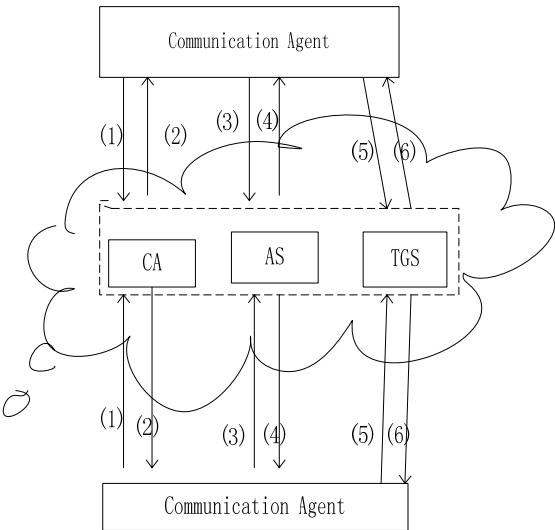


Figure 3. Certificate module

E. Communication Agent

In order to avoid the single point attack of failure, alert coming from the threshold component is passed to another cloud region's communication agent by the Internet. Then the received alert is further judged the reliability of alerts as the following formula:

$$\frac{\text{#number of IDSs sends the same alert}}{\text{#number of IDSs in the cloud}} > 0.5. \quad (2)$$

If the result is larger than 0.5, the packet is considered to be anomaly. Meanwhile its rule set is sent to database and the alert is passed to other cloud region communication agent.

So through the component we know that if the packet is passed the component frequently, the attack will be found and then this type of packet will be dropped. And the single point attack of failure will not be occurred.

IV. SIMULATION

To test the feasibility of the above mentioned architecture, we simulate the experiment. It consists of two servers I and II to simulate the two different cloud regions, at the same time, VM-based intrusion detection system is set up. Server-I with the IP address 202.113.73.150, executes two VMs including F-Secure sensor [25] and snort sensor and server-II with the IP address 202.113.73.183, executes three VMs including two Samhain sensors [26] and one snort sensor. The five different IDS sensors are separately used to supervise the local malicious or filter the network traffic. The basic configuration of each server is 2GB memory, 500GB hard disk and 2GHZ CPU and on-chip TPM v1.2. The third server as an attacker firstly launches attacks like TCP/IP packets, SYN flooding [27] to the servers in the format Network Mapper (NMAP) to scan the ports. We also compare with the NIDS system deployed in the cloud. Our measuring framework will incur additional, un-avoidable computing overhead. In cloud computing,

many applications run concurrently. Generally, the measuring framework would affect parallel programs more than sequential programs, so we tested parallel programs. The experiment results are as follows.

Sensor	Classification	Time	Source Address	Target Address
fsecure_sensor	malicious code found in file/root/test....	2011-06-20	null	null
snort_sensor	(portscan)TCP Partscan community sip	2011-06-20	202.113.76.183	202.113.73.150
snort_sensor	TCP/IP message flooding directed to ...	2011-06-20	202.113.73.150	202.113.76.183
fsecure_sensor	malicious code found in file/root/test....	2011-06-20	null	null

Figure 4. The simulation result

As shown in Fig. 4, the working state of the whole architecture can be viewed by user. The end user can also inspect each IDMEF alert directly for the detailed information of the alert. Further, the IDS and VM on this platform can also be detected by the end user.

TABLE I
THE DETECTION CONDITION OF SYN FLOODING

Simulation systems	Detection rate	False-positive rate	Negative rate
NIDS system	82%	0.8%	1.12%
Propose system	94.27%	0.55%	0.52%

The architecture's simulation results prove that it has a high detection rate, little false-positive rate and negative rate. It resists the failure of single point attack effectively.

V. FUTURE WORK

The realization of this proof-of-concept is only just beginning in the domain of the VM-based IDS system in the complicated cloud computing environment. First of all, the virtualization for alerts can be further regulated to support manual analysis. What is more, the next step needs to be researched is that the collect component needs to be upgraded to enable different approaches for analyze & compare component, such as labeling, filtering, classification. The other interesting research point is the possibility of implementing counter measures by use of VM management unit. It is possible that new correlation methods will also be needed, by taking the special allocation and related components of the cloud computing framework into consideration. Moreover, in order to truly realize the allocation, enforceability and scalability problems need further to be considered. So far, this frame-work has only been implemented in several physical machines. Because clouds involve multiple computers (servers) cooperating with each other, further research should focus on a measuring framework applicable to multiple physical machines followed by a number of unexpected problems.

VI. CONCLUSION

Considering the complexity of the cloud security architecture, a extensible VM-based multiple IDSs are deployed in each layer to monitor specific virtual component and the core management unit are constructed by multiple plugs satisfying the IDMEF standard to realize the ideas of virtualization and cloud alliance which is mainly used for avoiding the single point attack of failure. That is to say, the large scale attacks to several users may be detected easily by related and mutual alerts passed by communication agents. To further ensure the alerts could be sent to communication agent of another cloud region, the certificate module plays an important role here.

ACKNOWLEDGMENT

This work was supported by the Foundation of Tianjin for Science and Technology Innovation (No.10FDZDGX004 00&11ZCKFGX00900), Education Science and Technology Foundation of Tianjin (No.SB20080053 & SB20080055).

REFERENCE

- [1] Google Apps, <http://www.google.com/a>.
- [2] WindowsAzurePlatform,<http://www.microsoft.com/azure/default.mspx>.
- [3] Amazon Elastic Compute Cloud and Simple Storage Service, <http://aws.amazon.com>.
- [4] M.M. Yasin, A.A. Awan, "A study of host-based IDS using system calls," *Lahore, Pakistan*, vol. 3, pp. 36-41, June 2004
- [5] A.K. Ganame, J. Bourgeois, R. Bidou, F. Spies, "A global security architecture for intrusion detection on computer networks," *Montbeliard*, vol. 27, pp.30-47, March 2008,
- [6] Xiang Yang; Ke Li; Wanlei Zhou, " Low-rate DDoS attacks detection and traceback by using new information metrics," vol. 6 (2), pp.426-437,Jun 2011.
- [7] K. Sunil, L. Jun-yong, "A system architecture for high-speed deep packet inspection in signature-based network intrusion prevention," vol. 53, pp. 310-320, May-June 2007,
- [8] Q. Yan, W. Xie, B. Yan, et al. "An anomaly intrusion detection method based on HMM," vol. 38 (13), pp. 663-664, May 2006.
- [9] Xiantao. Zhang, Qi. Li; Sihan. Qing, "Building an IDS architecture using VMM-based non-intrusive approach," *Adelaide, SA, Australia* , pp. 594-600, Jan. 2008.
- [10] <http://radlab.cd.berkley.edu/UC> Berkley Reliable Adaptive Distributed Systems Laboratory (accessed on Feb. 2009).
- [11] P. Willmann, J. Shafer, D. Carr, A. Menon, S. Rixner, A. L. Cox, et al, "Concurrent direct network access for virtual machine monitors," *Lausann. Switzerland*, vol. 10-14, pp. 306-317, February 2007
- [12] <http://code.google.com/appengine/>,Google (accessed on Oct 2009).
- [13] G.W Dunlap, S.T King, S.Cinar, M.Basrai, " Enabling intrusion analysis through virtual-machine logging and replay," *Boston, MA, USA*, pp. 211-240, December 2002.
- [14] F. Cheng, S. Roschke, Ch. Meinel, "Implementing IDS management Lock-Keeper (ISPEC'09)," *Xi'an, China*, pp. 360-371, April 2009

- [15] S. Ranjan, R. Swaminathan, M. Uysal, et al., "DDoS-shield: DDoS-resilient scheduling to counter application layer attacks," vol. 17 (1), pp. 26-39, Feb. 2009,
- [16] M. Jelena, R. Peter, "A taxonomy of DDoS attack and DDoS defense mechanisms," vol. 34(2), pp. 39-53, 2004
- [17] A. Hussain, J. Heidemann, C. Papadopoulos, "A framework for classifying denial of service attacks," vol.33 (4), pp. 99-110, 2003
- [18] Alex Wun, Alex Cheung, Hans-Arno Jacobsen, "A taxonomy for denial of service attacks in content-based publish/subscribe systems," *Toronto, Ontario, Canada*, pp. 116-127, 2007
- [19] T. Usman, H. ManPyo, and L. Kyung-suk, "A comprehensive categorization of DDoS attack and DDoS defense techniques," *Berlin*, vol.4093, pp. 1025-1036, 2006
- [20] G. Michael, "A summary of DoS/DDoS prevention, monitoring and mitigation techniques in a service provider environment," 2003.
- [21] Qian. Liu , Chuliang. Weng , Minglu. Li , Yuan. Luo, "An In-VM measuring framework for increasing virtual machine security in clouds," vol. 8 , pp. 56 - 62 , Nov.-Dec, 2010
- [22] H. Debar, D. Curry, B. Feinstein, "The intrusion detection message exchange format," *Korea*, pp. 193-199, July 2004.
- [23] M. Kudo, S. Hada. "XML document security based on provisional authorization," vol. 33(4), pp. 379-389, 2004.
- [24] J.T. Kohl. "The evolution of the Kerberos authentication service," *Tromso, Norway*, pp. 295-313, 2009.
- [25] M. Komssi, M. Kauppinen, J. Heiskari, M. Ropponen, "Transforming a software product company into a service business: Case study at F-Secure issue," vol. 1, pp. 61 - 66 , July 2009
- [26] <http://www.f-secure.com/linux-weblog/> F-Secure Corporation (accessed Oct 2009).
- [27] D. Moore, C. Shannon, D.J. Brown, and et al. "Inferring internet denial-of-service activity," *California, San Diego*, vol. 24(2), pp.115-139, 2006.

Scheduling Program of the Track and Field Sports Competition Based on R_Timetable Algorithm

Yongxin Wang

Nanyang Institute of Technology/Department of Physical Education, Nanyang, China

Email: wyx_310310@163.com

Jianying Wang and Qiufen Wang

Nanyang Institute of Technology/ Department of Applied Mathematics, Nanyang, China

Nanyang Institute of Technology/ Department of Computer Science and Technology, Nanyang, China

Email: boocoo@163.com, w_qiufen @163.com

Abstract—The characteristics and several common open questions of the scheduling program in the track and field sports competition are analyzed in this paper, and then an automation method based on R-timetable algorithm, which is achieved effectively through the C++ program design, is provided to solve the corresponding problems. It has been found that the method in this paper immensely optimizes the organization, and perfectly improves the efficiency of arrangement of the sports competition.

Index Terms—R-timetable, Track and Field Sports Competition, Scheduling program, Algorithm.

I. INTRODUCTION

Competition schedule is very important. It determines whether the whole competition can carry out smoothly, since it shows all the arrangements for all of the athletes to compete in a contest, also for the judges to work and for the audiences to watch. Thus arranging the agenda is definitely complicated and difficult. The reasons lists showing as following.

Firstly arranging the competition schedule is a sort of problems which belongs to the time planning program or belongs to kind of NP problems, and the deterministic algorithm for solving this kind of problems has not yet been achieved till now.

Secondly the principally key part of arranging the competition schedule is to ascertain the best permutation, and it's generally a huge number of permutation for all of the items. Taking 10 items for example, the number of

Manuscript received August 20, 2011; accepted September 2, 2011.

Yongxin Wang(1978-), male, Master, Lecturer, Research direction: Sports teaching & training, Computer aided instruction & management.

Jianying Wang(1981-), Female, Bachelor, Lecturer, Research direction: Applied Mathematics, ordinary differential equation.

Qiufen Wang(1978-), Female, Master, Lecturer, Research direction: Computer software and algorithm theory.

permutation is 3,628,800.

Lastly due to the influence of a number of factors, which are related to each other and meanwhile restricted or impacted mutually, for instance, javelin item and shot item cannot be arranged at the same period, 3,000-meter item and 5,000-meter item cannot be conducted on the same day, etc. We are still looking forward to finding out one of the best way to arrange the competition schedule.

At present, there are still a series of typical problems against the arrangement of sports competition schedule, mainly showing as follow.

Firstly the schedule of the school sports competition is usually worked out just by simple hand working, so it's common to make mistakes with inefficiency.

Secondly the arrangement of competition schedule mostly bases on the past experiences, instead of carrying out according to the logical scientific theory.

Lastly there are only a few methods up to now, such as cut-and-try method, genetic algorithm[1], search method[2] etc, which are studied and applied to arrange the sports competition schedule. Although all these methods can speed the arrangements effectively, it's hard to ensure the approving results as well.

In view of the above-mentioned facts, an automation method based on R-timetable algorithm is given to optimize the organization and improve the efficiency of arrangement of the sports competition in this paper.

II. FOCUSED BACKGROUND

There are a series of particular characteristics of the track and field sports competition holding in the colleges or universities compared with other significant official sports competitions[3-7], showing as following.

Firstly, at present for many contemporary colleges and universities, there are usually at least two playgrounds can be used for athletic fields, thus track competition and field competition can be arranged at the same periods but

the different sports grounds. Consequently arrangements for sports competition can be carried out easily just by considering the conflicts of concurrently items.

Secondly, there are just three days by a definite date of the school sports competition, and the concurrently items are seldom regular ways since all of the athletes are students or teachers but not the professional athletes, thus they join the different items just according their hobbies.

Thirdly, school sports competitions commonly include many different kinds of groups, such as man and women items groups of students, man and women items groups of young teachers, men and women items groups of middle age teachers, men and women items groups of senior age teachers.

Lastly, it's flexible for arranging the agenda of school sports competition compared with the professional sports competitions. According to the actually special situation, there are some recreational items for senior age teachers arranged during the common items, such as the games of upland fishing, carrying things run.

Therefore, the principle of arrangement for the school sports competition is trying to avoid the clashes of the

concurrently items since it's really complicated because of much more different groups and special items. By considering the R-timetable algorithm method is one of the most effective way to solve time planning problems, the agenda arrangement is carried out by the R-timetable automation algorithm in this paper.

III. R_TIMETABLE ALGORITHM

A. Temporal Matrix of Relation

For a random event, let the starting time be s_i , and the finishing time be f_i ($s_i < f_i$), define the time interval of the event as $[s_i, f_i]$, denoted by I_i . Then the relations between any two random events can be shown by corresponding time intervals of each other.

For the random event I_1 and another random event I_2 , the two time intervals of these two random events should be denoted as $I_1(s_1, f_1)$ and $I_2(s_2, f_2)$ respectively, then there are 13 different time intervals for all of possible relations between these two events I_1 and I_2 , showing as the following table 1.

TABLE 1:
THE 13 DIFFERENT TIME INTERVALS FOR ALL POSSIBILITIES BETWEEN THE TWO EVENTS I_1 AND I_2

The relation between I_1 and I_2	number	Symbolic representation	The relation between s_i and f_i (for $i=1,2$)	sketch map of the time intervals (I_2 remains unchanged)
before	1	<	$s_1 < f_1 < s_2 < f_2$	
meet	2	m	$s_1 < f_1 = s_2 < f_2$	
overlap	3	o	$s_1 < s_2 < f_1 < f_2$	
finished by	4	fi	$s_1 < s_2 < f_1 = f_2$	
contains	5	di	$s_1 < s_2 < f_2 < f_1$	
started by	6	si	$s_1 = s_2 < f_2 < f_1$	
start	7	s	$s_1 = s_2 < f_1 < f_2$	
equal	8	$=$	$s_1 = s_2 < f_1 = f_2$	
during	9	d	$s_2 < s_1 < f_1 < f_2$	
finish	10	f	$s_2 < s_1 < f_1 = f_2$	
overlaped by	11	oi	$s_2 < s_1 < f_2 < f_1$	
meet by	12	mi	$s_2 < f_2 = s_1 < f_1$	
after	13	$>$	$s_2 < f_2 < s_1 < f_1$	

In the rectangular coordinate system, the ordered pair, such as (s_i, f_i) , can be denoted by a point. Then the time intervals of the events including all possibilities can be represented by coordinate points in the plane. There are 5 cases among s_1 and I_2 , which are $s_1 < s_2$, $s_1 = s_2$, $s_2 < s_1 < f_2$,

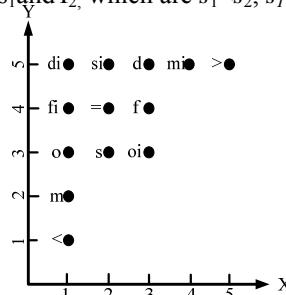
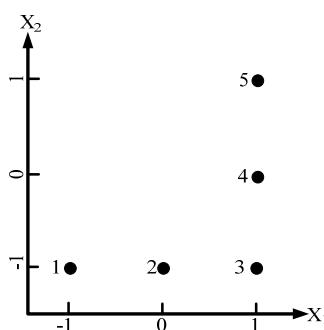


Figure 1. Figural relations between I_1 and I_2

$s_1 = f_2$, $s_1 > f_2$. Similarly, there are 5 cases among f_1 and I_2 , which are $f_1 < s_2$, $f_1 = s_2$, $s_2 < f_1 < f_2$, $f_1 = f_2$, $f_1 > f_2$. Number 1, 2, 3, 4, 5 represent the corresponding intervals or points $(-\infty, s_2)$, $[s_2]$, (s_2, f_2) , $[f_2]$, $(f_2, +\infty)$ respectively, thus the figural relations between I_1 and I_2 show as figure 1.

Moving forward to follow the same method, assume that $X_1 = \{-1, 0, 1\}$ and number -1, 0, 1 represent intervals or points of $(-\infty, s_2)$, $[s_2]$, $(s_2, +\infty)$ respectively.

Similarly, let $X_2 = \{-1, 0, 1\}$ and number -1, 0, 1 represent the intervals or points of $(-\infty, f_2)$, $[f_2]$, $(f_2, +\infty)$ respectively. Using number 1, 2, 3, 4, 5 to represent the subsets of the product space among X_1 and X_2 , then X can be expressed as the subset of the space domain $X_1 \times X_2$, showing as the following figure 2.

Figure 2. Subset of the product space $X_1 \times X_2$

In the same way, let $Y_1 = \{-1, 0, 1\}$ and the number -1, 0, 1 represent the intervals or points of $(-\infty, s_1]$, $[s_1], (s_1, +\infty)$ respectively; Similarly, let $Y_2 = \{-1, 0, 1\}$ and the number -1, 0, 1 represent the intervals or points $(-\infty, f_2]$, $[f_2], (f_2, +\infty)$ respectively, then Y can be expressed as the subset of the space domain $Y_1 \times Y_2$.

Supposing that if $R(1, 2)$ is the simple-ingredient time relational constraints [8] between I_1 and I_2 , then $R(1, 2)$ can also be denoted by using a 2×2 matrix. For example, if $R(1,2)=\{<\}=(1)\times(1)=((-1)\times(-1))\times((-1)\times(-1))$, then it can be expressed as a matrix

$$\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}.$$

In general, supposing that if the set of events is $\{(i, I_i), i=1, 2, 3, \dots, n\}$ and the time interval I_i is (s_i, f_i) for any random event $i(i=1, 2, 3, \dots, n)$, then for any two random events I_i and I_j , the time interval relational constraints between them should be

$R(i,j)=(A_1(i,j) \times A_2(i,j)) \times (B_1(i,j) \times B_2(i,j))$, since A_1, A_2, B_1, B_2 are the subsets of the set $\{-1, 0, 1\}$, the corresponding matrix should be

$$M(i,j)=\begin{pmatrix} A_1(i,j) & A_2(i,j) \\ B_1(i,j) & B_2(i,j) \end{pmatrix}, \quad i \neq j,$$

especially when $i=j$ we get

$$M(i, i)=\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The matrix

$$M=\begin{pmatrix} M(1,1) & M(1,2) & \cdots & M(1,n) \\ M(2,1) & M(2,2) & \cdots & M(2,n) \\ \vdots & \vdots & \vdots & \vdots \\ M(n,1) & M(n,2) & \cdots & M(n,n) \end{pmatrix}$$

is called the corresponding relational matrix of $\{R(i, j)\}$, or time matrix of relation for short, and $M(j, i)=-M(i, j)^T$. It contains n^2 2×2 matrices, in other words, it's a matrix of $2n \times 2n$ and any element can be written in the form of $M(i, j) \quad i, j=1, 2, \dots, 2n$.

B. Description of the Algorithm

a. Data structure

The main data structure in the R-timetable algorithm

includes structure types, set types and array types. All of the specific details show as following.

The 1-dimension structure types data for saving the time intervals $p[i]$ and x_1 and x_2 , the relations between them showing in the figure 1 and figure 2, and the data of storage listing in the following table 2 and table 3.

TABLE 2:
THE ELEMENTS OF ARRAY P

i	1	2	3	4	5	6	7	8	9	10	11	12	13
p[i].x	1	1	1	1	1	2	2	2	3	3	3	4	5
p[i].y	1	2	3	4	5	5	3	4	3	4	5	5	5

TABLE 3:
THE ELEMENTS OF ARRAY X_1 AND X_2

i	1	2	3	4	5
x1x2[i].x	-1	0	1	1	1
x1x2[i].y	-1	-1	-1	0	1

R is the 2-dimension integral types array for storing the data of time intervals. For the specific event sets E , pick up any random element event i and another element event $j(j \in E - \{i\})$, the relation of the event i and j are stored in the 2-dimension integral types array $R(i, j)$, which should be denoted as

$$R(i, j) \subset \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}.$$

Substitute the corresponding value for example, it can be $R(1, 2)=\{1, 2\}$, $R(1, 3)=\{1\}$, $R(2, 3)=\{13\}$.

M is the 2-dimension integral array for storing the data of time matrix of relation.

HJ is the set of the corresponding line numbers related to the temporal matrix of relation M .

e is the line number of the corresponding non-positive row vector.

S keeps the records of consistent subsets of e in the 1-dimension integral array.

b. Design of the algorithm

Assume that there are n random events, then the steps of the algorithm show as following.

Step 1, Input temporal matrix of relation M and the set of the corresponding line numbers $HJ(M)$.

Step 2, Let $S=\Phi$.

Step 3, If $HJ(M)=\Phi$, then output S , since S is the permutation number of the events, which can meet the conditions of time relational constraints, the program ends. If not, move on to the step 4.

Step 4, If M has non-positive row vectors, then e_i (starting from $i=1$) represents the set of the line numbers for these row vectors, and move on to step 5. If not, the program ends.

Step 5, Work out the compatible subset of e_i , denoted by d_i , move on to step 6. If not, the program ends.

Step 5-1, Let $e_i=\{x_1, x_2, \dots, x_m\}, i=1$;

Step 5-2, If $j \geq m+1$, the program ends. Otherwise, let

$L_0=\{x_j\}, L_1=\{y/m(x_j, y)=\{0, 1\} \text{ or } =\{0\}, y \in HJ(M)\}$, Generally,

$L_{k+1} = \{y | \exists x \in L_k, m(x,y) = \{0,1\} \text{ or } =\{0\}, y \in HJ(M)\}$.
Step 5-3,
 $L(x_j) = \bigcup_k L_k$. For $\forall z \in HJ(M)$,
 $\bigcap_{y \in L(x_j)} m(y,z) \neq \emptyset \text{ and } \neq \{1\}$.

Then let $d_i = L(x_j)$, and d_i is the compatible subset of e_i including x_j , if it's done, move on to step 6. If not, move back to step 5-1 for $j++$.

Step 6, Delete all of the row elements and the column elements related to d_i from the matrix M, then we'll get the corresponding sub matrix M_i .

Step 7, Adding all of the elements of d_i to S , giving M the value again with M_i for $i++$, move back to step 3.

Note: Non-positive row vectors, which are denoted by $(a_{i1}, a_{i2}, \dots, a_{i2n})$, for
 $a_{ij} = \min\{t/t \in m(i,j)\} (j=1,2,\dots,2n)$, and $a_{ij} \leq 0$.

IV. CASE STUDY OF ARRANGING THE COMPETITION SCHEDULE

A. Set Up Temporal Matrix of Relation

a. Elementary principles for arranging the schedule

1) Short distance sport items first, and then the long distance ones.

2) Preliminary contest first, and then finals.

3) Track first, and then the field events.

4) Spacing interval during the same sport events should be reasonable to make sure athletes have enough time to rest or get ready for next item[9-10].

5) Try to arrange attractive or interesting items and the others alternately.

6) Minimize the clashes in a possible way according to the situation of the multiple occupied athletes, and try to prolong the time span among any items if simultaneous entries happen.

One of the effective way to minimize the clashes of the multiple occupations is to interlude items and hold the concurrently item in the different time spans, which is the precedence essentially related to each of the time interval of sport events.

According to the statistical data of the entry form list, for the items of Men's 100-meter and Men's 400-meter, Women's 100-meter and Women's 200-meter, athletes usually take part in the items at the same time. In order to leave enough time for the multiple occupied athletes to rest, the schedule can be arranged as following way

Men's 100-meter → Women's 100-meter → Men's 400-meter → Women's 200-meter.

The contradistinctive nexus among the four items can be shown as the following table 4.

TABLE 4.

THE TIME INTERVALS OF CONCURRENTLY ITEMS

	Men's 100-meter	Men's 400-meter	Women's 100-meter	Women's 200-meter
Men's 100-meter	=	<	<	<
Men's			=	>

400-meter	=	<
Women's		=
100-meter		
Women's		
200-meter		

From the above table, we know that Men's 100-meter < Men's 400-meter, which means that the item of Men's 100-meter is arranged prior to Men's 400-meter. In the same way, Men's 400-meter > Women's 100-meter means that the item of Men's 400-meter is arranged behind the item of Women's 100-meter. Consequently, the clash of concurrent items is removed in a perfect way by putting the item Women's 100-meter in the item Men's 100-meter and Men's 400-meter.

b. Case study of the competition schedule

Taking the agenda of the 17th track and field sports competition of Nanyang Institute of Technology for the example, we worked out one of the best scheme of the competition schedule. Preliminary contests and finals of the typical items in the track items and field items were principally researched in what follows.

Typical items mentioned above include Men's 100 meter , Men's 400-meter , Men's 3000-meter , Women's 100-meter , Women's 400-meter , Women's 3000-meter , Men's shot-put , Women's shot-put , Men's long jump, Women's long jump. Now regard all of these items as individual events and put them in order, the opening and closing ceremony both are ordered in the serial number. There are 18 events in total, showing as follow table 5.

TABLE 5.
SPORTS EVENT AND SERIAL NUMBER

Serial number	Sports event
1	Men's 100-meter preliminary
2	Women's 100-meter preliminary
3	Men's 400-meter preliminary
4	Women's 400-meter preliminary
5	Men's 3000-meter final
6	Women's 3000-meter final
7	Men's 100-meter final
8	Women's 100-meter final
9	Men's 400-meter final
10	Women's 400-meter final
11	Men's shot-put qualifying round
12	Women's shot-put qualifying round
13	Men's long jump qualifying round
14	Women's long jump qualifying round
15	Men's long jump final
16	Women's long jump final
17	Opening ceremony

According to the principle of the arrangement for the school competition schedule, all of the time intervals co-relations of eighteen events mentioned above show as following table 6.

TABLE 6.
CO-RELATIONS OF 18 EVENTS' TIME INTERVAL

Obtain the temporal matrix of relation according to the time interval co-relations of 18 events as following figure 3.

Figure 3 : Temporal matrix of relation of case study

B. Consequence and Analysis of the Case Study Arrangement

a. Consequence of case study

According to the temporal matrix of relation, run the $d_3=\{1,21\}, e_4=\{2\}, d_4=\{2\}, e_5=\{3\}, d_5=\{3\}, e_6=\{4\}, d_6=\{4\}, e_7=\{22\}, d_7=\{22\}, e_8=\{5,25\}, d_8=\{5,25\}, e_9=\{6\}, d_9=\{6\}, e_{10}=\{7\}, d_{10}=\{7\}, e_{11}=\{8\}, d_{11}=\{8\}, e_{12}=\{26\}, d_{12}=\{26\}, e_{13}=\{13,23\}, d_{13}=\{13,23\}, e_{14}=\{14\}, d_{14}=\{14\}, e_{15}=\{15\}, d_{15}=\{15\}, e_{16}=\{16\}$,

$d_{16}=\{16\}, e_{17}=\{24\}, d_{17}=\{24\}, e_{18}=\{9,27\}, d_{18}=\{9,27\}, e_{19}=\{10\}, d_{19}=\{10\}, e_{20}=\{28\}, d_{20}=\{28\}, e_{21}=\{11,29\}, d_{21}=\{11,29\}, e_{22}=\{12\}, d_{22}=\{12\}, e_{23}=\{17\}, d_{23}=\{17\}, e_{24}=\{18\}, d_{24}=\{1$

$\overline{\underline{-I_{17}}}$	$\overline{\underline{-I_1}}$	$\overline{\underline{-I_2}}$	$\overline{\underline{-I_3}}$	$\overline{\underline{-I_4}}$	$\overline{\underline{-I_5}}$	$\overline{\underline{-I_6}}$	$\overline{\underline{-I_7}}$	$\overline{\underline{-I_8}}$	$\overline{\underline{-I_9}}$	$\overline{\underline{-I_{10}}}$	$\overline{\underline{-I_{11}}}$	$\overline{\underline{-I_{12}}}$	$\overline{\underline{-I_{13}}}$	$\overline{\underline{-I_{14}}}$	$\overline{\underline{-I_{15}}}$	$\overline{\underline{-I_{16}}}$	$\overline{\underline{-I_{17}}}$	$\overline{\underline{-I_{18}}}$	$\overline{\underline{-I_{19}}}$	$\overline{\underline{-I_{20}}}$	$\overline{\underline{-I_{21}}}$	$\overline{\underline{-I_{22}}}$	$\overline{\underline{-I_{23}}}$	$\overline{\underline{-I_{24}}}$	$\overline{\underline{-I_{25}}}$	$\overline{\underline{-I_{26}}}$	$\overline{\underline{-I_{27}}}$	$\overline{\underline{-I_{28}}}$	$\overline{\underline{-I_{29}}}$	$\overline{\underline{-I_{30}}}$	$\overline{\underline{-I_{31}}}$	$\overline{\underline{-I_{32}}}$	$\overline{\underline{-I_{33}}}$	$\overline{\underline{-I_{34}}}$
----------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------

Figure 4 : The marshalling sequence of the case study events

b. Analysis of case study resolution

According to the above conclusion data, the whole procedure should be: Opening ceremony first.

For track items, the sequence should be:

Men's 100-meter preliminary→Women's 100-meter preliminary→Men's 400-meter preliminary→Women's 400-meter preliminary item →Men's 100-meter final item→ Women's 100-meter final→Men's 3000-meter final→ Women's 3000-meter final→Men's 400-meter final→Women's 400-meter final;

For field events, the sequence should be:

Men's shot-put qualifying round→Men's long jump qualifying round item →Women's shot-put qualifying round→Women's long jump qualifying round→Men's long jump final→Women's long jump final.

For interlude items, the sequence should be:

Men's shot-put qualifying round item should start while Men's 100-meter preliminary item is being in play, Men's long jump qualifying round item should start while Men's 400-meter preliminary item is being in play, Women's shot-put qualifying round item should start while Men's 100-meter final item is being in play, Men's long jump qualifying round item should start while Men's 3000-meter final item is being in play, Men's long jump final item should start while Women's 3000-meter final item is being in play, Women's long jump final item should start after Women's 3000-meter final item ends.

This agenda is highly logical and rigorously scientific by following almost all of the elementary principles of arrangement and taking account of the other factors, such as, consider to collocate the short distance items and the long distance ones, combine the preliminary contests and the finals, intersect the track and the field events, of course leave the reasonable spacing interval during the same sport events so that athletes have enough time to rest or get ready for next item, and also arrange attractive or interesting items and the others alternately, etc.

V. CONCLUSION

program based on R_time table algorithm, we can obtain the data as following:

$e_1=\{33\}, d_1=\{33\}, e_2=\{34\}, d_2=\{34\}, e_3=\{1,21\},$

$8\}, e_{25}=\{30\}, d_{25}=\{30\}, e_{26}=\{19\}, d_{26}=\{19\}, e_{27}=\{20\}, d_{27}=\{20\}, e_{28}=\{31\}, d_{28}=\{31\}, e_{29}=\{32\}, d_{29}=\{32\}, e_{30}=\{35\}, d_{30}=\{35\}, e_{31}=\{36\}, d_{31}=\{36\}$

Then the corresponding R_time table should be:

$(33,34,(1,21),2,3,4,22,(5,25),6,7,8,26,(13,23),14,15,16,24,(9,27),10,28,(11,29),12,17,18,30,19,20,31,32,35,36)$

and the marshalling sequence of all these events shows as figure 4.

$\overline{\underline{-I_{17}}}$	$\overline{\underline{-I_1}}$	$\overline{\underline{-I_2}}$	$\overline{\underline{-I_3}}$	$\overline{\underline{-I_4}}$	$\overline{\underline{-I_5}}$	$\overline{\underline{-I_6}}$	$\overline{\underline{-I_7}}$	$\overline{\underline{-I_8}}$	$\overline{\underline{-I_9}}$	$\overline{\underline{-I_{10}}}$	$\overline{\underline{-I_{11}}}$	$\overline{\underline{-I_{12}}}$	$\overline{\underline{-I_{13}}}$	$\overline{\underline{-I_{14}}}$	$\overline{\underline{-I_{15}}}$	$\overline{\underline{-I_{16}}}$	$\overline{\underline{-I_{17}}}$	$\overline{\underline{-I_{18}}}$	$\overline{\underline{-I_{19}}}$	$\overline{\underline{-I_{20}}}$	$\overline{\underline{-I_{21}}}$	$\overline{\underline{-I_{22}}}$	$\overline{\underline{-I_{23}}}$	$\overline{\underline{-I_{24}}}$	$\overline{\underline{-I_{25}}}$	$\overline{\underline{-I_{26}}}$	$\overline{\underline{-I_{27}}}$	$\overline{\underline{-I_{28}}}$	$\overline{\underline{-I_{29}}}$	$\overline{\underline{-I_{30}}}$	$\overline{\underline{-I_{31}}}$	$\overline{\underline{-I_{32}}}$	$\overline{\underline{-I_{33}}}$
----------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------

This paper provides an automation method based on R-timetable algorithm, which is achieved effectively through the C++ program design, since it's proved to be logically feasible, according to the characteristics and several common questions of the scheduling program in the track and field sports competition.

The data of experimental results indicate that the method contributes many merits and advantages with respect to many common and typical problems, mainly show as follows:

First, the method optimizes the assembly of programs perfectly, including less clashes of merged programs, reasonable interludes of attractive programs, etc.

Second, it also remarkably increases the efficiency, makes the arrangement easier and more concise than before.

REFERENCE

- [1] Feng Lin and Jin-yuan Yang. "Track and field sports meeting scheduling based on genetic algorithm". *JOURNAL OF JILIN INSTITUTE OF CHEMICAL TECHNOLOGY*, 4thed,vol 27,2010,pp.88-90.
- [2] Dan Li and Li-chen Wang, "An Algorithm Model of Popular Track and Field Sports Competition Schedule" ,*JORUNAL OF PHYSICAL EDUCATION INSTITUTE OF SHANXI TEACHERS UNIVERSITY*, 4thed,vol 17, 2002,pp.55-57.
- [3] Beijing Physical Education DepartmentsTeaching Materials Compilation Committee. *track and field*.Beijing: People's Sports Publishing House,1978.
- [4] Chinese Athletic Association Validation. *Track and Field Competition Rules*. Beijing: People's Sports Publishing House,1995.
- [5] Rong-fang GAO, Li-jun YIN and Jing ZHANG. "Design and Realization of a Dynamic Grouping Algorithm Dynamic Grouping Algorithm Based on Track Events". *COMPUTER ENGINEERING & SOFTWARE*,1sted,vol31,2011,pp.1-3.
- [6] Lin HE. "On the Adoption of a Computerized Managerial System for a Sports Meet at Grass-roots Level". *JOURNAL OF ZHEJIANG SHUREN UNIVERSITY*,3rded,vol 1,2001,pp.61-64.

- [7] Hui HE. "On the Determination of Time Duration of Each Stage of a Track Event to Improve Scheduling of Track and Filed Meet". JOURNAL OF XI'AN INSTITUTE OF PHYSICAL EDUCATION, 2nded, vol 18, 2001, pp. 110-112.
- [8] Bo Zhang and Ling Zhang, "Relational matrix method of temporal planning", CHINESE JOURNAL OF COMPUTERS, 6thed, vol 14, 1991, pp.411-422.
- [9] Xianjun Lan, "Description of scheduling program for the track and field sports meeting", READ AND WRITE PERIODICAL, 11thed, vol 5, 2008, pp.82.
- [10] Laoming Li, "A manual for arbiters of track and field". Beijing: BEIJING SPORT UNIVERSITY PRESS, 2005.

Yongxin Wang was born in Henan, China, in 1978, and received the B.S and M.S degree in physical education from Henan University, Kaifeng, China, in 2002 and 2010 respectively.

He has been a lecturer for 3 years at the department of Physical Education of Nanyang Institute of Technology in Nanyang, China. He researched in martial art and also focused on network, software guide. He joined the projects researched in volleyball teaching and studying, and in comparative research in the current physical education of the universities between China and USA.

Mr wang was proficient in many kinds of physical arts and also interested in computer science and algorithm .

Jianying Wang was born in Gansu, China, and received the B.S degree in applied mathematics from Lanzhou University in 2004, Lanzhou, China.

She has been a lecturer for 2 years and worked at the department of applied mathematics in Nanyang Institute of Technology, Nanyang, China. She researched at the ordinary differential equation, published papers about Modified Models for Time Series and compiled the Advanced mathematics textbook, she also took part in the projects of research and application related to Nonconforming Finite Element Method for a kind of Nonlinear Parabolic Equations, and Intelligent Control System based on Fan Boolean Algebra, etc.

Ms wang also achieved the first place in the competition of Teaching Technique and the first Award of Henan Province Zone in the China Undergraduate Mathematical Contest in Modeling in 2005 and 2007. Ms Wang focused on the applied mathematics and she was always interested in the computer scince and algorithm.

Qiufen Wang was born in Henan, China, in 1978, and received the B.S degree in computer science and application in 2002 from Henan University, Kaifeng, China, and M.S degree in computer science in 2005 from Zhejiang Science and Technology University, Hangzhou, China.

She has been a lecturer for 3 years and worked at the department of Computer Science and Technology in Nanyang Institute of Technology, Nanyang, China.

She researched at the software science and the design of web-based inquiry learning platform for college science education, signal processing systems software. she published several papers about operating system principles and the application of competition mechanism in teaching operating system. She was in charge of several important computer projects of Research and Application related to Ajex technique.

Ms Wang focused on the applied technique of computer science, and she was well up to the computer scince and algorithm.

Research on a New Kind of Robust Backstepping Filter Derivative Control Method

Yue Zhi

Dept of Control Engineering, Naval Aeronautical and Astronautical University Yantai, China
Email: zyue060610@126.com

Junwei Lei and Jinyong Yu

Dept of Control Engineering, Naval Aeronautical and Astronautical University Yantai, China
Email: {leijunwei, yujinyong1024}@126.com

Abstract—A novel Lyapunov function, which contains a concept of transfer function, was constructed to prove the rightness of a new kind of robust backstepping filter derivative control method. Meanwhile, the relationship between Lyapunov function and transfer function was established, which is an important concept that can be applied in a large family of control systems. Also, the backstepping design technology is perfectly integrated with the PID control method, which was testified by the simulation result. And comparing with pure derivative method, better performance was achieved by the adopting of filter derivative method.

Index Terms—Lyapunov function, Transfer function, Filter Derivative, Backstepping

I. INTRODUCTION

Transfer function is the most useful and important concept in classic control theory. We have many methods to analysis the stability of transfer functions such as solving the root of polynomial of denominator[1,2]. Lyapunov function is one of the most important tools in modern control theory. It can be used for the design of all kinds of systems, especially for nonlinear complex big systems. It is also a necessary tool to analysis the stability of control system designed with modern control strategies such as adaptive control or backstepping technology or robust control, etc[3-6]. Although there is no universal rules to construct a Lyapunov function for a real system, the Lyapunov function method is used widely for the design and analyze of nonlinear systems and most researchers like to use it because it is almost the only effective meant to analyze a complex nonlinear system[7-13]. Also it is a meaningful method because it reveals the relationship between the stability of a system and the virtual energy of a system, so it make the analysis of the stability of a system to be a obvious simple question from the energy point of view sight.

It is obvious that transfer function can be a useful and meaningful part if it can be included in the controller designed with modern control theory. But how to integrate the above two kinds of ideas? And how to analysis the stability of such kind of systems designed with modern control theory and transfer function theory?

It is an interesting problem. Some of linear or nonlinear control methods can be explained by the introducing of a Lyapunov function, so it can make those methods easier to be accepted by researchers. Also, it is meaningful to find a Lyapunov function for an accepted old method because we can get more close to the essence of stability of nonlinear systems by thinking this problem from two or more different angles[14-47]. In this paper, the traditional design method, such as PID control and filters, were perfectly integrated with the modern control method such as backstepping technology. And a filter derivative was constructed, which had better performance than pure derivative control according to the numerical simulation at the end of this paper. And the most important part is that the relationship between transfer function and Lyapunov function was revealed by constructing a new type of Lyapunov functions which contains the traditional transfer function.

II. MODEL DESCRIPTION

The following two-order system is taken as an example to illustrate the hybrid control of backstepping and integral approach.

$$\begin{aligned}\dot{x}_1 &= f_1(x_1) + \Delta f_1(x) + x_2 \\ \dot{x}_2 &= f_2(x) + \Delta f_2(x) + u\end{aligned}\quad (1)$$

Assumption 1: there exists known constants c_{i0} and c_{ii} such that $|\Delta f_i(x) - \dot{x}_i^d| \leq c_{i0} + c_{ii} |z_i|$.

III. DESIGN OF CONTROLLER WITH PURE DERIVATIVE METHOD

Define a new variable as $z_1 = x_1 - x_1^d$, then the first order subsystem can be written as

$$\dot{z}_1 = f_1(x_1) + \Delta f_1(x) + x_2 - \dot{x}_1^d \quad (2)$$

Design a virtual control as

$$\begin{aligned}x_2^d &= -f_1(x_1) - k_{p1}z_1 - k_{d1}\dot{z}_1 - k_{i1} \int z_1 dt \\ &\quad - k_{s1} \text{sign}(z_1) - k_{t1} \frac{z_1}{|z_1| + \varepsilon_1}\end{aligned}\quad (3)$$

Define a new variable as $z_2 = x_2 - x_2^d$, then the following equation holds

$$(1+k_{D1})z_1\dot{z}_1 + k_{I1}z_1 \int z_i dt = z_1(\Delta f_1(x) - \dot{x}_1^d) + \\ z_1z_2 - k_{P1}z_1^2 - k_{s1}|z_1| - k_{i1}\frac{z_1^2}{|z_1| + \varepsilon_1} \quad (4)$$

Considering the second subsystem, we have

$$\dot{z}_2 = f_2(x) + \Delta f_2(x) + u - \dot{x}_2^d \quad (5)$$

Design the control as

$$u = -f_2(x) - z_1 - k_{P2}z_2 - k_{D2}\dot{z}_2 - k_{I2} \int z_2 dt \\ - k_{s2}sign(z_2) - k_{i2}\frac{z_2^2}{|z_2| + \varepsilon_2} \quad (6)$$

So the equation can be arranged as

$$(1+k_{D2})z_2\dot{z}_2 + k_{I2}z_2 \int z_2 dt = z_2(\Delta f_2(x) - \dot{x}_2^d) - \\ z_1z_2 - k_{P2}z_2^2 - k_{s2}|z_2| - k_{i2}\frac{z_2^2}{|z_2| + \varepsilon_2} \quad (7)$$

Choose the Lyapunov function as

$$V = \frac{1}{2} \sum_{i=1}^2 (1+k_{Di})z_i^2 + k_{Ii}(\int z_i dt)^2 \quad (8)$$

Solve the derivative of the Lyapunov function and get

$$\dot{V} = \sum_{i=1}^2 z_i (\Delta f_i(x) - \dot{x}_i^d) - k_{Pi}z_i^2 - k_{si}|z_i| - k_{Ii}\frac{z_i^2}{|z_i| + \varepsilon_i} \quad (9)$$

According to the assumption, there exist parameters k_{Pi} and k_{si} which is big enough such that

$$\dot{V} \leq 0 \quad (10)$$

Now, it is easy to prove that the system is stable.

IV. EXAMPLE AND SIMULATION

Considering the above two-order system, we choose

$$\begin{aligned} f_1(x_1) &= 3x_1, \Delta f_1(x) = \sin(x_1 x_2) + x_1 \\ f_2(x) &= 4x_2, \Delta f_2(x) = x_2 \cos(x_1 x_2) \end{aligned} \quad (11)$$

Then the two order system can be written as

$$\begin{aligned} \dot{x}_1 &= 3x_1 + \sin(x_1 x_2) + x_1 + x_2 \\ \dot{x}_2 &= 4x_2 + x_2 \cos(x_1 x_2) + u \end{aligned} \quad (12)$$

We define the desired signal as $x_1^d = 5$ set the control parameters as $k_{P1} = 5$, $k_{I1} = k_{s1} = k_{i1} = 5$, $k_{P2} = 20$, and do the simulation with parameter as $k_{D1} = 0$ and $k_{D2} = 1$ respectively, the simulation result can be shown as Fig.1 and Fig. 2. We can know from the above figures that the overshoot can be reduced by adopting the backstepping and PD hybrid control method, also the performance of the system can be improved. But we

should also point out that if the derivative coefficient does not increase properly, oscillation will be caused (For example, we choose $k_{D1} = k_{D2} = 1$ to do the simulation and the bad performance can be shown as Fig 3) or the system will become unstable (For example, we choose $k_{D1} = 2, k_{D2} = 1$ to do the simulation and the bad performance can be shown as Fig.4). Also, in some situation, the demand for simulation algorithm will become very strict if the pure derivative item is used. In other hand, it is very difficult to get some derivative signals in many actual systems, so we will consider adopting a filter derivative to take place the pure derivative.

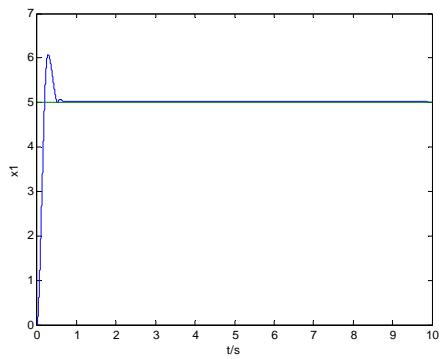


Figure 1. Curve of $x1(k_{D1} = 0)$

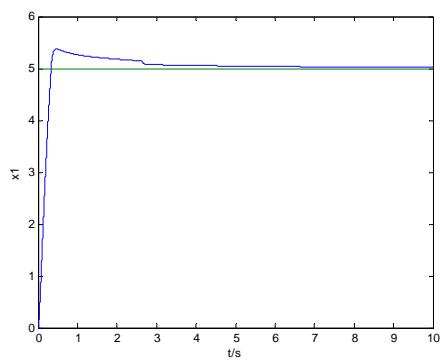


Figure 2. Curve of $x1 (k_{D1} = 1)$

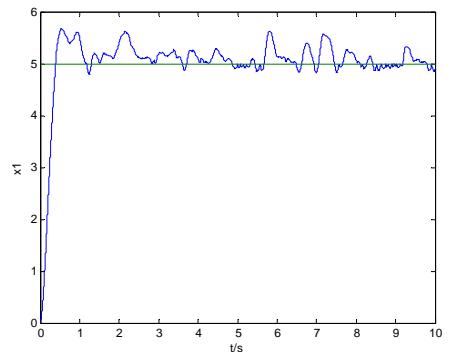
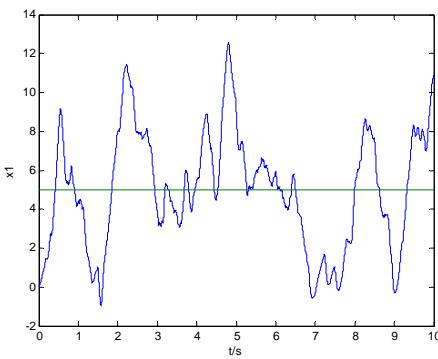


Figure 3. Curve of $x1 (k_{D1} = k_{D2} = 1)$

Figure 4. Curve of x_1 ($k_{D1} = 2, k_{D2} = 1$)

V. PROOF OF FILTER DERIVATIVE METHOD

Considering the above system

$$\begin{aligned}\dot{x}_1 &= f_1(x_1) + \Delta f_1(x) + x_2 \\ \dot{x}_2 &= f_2(x) + \Delta f_2(x) + u\end{aligned}\quad (13)$$

We define $z_1 = x_1 - x_1^d$ and the first subsystem can be written as

$$\dot{z}_1 = f_1(x_1) + \Delta f_1(x) + x_2 - \dot{x}_1^d \quad (14)$$

Design a virtual control as

$$\begin{aligned}x_2^d &= -f_1(x_1) - k_{p1}z_1 - k_{D1} \frac{s}{\tau_1 s + 1} z_1 \\ &\quad - k_{I1} \int z_1 dt - k_{s1} \text{sign}(z_1) - k_{i1} \frac{z_1}{|z_1| + \varepsilon_1}\end{aligned}\quad (15)$$

where $\frac{s}{\tau_1 s + 1}$ is a filter which is used to get an approximate derivative of the error

$$L = \frac{1}{2} \frac{z^2}{\tau s + 1} = \frac{J}{\tau s + 1} \geq 0, J = \frac{1}{2} z^2 \quad (16)$$

Then we have $\tau \dot{L} + L = J$ and solve the derivative we get

$$\begin{aligned}\dot{L} &= \frac{1}{\tau} J - \frac{1}{\tau} V = \frac{1}{\tau} \frac{1}{2} z^2 - \frac{1}{\tau} \frac{z^2}{\tau s + 1} \\ &= \frac{1}{2} \frac{1}{\tau} z \left[z - \frac{z}{\tau s + 1} \right] = \frac{1}{2} z \left[\frac{\tau s z}{\tau s + 1} \right]\end{aligned}\quad (17)$$

Design $z_2 = x_2 - x_2^d$ then the following equation holds

$$\begin{aligned}z_1 \dot{z}_1 + k_{I1} z_1 \int z_1 dt + \left(\frac{1}{2} \frac{k_{D1} z_1^2}{\tau_1 s + 1} \right)' \\ = z_1 (\Delta f_1(x) - \dot{x}_1^d) + z_1 z_2 - k_{p1} z_1^2 - k_{s1} |z_1| - k_{i1} \frac{z_1^2}{|z_1| + \varepsilon_1}\end{aligned}\quad (18)$$

Considering the second subsystem

$$\dot{z}_2 = f_2(x) + \Delta f_2(x) + u - \dot{x}_2^d \quad (19)$$

We design the control as

$$\begin{aligned}u &= -f_2(x) - z_1 - k_{p2} z_2 - k_{D2} \frac{s}{\tau_2 s + 1} z_2 \\ &\quad - k_{I2} \int z_2 dt - k_{s2} \text{sign}(z_2) - k_{i2} \frac{z_2}{|z_2| + \varepsilon_2}\end{aligned}\quad (20)$$

Then the system can be arranged as

$$\begin{aligned}z_2 \dot{z}_2 + k_{I2} z_2 \int z_2 dt + \left(\frac{1}{2} \frac{k_{D2} z_2^2}{\tau_2 s + 1} \right)' \\ = z_2 (\Delta f_2(x) - \dot{x}_2^d) - z_1 z_2 - k_{p2} z_2^2 - k_{s2} |z_2| - k_{i2} \frac{z_2^2}{|z_2| + \varepsilon_2}\end{aligned}\quad (21)$$

And the Lyapunov function can be chosen as

$$V = \frac{1}{2} \sum_{i=1}^2 \left(z_i^2 + k_{ii} (\int z_i dt)^2 + k_{Di} \frac{z_i^2}{\tau_i s + 1} \right) \quad (22)$$

The derivative of Lyapunov function can be computed as

$$\dot{V} = \sum_{i=1}^2 z_i (\Delta f_i(x) - \dot{x}_i^d) - k_{pi} z_i^2 - k_{si} |z_i| - k_{ii} \frac{z_i^2}{|z_i| + \varepsilon_i} \quad (23)$$

According to the assumption, there exist two big enough parameters k_{pi} and k_{si} such that

$$\dot{V} \leq 0 \quad (24)$$

So it is easy to prove that the system is stable.

VI. EXAMPLE AND SIMULATION

Now the above system is used to do the simulation and we set

$$\begin{aligned}f_1(x_1) &= 3x_1, \Delta f_1(x) = \sin(x_1 x_2) + x_1 \\ f_2(x) &= 4x_2, \Delta f_2(x) = x_2 \cos(x_1 x_2)\end{aligned}\quad (25)$$

then the second order system can be written as

$$\begin{aligned}\dot{x}_1 &= 3x_1 + \sin(x_1 x_2) + x_1 + x_2 \\ \dot{x}_2 &= 4x_2 + x_2 \cos(x_1 x_2) + u\end{aligned}\quad (26)$$

Define the desired value as $x_1^d = 5$ and use the backstepping and filter derivative hybrid control method, and choose the control parameters as $k_{p1} = 5$, $k_{ii} = 5$, $k_{si} = k_{ii} = 5$, $k_{p2} = 20$, $k_{D1} = 2$, $\tau_1 = 0.001$ and $k_{D2} = 1$, $\tau_2 = 0.1$. The simulation result can be shown as Fig.5. If we use the pure derivative but not the filter derivative, the system is unstable even though the same coefficient for derivative is used. The comparison means that the filter derivative can better

improve the performance of the system compared with pure derivative method.

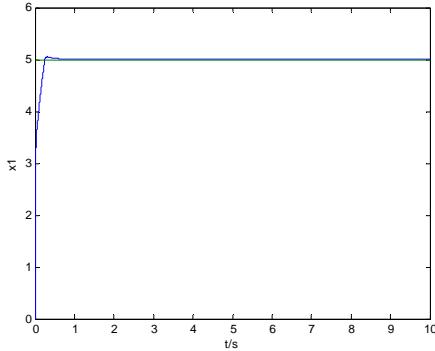


Figure 5. Curve of x_1

Considering that we further increase the uncertainties of the system, we choose $\Delta f_2(x) = x_2 \cos(x_1 x_2) x_1^2$ and use pure derivative to do the simulation. Through a series of

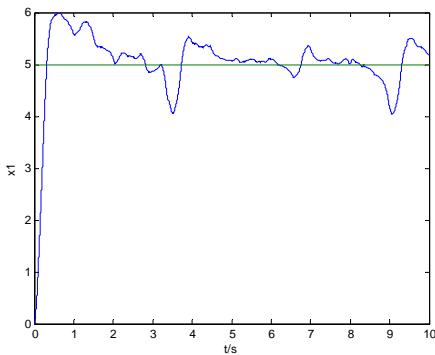


Figure 6. Curve of x_1

simulations, we found a better one with parameters as $k_{D1} = 0.5, k_{D2} = 1$, and the simulation result can be shown as Fig 6. It is obvious that if we want to reduce the oscillation and make the curve smooth, we need to increase the derivative coefficients. But the increase of the derivative caused the system to be unstable unexpectedly, which can be shown as Fig 7 where the

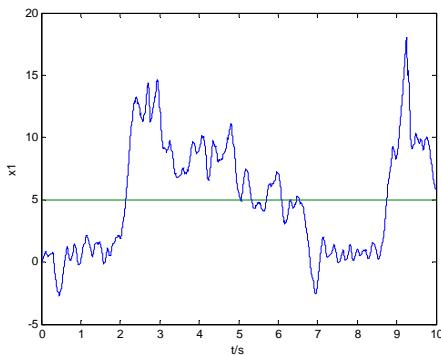


Figure 7. Curve of x_1

parameters are chosen as $k_{D1} = 2, k_{D2} = 1$.

Later, used the filter derivative algorithm and chose parameters as $k_{D1} = 2, \tau_1 = 0.001$, $k_{D2} = 1, \tau_2 = 0.1$, the control effect is not very good, which can be shown as Fig 8. And we changed the parameters as

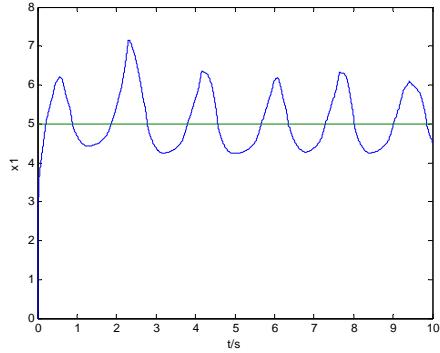


Figure 8. Curve of x_1

$k_{D1} = 5, \tau_1 = 0.001$ and $k_{D2} = 3, \tau_2 = 0.1$, and the simulation result can be shown as Fig 9. Also, the

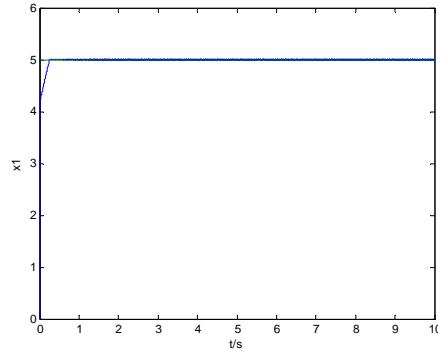
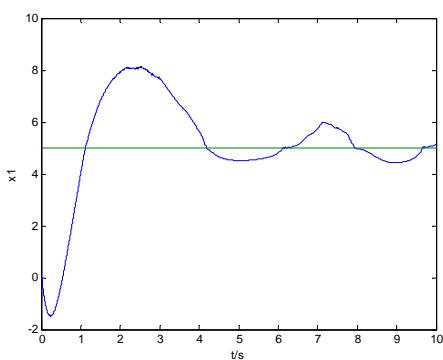


Figure 9. Curve of x_1

parameters can be chosen in a big interval, and the possibility that the system become unstable is reduced.

After that, when consider the situation that the input is limited, and assume the saturation value of input is 2500, then the response speed is slow and overshoot and oscillation are appeared as Fig 10. That is because of the uncertainties of the system and meanwhile the energy of the input is limited. Further, we increase the uncertainty of the system and choose $\Delta f_2(x) = x_2 \cos(x_1 x_2) x_1^2 + x_2^2$, and then the system is unstable. And this problem can not be solved by simply increasing the coefficient of integral item or derivative item. Another reason is that the assumption of this paper does not satisfied, so it is an important part of our future work.

Figure 10. Curve of x_1

VII. CONCLUSIONS

Above of all, the main contribution of this paper can be concluded as following points. First, a novel type of filter derivative method was introduced to robust backstepping design. Second, the relationship between transfer function (an important concept in classic control theory) and Lyapunov function (one of the most popular concept in modern control theory) was established. What is very different from other papers is that detailed and elaborate simulations were done in this paper to research the problem that how much uncertainties and how big an amplitude of uncertainties can be solved by this filter derivative method.

At last, the paper also shows some situations that the proposed new method can not cope with and the reasons are given to explain the result. Also, it points out our future research field.

ACKNOWLEDGMENT

The authors wish to thank their friend Heidi in Angels (a town of Canada) for her help , and thank Amado for his many helpful suggestions.

REFERENCES

- [1] Hao Lei,Wei Lin, Universal adaptive control of nonlinear systems with unknown growth rate by output feedback, *Automatica* 42 (2006) 1783-1789
- [2] Hao Lei,Wei Lin , Adaptive regulation of uncertain nonlinear systems by output feedback: Auniversal control approach , *Systems & Control Letters* 56 (2007) 529 - 537
- [3] Swaroop D, Gerdes J C, Yip P P, Hedrick J K. Dynamic surface control of nonlinear systems. In: Proceedings of the American Control Conference. Albuquerque, New Mexico: IEEE, 1997. 3028-3034
- [4] Swaroop S, Hedrick J K, Yip P P, Gerdes J C. Dynamic surface control for a class of nonlinear systems. *IEEE Transactions on Automatic Control*, 2000, 45(10): 1893-1899
- [5] Zhang, T., Ge, S.S. & Hang, C.C. Adaptive neural network control for strict-feedback nonlinear systems using backstepping design. *Automatica*, Vol. 36. No. 12 (2000) 1835-1846
- [6] Chiman Kwan and F. L. Lewis. Robust backstepping control of nonlinear systems using neural networks. *IEEE Transactions on systems, man and cybernetics*,Vol. 30. No. 6 (2000) 753-766
- [7] Junwei Lei, Xinyu Wang, Yinhua Lei, How many parameters can be identified by adaptive synchronization in chaotic systems? [J] *Physics Letters A*, Volume 373, Issue 14, 23 March 2009, Pages 1249-1256
- [8] Junwei Lei, Xinyu Wang, Yinhua Lei, A Nussbaum gain adaptive synchronization of a new hyperchaotic system with input uncertainties and unknown parameters, [J] *Communications in Nonlinear Science and Numerical Simulation*, Volume 14, Issue 8, August 2009, Pages 3439-3448
- [9] Xinyu Wang,Junwei Lei, Yinhua Lei,Trigonometric RBF Neural Robust Controller Design for a Class of Nonlinear System with Linear Input Unmodeled Dynamics[J]. *Applied Mathematical and Computation*, 185 (2007) 989 - 1002
- [10] Wang Xinyu, Lei Junwei, Design of Global Identical Terminal Sliding Mode Controller for a Class of Chaotic System,2006 IMACS Multiconference on Computational Engineering in Systems Applications(CESA ' 2006), Tsinghua University Press,2006.10
- [11] Xinyu Wang, Hongxing Wang, Hong Lei, Junwei Lei , Stable Robust Control for Chaotic Systems Based on Linear-parameter-neural-networks, The 2nd International Conference on Natural Computation (ICNC'06) and the 3rd International Conference on Fuzzy Systems and Knowledge Discovery (FSKD'06), Sep 24-28, 2006, Xi-an, China
- [12] GE S S , Wang C , Lee T H. Adaptive backstepping control of a class of chaotic systems[J]. *Int J Bifurcation and chaos*. 2000, 10 (5): 1140-1156
- [13] GE S S , Wang C , Adaptive control of uncertain chus's circuits[J]. *IEEE Trans Circuits System*. 2000, 47(9): 1397-1402
- [14] Alexander L, Fradkov, Markov A Yu. Adaptive synchronization of chaotic systems based on speed gradient method and passification[J]. *IEEE Trans Circuits System* 1997,44(10):905-912
- [15] Dong X. Chen L. Adaptive control of the uncertain Duffing oscillator[J], *Int J Bifurcation and chaos*. 1997,7(7):1651-1658
- [16] Tao Yang, Chun-Mei Yang and Lin-Bao Yang, A Detailed Study of Adaptive Contorl of Chaotic Systems with Unknown Parameters[J] . *Dynamics and Control*. 1998,(8):255-267
- [17] M.T. Yassen, Adaptive chaos control and synchronization for uncertain new chaotic dynamical system , *Physics Letters A* 350 (2006) 36 - 43
- [18] H.Nijmeijer,M.Y.mareels.An observer looks at synchronization.IEEE Trans.on Circuits Syst.I.1997, (44):882-890
- [19] M.T.Yassen.Chaos synchronization between two different chaotic systems using active control.*Chaos,Solitons & Fractals*.2005;23(4):131-140.
- [20] H.N.Agiza,M.T.Yassen. Synchronization of Rossler and Chen chaotic dynamical systems using active control.*Physics Letters A*.2001,(278):191-197.
- [21] Ming-chung Ho,yao-Chen Hung.Synchronization of two different systems by Using generalized active control. *Physics Letters A*.2002,(301):424-428.
- [22] E.M.Elbabbasy,H.N.Agiza,M.M.El-dessoky.Adaptive synchronization of lv System with uncertain parameters. *Chaos,solitons and fractals*. 2004,(21):657-667

- [23] Bemdarom. An Adaptive approach to the control and synchronization of continuous time chaotic systems[J]. Int J Bifurcation and chaos. 1996, 6(3):557-568.
- [24] GE S S , Wang C , Lee T H . Adaptive backstepping control of a class of chaotic systems[J]. Int J Bifurcation and chaos. 2000, 10 (5): 1140-1156
- [25] GE S S , Wang C , Adaptive control of uncertain Chua's circuits[J]. IEEE Trans Circuits System. 2000, 47(9): 1397-1402
- [26] Alexander L, Fradkov, Markov A Yu. Adaptive synchronization of chaotic systems based on speed gradient method and passification[J]. IEEE Trans Circuits System 1997,44(10):905-912
- [27] Dong X. Chen L. Adaptive control of the uncertain Duffing oscillator[J], Int J Bifurcation and chaos. 1997,7(7):1651-1658
- [28] Tao Yang, Chun-Mei Yang and Lin-Bao Yang, A Detailed Study of Adaptive Control of Chaotic Systems with Unknown Parameters[J] . Dynamics and Control. 1998,(8):255-267
- [29] Oscar. Fuzzy control of chaos [J] . Int J Bifurcation and chaos. 1998.8(8): 1743-1747
- [30] Liang Chen, Guanrong Chen, Lee Yang-Woo, Fuzzy modeling and adaptive control of uncertain chaotic systems[J]. Information Sciences. 1999, 121(1):27-37
- [31] Liang Chen,Guanrong Chen.Fuzzy prescriptive control of uncertain chaotic Systems using time series[J].Int J Bifurcation and chaos.1999,9(40):757-767.
- [32] Chin Teng Lin.Controlling chaos by GA-based reinforcement learning neural network [J]. IEEE Trans On neural networks. 1999,10:846-859
- [33] Gao Ping Jing et al. A simple global synchronization criterion for coupled chaotic systems. Solitons and fractals. 2003,(15):925-935
- [34] Jianping Yan, Changpin Li. On synchronization of three chaotic systems. Chaos,Solitons and fractals.2005,(23):1683-1688
- [35] C.Sarasola,F.J.Torrealdea,A.D Anjou. Feedback synchronization of chaotic Systems[J].Int.J.Bifurcation and chaos.2003,13(1):177-191
- [36] M.T.Yassen.Chaos synchronization between two different chaotic systems using active control.Chaos,Solitons & Fractals.2005:23(4):131-140
- [37] H.N.Agiza,M.T.Yassen. Synchronization of Rossler and Chen chaotic dynamical systems using active control.Physics Letters A.2001,(278):191-197
- [38] Ming-chung Ho, Yao-Chen Hung.Synchronization of two different systems by Using generalized active control. Physics Letters A.2002,(301):424-428.Jiang G.P., Tang K.S. A global synchronization criterion for coupled chaotic systems via unidirectional linear error feedback approach. Int. J. Bifurcation Chaos, 2002, 12(10):2239-2253
- [39] Bu S.L., Wang S.Q.An algorithm based on variable feedback to synchronize chaotic and hyperchaotic systems. Physica D,2002,164:45-52
- [40] Suykens. J.A.K, Robust nonlinear H^∞ synchronization of chaotic Lur's system, IEEE Trans Circuit & System1997, 44(10):891-903
- [41] Michele B, Donato C. Synchronization of hyper-chaotic circuits via continuous feedback control with application to secure communication. Int. J. Bifurcation and Chaos, 1998,8(10):2031-2040
- [42] Alexander L.Fradkov, A Yu.Markov. Adaptive synchronization of chaotic system phased on speed gradient method and passification. IEEE Trans on Circuit Syst.(I),1997,44(10):905-912
- [43] Young_Hoon Joo, Leang-San shieh, Guangrong Chen, Hybrid state-space fuzzy model-based controller with dual-rate sampling for digital control of chaotic systems, IEEE Trans on Fuzzy Syst.(I),1999,7(4):394-408
- [44] Alexander Pogromsky, Henk Nijmeijer. Observer-based robust synchronization of dynamical systems. Int,J. Bifurcation and Chaos,1998,8(11):2243-2254
- [45] E.M.Elabbas,H.N.Agiza,M.M.El-dessoky.Adaptive synchronization of lv System with uncertain parameters. Chaos,solitons and fractals. 2004,(21):657-667
- [46] M.T.yassen. Adaptive synchronization of Rosser and lü systems with fully uncertain parameters. Chaos,solitons and fractals.2005,(23):1527-1536
- [47] Ju H.Park. Adaptive synchronization of Rossler system with uncertain Parameters.Chaos, solitons and fractals.2005,(22):1-6.



Yue Zhi was born in Zouping, Shandong province of China on Dec, 1982. He received the B.S. degree in Electronic Automation and received the M.S. degree in Guidance, Control and Navigation from Naval Aeronautical and Astronautical University, China, in 2005 and 2008 respectively.

Now he is a Ph.D. candidate of Control Science and Engineering, Naval Aeronautical and Astronautical University, China. His current research interests include actuator and sensor nonlinearities, robust control and isolation of nonlinear systems.



Junwei Lei was born in Chibi, Hubei province of China on 9th Nov, 1981. He received the B. Eng degree in Missile Control and Testing and the Master Degree in Control Theory and Control Engineering from Naval Aeronautical Astronautical University, Yantai of China in 2003 and 2006 respectively. After that he continued his study there and received the Doctor degree in Guidance , Control and Navigation in 2010 .

He worked in NAAU as an assistant teacher in 2009 and became a lecture in 2010. His present interests are neural networks, chaotic system control, variable structure control and adaptive control.

Mr Lei studied in Canada Military Force and Language School in Base Bodern of Toronto from Jan to Jun in 2010.

Jinyong Yu was born in Haiyang, Shandong province of China in 1976. He received the B. Eng degree in Automation Engineering in Central South University in 1999 and received the Master Degree in Control Theory and Control Engineering from Naval Aeronautical Astronautical University, Yantai of China in 2002. He received the Doctor degree in Guidance, Control and Navigation from NAAU in 2006 respectively.

He worked in NAAU as an assistant teacher in 2003 and became a lecture in 2005. In 2008, he was promoted to be a vice professor. His present interests are aircraft control and adaptive control, guidance and navigation.

Intelligent PID Controller on Soft Computing

Honghua Xu, Xiaoqiang Di

Changchun University of Science and Technology, Changchun, China

Email :{honghuax,xiaoqiangd}@126.com

Huamin Yang, Guangcai Cui

Changchun University of Science and Technology, Changchun, China

Email : {huaminyan,guangcaicu}@yahoo.com

Abstract—The PID control algorithm is used for the control of almost all loops in the process industries, and is also the basis for many advanced control algorithms and strategies. For complex systems, controller's structure and parameters need to rely on experience and on line tuning to determine. Application of soft computing techniques can realize self-tuning of controller parameters. In this paper, the author applies fuzzy system, neural network technique and genetic algorithm to PID controller in order to achieve intelligent adjustment the controlled variables. Simulation results show that the application of soft computing techniques in the controller has better control quality.

Index Terms—intelligent, fuzzy system, neural network, proportion integration differentiation, genetic algorithms

I. INTRODUCTION

PID(Proportion Integration Differentiation) control is one of the best widely applied controlling methods in industrial producing process because it is simple and robust[1].

Conventional PID control is a most common method used in the process of manufacturing. It has found its application in many trades. It has many advantages such as ease of controlling, ease of realizing and debugging on the spot etc[2]. In the actual control process, the controlled process is nonlinear, time-varying and indefinite, especially for some more complicated systems. Conventional PID is hard to meet this requirement.

In the actual design of control system, there still exists trial and error method to solve[3]. With quick development of relative technology, control system needs higher norms and self-adaptation technique of PID control parameters in order to suit for more complicated situations and higher demand for norms. So the control engineers are on look for automatic intelligent tuning procedures.

In this paper, the parameters of PID controller are tuned for controlling some rock mechanics testing machine[4]. Conventional PID performance has been compared and analyzed with the intelligent tuning techniques like fuzzy system, neural network, and Genetic algorithm. Soft computing techniques based tuning methods have proved their excellence in giving better results by improving the steady state characteristics.

II. PID CONTROL MECHANISM

A. PID algorithm in analog system

In simulation system, the equation of PID algorithm[5] is:

$$U(t) = K_p \left[e(t) + \frac{1}{T_I} \int e(t) dt + T_D \frac{de(t)}{dt} \right] \quad (1)$$

Where, $U(t)$ is the output signal of regulator; $e(t)$ is the error signal of regulator, which equals to the difference between measurement value and preset value; K_p is the proportional coefficient of regulator; T_I is the integral time of regulator; T_D is the differential time of regulator.

B. Digital PID Algorithm

Because computer control is a sampling control, it can only calculate control volume according to the error of the sampling moments[2,6]. Therefore, in the process of computer control, we must make dispersed treatment to equation (1) to use digital difference equation to take the place of differential equation in successive system. Then integral term and differential term can be expressed with summation and increment equation.

$$\int_0^n e(t) dt = \sum_{j=0}^n e(j) \Delta t = T \sum_{j=0}^n e(j) \quad (2)$$

$$\frac{de(t)}{dt} \approx \frac{e(k) - e(k-1)}{\Delta t} = \frac{e(k) - e(k-1)}{T} \quad (3)$$

We put (2) and (3) into (1) and then we can get dispersed PID equation.

$$U(k) = K_p \left\{ e(k) + \frac{T}{T_I} \sum_{j=0}^k e(j) + \frac{T_D}{T} [e(k) - e(k-1)] \right\} \quad (4)$$

Where, $\Delta t = T$ is the sampling period and T should be sufficiently small in order to obtain a definite precision; $e(k)$ is the error value of No.k sampling moment; $e(k-1)$ is the error value No.(k-1) sampling

moment; k is sampling serial number, $k = 0, 1, 2, \dots$; $U(k)$ is the regulator output of No. k sampling moment.

From equation (4), if we want to get $U(k)$, not only do we need the error signals $e(k)$ and $e(k-1)$, of this time and last time but we need plus previous error signals $e(j)$ from integral term, i.e. $\sum_{j=0}^k e(j)$. By this

way, it is quite complicated to calculate, however, there needs much memory to store $e(j)$. So it is quite inconvenient to control directly with equation (4). Let's adjust it as follows:

$$U(k-1) = K_p \left\{ E(k-1) + \frac{T}{T_i} \sum_{j=0}^{k-1} E(j) + \frac{T_p}{T} [E(k-1) - E(k-2)] \right\} \quad (5)$$

Subtract equation (5) from equation (4), we can get:

$$\begin{aligned} U(k) &= U(k-1) + K_p [e(k) - e(k-1)] + \\ &K_I e(k) + K_D [e(k) - 2e(k-1) + e(k-2)] \end{aligned} \quad (6)$$

Where, $K_I = K_p \frac{T}{T_i}$ is integral coefficient;

$K_D = K_p \frac{T_p}{T}$ is differential coefficient.

From equation (6), to get the output value $U(k)$, it is ok to get $U(k-1)$, $e(k)$, $e(k-1)$ and $e(k-2)$. It is much easier than using equation (4).

Although equation (6) is revised only a little bit, it brings the following advantages.

- (1). Because the output is increment, it has small influence of malfunction.
- (2). It's easy to realize the non-interfering switches of control patterns.
- (3). No integral out of control. It's easy to get better regulation quality.

III. SOFT COMPUTING TECHNIQUES

A. PID Controller on Fuzzy System

Fuzzy system is a system based on knowledge or rules. Its core is knowledge base made up of so-called IF-THEN rules which come from experts or the knowledge in this field[7-9]. Then, we apply all these rules to single system. Different fuzzy systems should adapt to different combination rules.

1. fuzzy system

The basic structure of a pure fuzzy system is as shown in Figure 1. Rule base can be shown as a set of several IF-THEN rules. Through combining these IF-THEN rules, fuzzy inference engine will decide how to map the fuzzy set of input domain of discourse $U \subset R^n$ on the fuzzy set of output domain of discourse $V \subset R$.

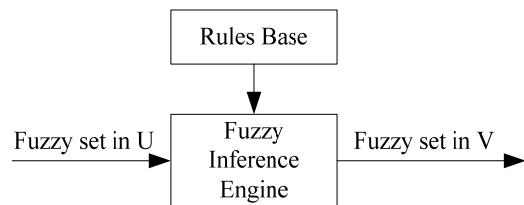


Figure 1. The structure of a pure fuzzy system.

The leading problem of a pure fuzzy system is that its input variable and output variable are all fuzzy sets, whereas, the input and output in engineering system are all real-valued variables[10,11].

To apply fuzzy system to engineering system, a simple method is to add a fuzzifier at the input end. Its purpose is to change the real-valued variable to a fuzzy set. At the output end, we add a defuzzifier, which is to switch the fuzzy set to real-valued variables. Figure 2 shows us a fuzzy system with a fuzzifier and a defuzzifier.

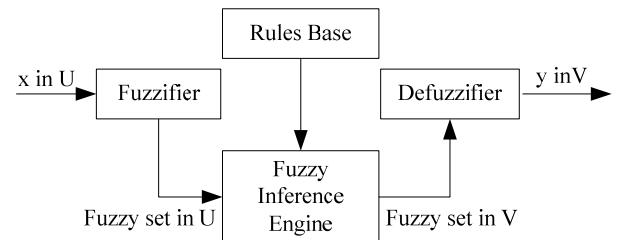


Figure 2. The structure of a fuzzy system with a fuzzifier and a defuzzifier.

Rule base is composed of fuzzy IF-THEN rule set. It is the core of a fuzzy system. Other components of a fuzzy system carry out these rules in a reasonable and efficient way. Rule base consists of the following IF-THEN rules.

$$R_u^{(l)} : \text{If } x_1 \text{ is } A_1^l \text{ and...and } x_n \text{ is } A_n^l, \text{ then } y \text{ is } B^l \quad (7)$$

Where, A_i^l and B^l are fuzzy set of $U_i \subset R$ and $V \subset R$ respectively, $x = (x_1, x_2, \dots, x_n)^T \in U$ and $y \in V$ are input and output variables of the fuzzy system.

To provide a general knowledge equation, equation (7) consists of the following rules:

- (1). “incomplete rule”

If x_1 is A_1^l and...and x_m is A_m^l , then y is B^l in $m < n$

- (2). “or rule”

If x_1 is A_1^l and...and x_m is A_m^l or x_{m+1} is A_{m+1}^l and...and x_n is A_n^l , then y is B^l

- (3). “single fuzzy statement”

y is B^l

- (4). “stepwise variation rule”

If x is smaller, then y is bigger

The function of a fuzzy inference engine is to compound the mapping of a fuzzy IF-THEN rule which exists in rule base. This mapping is from fuzzy set A^l in

U on fuzzy set B' in V. Fuzzy inference engine is designed according to the following rules.

- (1). In conformity with the experience of experts.
- (2). High computer efficiency.
- (3). Specific requirements of some features.

Product inference engine has the advantage of ease of computing and can directly reflect actual problems. Its equation is as shown in equation (8).

$$\mu_{B'}(y) = \max_{l=1}^M \left[\sup_{x \in U} (\mu_{A'}(x) \prod_{i=1}^n \mu_{A'_i}(x_i) \mu_{B'}(y)) \right] \quad (8)$$

If you offer a fuzzy set A' in U, the inference engine can produce a fuzzy B' in V.

A fuzzifier can be defined as a mapping of $x^* \in U \subset R^n$ on fuzzy set A' in U. The fuzzifier is designed according to the following rules.

- (1). The fuzzifier should consider the fact that it is at clear point x^* that input occurs.
- (2). If the input of a fuzzy system is interfered by noise, the fuzzifier needs the ability to overcome the influence of noise.
- (3). The fuzzifier should contribute to simplify the computing of fuzzy inference engine.

The triangle fuzzifier maps $x^* \in U$ on fuzzy set A' in U. It has the following triangle membership function.

$$\mu_A(x) = \begin{cases} \left(1 - \frac{|x_1 - x_1^*|}{b_1}\right) * ... * \left(1 - \frac{|x_n - x_n^*|}{b_n}\right) & \text{if } |x_i - x_i^*| \leq b_i \ (i=1,2,...,n) \\ 0 & \text{others} \end{cases} \quad (9)$$

Where, parameter $b_i (i=1,2,...,n)$ is positive number, t-nom * usually chooses algebraic product operator or min operator.

A defuzzifier can be defined as a mapping of fuzzy set B' (output of the fuzzy inference engine) in $V \subset R$ on $y^* \in V$. Defuzzifier methods should consider the following rules.

- (1). Point y^* should visually represent B' .
- (2). Ease of computing is quite important because of controller's real-time operation.
- (3). Continuity is needed. Slight change of B' shouldn't cause substantial variation of y^* .

The center average defuzzifier can be defined as the following equation (10).

$$y^* = \frac{\sum_{l=1}^M \bar{y}' \omega_l}{\sum_{l=1}^M \omega_l} \quad (10)$$

Where, \bar{y}' is the center of the l th fuzzy set; ω_l is the height. The center average defuzzifier is ease of computing, reasonable and visual.

2. The structure of pid controller with fuzzy system

The PID controller with a fuzzy system takes error e and error variation ec as its input. It utilizes fuzzy system to tune PID parameters on line, which can satisfy the requirements of different e and ec to PID parameters' adaptation. Its structure is as shown in Figure3.

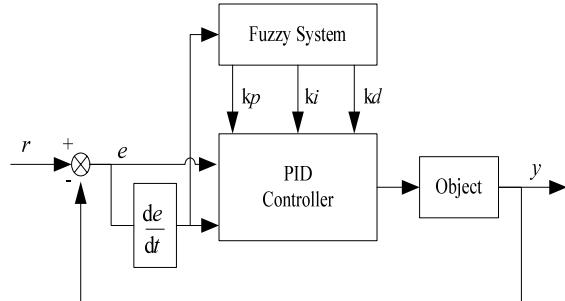


Figure 3. The structure of PID controller with a fuzzy system.

where the core of the PID controller is to build up fuzzy rule base on the basis of engineering technicians' knowledge and actual operation equation. The fuzzy sets of e and ec are {NB,NM,NS,O,PS,PM,PB} in which all the components stand for negative big, negative medium, negative small, zero, positive small, positive medium, positive big. The following fuzzy rules are established.

If (e is NB) and (ec is NM) then (kp is PB)(ki is NB)(kd is NS)

If (e is NM) and (ec is PS) then (kp is PS)(ki is NS)(kd is NM)

.....

If (e is PM) and (ec is NB) then (kp is PS)(ki is Z)(kd is PB)

If (e is PB) and (ec is PM) then (kp is NB)(ki is PB)(kd is PS)

B. PID Controller on RBFNN

Radial basis function neural network (RBFNN) is a three-layer feed forward[12]. With radial basis function network, we can construct a self-learning PID controller with parameters k_p, k_i and k_d .

1. Structure of Neural Network PID Controller

The PID control strategy based on neural network is shown in Figure 4. The regulator consists of two parts:

PID controller can directly closed-loop control the objects and three parameters k_p, k_i, k_d can be adjusted on line.

According to the operation condition of the system, neural network regulates the parameters of PID controller to optimize some performance index[13,14]. Through self-learning of neural network and adjustment of weight coefficient, neural network can obtain some optimized PID control parameter.

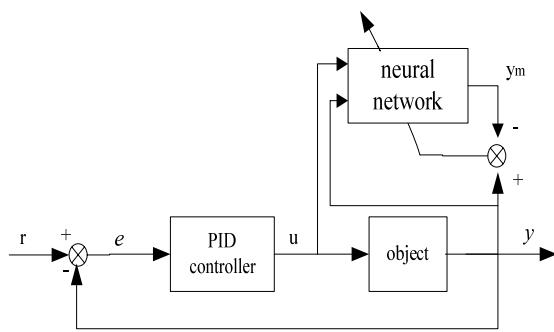


Figure 4. Structure of neural network PID controller

Where, r is the system input, and y is the system output. y_m is the output of neural network. u is the output of PID controller.

2. Neural Network Identification Algorithm

Radial basis function neural network consists of input layer, hidden layer and output layer, in which input layer consists of signal source nodes, the number of units in hidden layer is decided by the need of the described questions and output layer responds to the input layer[15]. The mapping from input to output is nonlinear, whereas the mapping from hidden layer space to output space is linear. Therefore, the learning speed will be much faster and at the same time local minimum problems will be avoided.

Suppose neural network is made up of n input nodes, m hidden nodes and one output node. In the structure of neural network, the input vector of network input layer can be expressed as $X = [x_1, x_2, \dots, x_n]^T$, the radial basis vector of the network hidden layer can be expressed as $H = [h_1, h_2, h_3, \dots, h_m]^T$, where h_j is radial basis function may choose different algorithms. Here, we'd like to choose Gaussian function, i.e. (11).

$$h_j = \exp\left(-\frac{\|X - C_j\|^2}{2b_j^2}\right), j = 1, 2, \dots, m \quad (11)$$

Where, C_j is the center vector of No.j node in network. The expression is as follows:

$$C_j = [c_{j1}, c_{j2}, \dots, c_{jn}]^T, (i = 1, 2, \dots, n) \quad (12)$$

Suppose the base width vector of the network is $B = [b_1, b_2, \dots, b_m]^T$, b_j is the base width parameter of node j and b_j is greater than zero. The network weight vector is $W = [w_1, w_2, \dots, w_m]^T$. The output expression of identification network is as follows.

$$y_m(k) = \sum_{j=1}^m w_j h_j \quad (13)$$

In order to obtain better result of network identification, while revising weight coefficient, an inertia term should be added in. According to low

gradient descent, the interactive algorithm of output weight, node center and node width parameter can be expressed as equations (14) (15) and (16)

$$\begin{aligned} \Delta W_j &= [y(k) - y_m(k)]h_j \\ W_j(k) &= W_j(k-1) + \eta \Delta W_j + \alpha[w_j(k-1) - w_j(k-2)] \\ &+ \beta[W_j(k-2) - W_j(k-3)] \end{aligned} \quad (14)$$

$$\begin{aligned} \Delta b_j &= [y(k) - y_m(k)]w_j h_j \frac{\|X - C_j\|^2}{b_j^3} \\ b_j(k) &= b_j(k-1) + \eta \Delta b_j + \alpha[b_j(k-1) - b_j(k-2)] \\ &+ \beta[b_j(k-2) - b_j(k-3)] \end{aligned} \quad (15)$$

$$\begin{aligned} \Delta c_{ji} &= [y(k) - y_m(k)]w_j \frac{X_j - C_{ji}}{b_j^2} \\ c_{ji}(k) &= c_{ji}(k-1) + \eta \Delta c_{ji} + \alpha[c_{ji}(k-1) - c_{ji}(k-2)] \\ &+ \beta[c_{ji}(k-2) - c_{ji}(k-3)] \end{aligned} \quad (16)$$

Where, η is learning rate, and α, β is inertia coefficients whose value range is between 0 and 1.

The sensitivity information algorithm of the output of the controlled object to the control input is as follows.

$$\frac{\partial y(k)}{\partial u(k)} \approx \frac{\partial y_m(k)}{\partial u(k)} = \sum_{j=1}^m w_j h_j \frac{c_{1j} - x_1}{b_j^2} \quad (17)$$

In the actual application, x_1 is the input term of identification network and its value is $u(k)$.

3. PID Parameter Tuning Principle online

According to Figure 4, control error is:

$$e(k) = r(k) - y(k) \quad (18)$$

Suppose there are parameters c_1, c_2, c_3 .

$$c_1 = e(k) - e(k-1)$$

$$c_2 = e(k)$$

$$c_3 = e(k) - 2e(k-1) + e(k-2)$$

PID control strategy uses equation (6) to calculate.

$$u(k) = u(k-1) + k_p c_1 + k_i c_2 + k_d c_3$$

(19)

Neural network tuning target is :

$$E(k) = \frac{1}{2} e^2(k)$$

Low gradient descent should be adopted to adjust parameters k_p, k_i and k_d . Its calculation process is as follows:

$$\Delta k_p = -\eta \frac{\partial E}{\partial k_p} = -\eta \frac{\partial E}{\partial y} \frac{\partial y}{\partial u} \frac{\partial u}{\partial k_p} = \eta e(k) \frac{\partial y}{\partial u} c_1 \quad (20)$$

$$\Delta k_i = -\eta \frac{\partial E}{\partial k_i} = -\eta \frac{\partial E}{\partial y} \frac{\partial y}{\partial u} \frac{\partial u}{\partial k_i} = \eta e(k) \frac{\partial y}{\partial u} c_2 \quad (21)$$

$$\Delta k_d = -\eta \frac{\partial E}{\partial k_d} = -\eta \frac{\partial E}{\partial y} \frac{\partial y}{\partial u} \frac{\partial u}{\partial k_d} = \eta e(k) \frac{\partial y}{\partial u} c_3 \quad (22)$$

Where, $\frac{\partial y}{\partial u}$ is information of controlled object which could be obtained through neural network identification.

C. PID Controller on GA

1. Genetic algorithms

Genetic algorithms (GA), which imitate genetic mechanism and the theory of evolution, is the most optimized parallel random searching method[16,17]. GA introduces the rule of the theory of evolution "survival of the fittest" to the coded string group composed of the most optimized parameters. According to adaptation value functions, GA which selects units through copying, crossing and varying retains high adaptation value units which form new units that inherit the information from previous generation and optimize the previous generation in turn. Repeatedly in this way, the adaptation ability of units is continuously improved till a satisfactory situation. Its algorithm is quite easy and can be concurrently dealt with to get overall optimization.

PID controller to which GA is applied can employ GA to optimize parameters. This method is the one which doesn't need any initial information and can find the most optimized and the most efficient combination method.

The elements for GA:

- (1) Chromosome coded method. Basic GA employs fixed length binit to mark units in the group and its allele is composed of symbol set {0,1}.
- (2) Unit's sufficiency evaluation. The directly proportional probability of GA and unit sufficiency can decide how much probability it will be from the units of current group to the group of the next generation. To correctly calculate the probability, all the sufficiency must be firstly defined from objective function to unit sufficiency.
- (3) Genetic operator. Basic GA employs genetic operators which include selection operator which selects operation ratio, one-point crossover operator which employs crossover operation, simple mutation operator which employs mutation operation or uniform mutation operator.
- (4) The operation of basic GA demands predetermined parameters:

M: group size i.e. the number of units in the group.
G: the last evolution algebra of GA.

P_c : crossover probability.

P_m : mutation probability.

2. PID setting principle based on GA

- (1). Parameter determination and representation.

At first parameter scope should be determined and this scope should be given by users. Next, in terms of precision, coding should be made. Binary character string should be chosen to represent each parameter and relation should be set up. At last, binary character string should be attached to form a long binary character string which is that object that GA may operate.

(2). Selecting initial population.

Because all the processes can be realized through programming, initial populations are generated by computer random. As for binary coding, firstly evenly distributed random numbers between 0 and 1 are generated[18]. Then 0 stands for the generated numbers between 0 and 0.5; 1 stands for the generated numbers between 0.5 and 1. Besides, the size of population is decided by the complexity of computer.

(3). Fitting function determination.

Under constraint conditions, common optimizing algorithm can obtain a group of satisfactory parameters. In the process of designing, the best parameter should be chosen from this group of parameters. There are three aspects we use to measure the indicators of the controlling system, i.e. stability, accuracy and rapidity. Rise time reflects the rapidity of the system. The shorter the rise time spends, the faster the controlling carries out.

(4). GA operation.

At first, we should employ sufficiency ration to copy, i.e. through fitting function we can get fitting value to get corresponding copy probability of each string. The product of copy probability and the number of strings of each generation is the number of strings copied in the next generation. The bigger copy probability is, the more offspring the next generation will have. On the contrary, it will die out.

Through copy, crossover and mutation, initial population obtains a new population, which can be taken into fitting function after decoding. Then we should check if it satisfies termination condition. If not, the above operation will repeat till making it.

In order to obtain satisfactory dynamic characteristics of transient process, performance indicator of error absolute value time should be adopted as the minimum objective function of parameter selection. To avoid excessive control energy, quadratic term of control input should be added to objective function. The following equation is chosen as the most optimized indicator of parameter selection.

$$J = \sum_{t=0}^T (w_1 |e(t)| + w_2 u^2(t)) dt + w_3 t_u \quad (23)$$

Where, $e(t)$ is the system error. $u(t)$ is controller output. t_u is rise time. w_1, w_2, w_3 are weights.

To avoid overshoot, punitive function is adopted, i.e. once overshooting, the overshoot is considered as a term of the most optimized indicator. At that time, the most optimized indicator is as follows:

$$J = \hat{\Delta}_0^*(w_1 |e(t)| + w_2 u^2(t) + w_4 |e(t)|)d_t + w_3 t_u$$

if $e(t) < 0$

(24)

Where, w_4 is a weight and $w_4 \gg w_1$.

IV. EXPERIMENT'S RESULT AND ANALYSIS

Some rock mechanics testing machine has the functions of real-time control, monitoring and feedback. Its control process is nonlinear and highly complicated. The control system asks for realizing automated control with high accuracy and high reliability to the testing machine.

The transfer function of the rock mechanics testing machine system is equation below:

$$G(s) = \frac{0.4}{0.015s^3 + 0.523s^2 + 0.12s}$$

Here, traditional PID control algorithm and intelligent PID control algorithm based on soft computing are adopted to respectively make simulation experiment and analysis to the system.

The executive body of the system adopts driving step motor to control. Its control algorithm is as shown in equation (6). The process of control algorithm is as follows:

- (1) initializing parameter.
- (2) sample $r(k)$ and $y(k)$.
- (3) calculating error value.
- (4) calculating controller output;
- (5) renewing parameters and returning to procedure (2).

Controller parameters are set as follows:

$k_p = 20$, $k_i = 0.2$ and $k_d = 8$. The simulation result is shown in Figure 5.

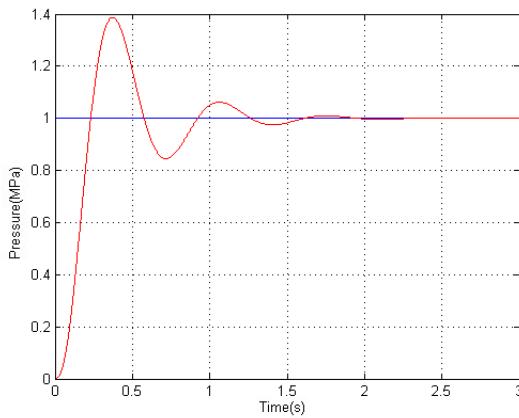


Figure 5. The result of conventional PID.

During intelligent tuning, we define the variation scopes of system error e and error variation rate ec as the domain of discourse on fuzzy set and then decide its fuzzy subsets and their membership. While controlling on line, control system together fuzzy system's handling and computing finishes proofreading to PID parameters on line. Its working flow is as follows:

- (1) Taking current sampling value.
- (2) Computing $e(k)$, $ec(k)$ and $e(k-1)$.
- (3) Fuzzifying $e(k)$ and $ec(k)$.
- (4) Fuzzy adaptation to Δk_p , Δk_i and Δk_d .
- (5) Computing current k_p , k_i and k_d .
- (6) Effect of PID controller's output on the controlled system.

The simulation result is shown in Figure 6.

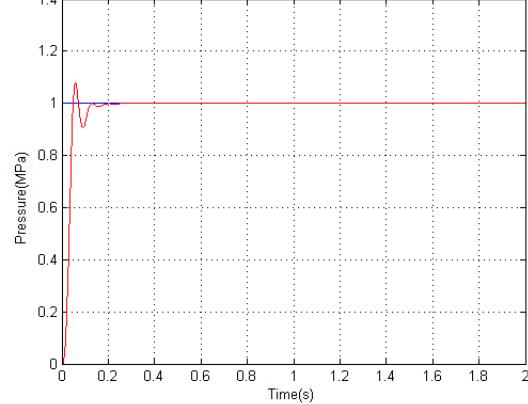


Figure 6. The result of PID controller on fuzzy system.

The procedures of control strategy using RBFNN are as follows:

- (1) Making sure the number of network input nodes, the number of hidden layers, learning rate and inertia coefficient to get center vector of hidden nodes, base width vector of the network and the initial value of weight vector.
- (2) Sampling to get input r and output y , and calculating error e in terms of equation (18).
- (3) Calculating the output u of regulator according to equation (19) to control the controlled object real time.
- (4) Calculating network output ym and adjusting center vector of network hidden nodes and base width vector and weight vector to obtain network identification information.
- (5) Adjusting parameters of regulator in terms of equations (20) (21) and (22).
- (6) Going back to procedure (2).

The structure of neural network is 3-6-1. The three inputs of network identification are $u(k)$, $y(k)$ and $y(k-1)$. Learning rate $\eta = 0.25$. Inertia coefficients $\alpha = 0.05$, $\beta = 0.01$. The simulation result is shown in Figure 7.

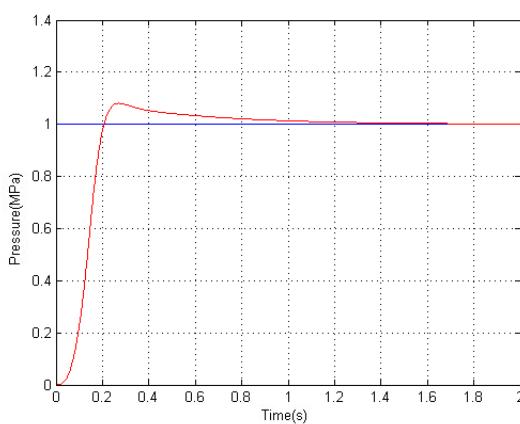


Figure 7. The result of PID controller on RBFNN.

The specific procedures of optimizing PID parameters using GA are as follows:

- (1) making sure the scope and coding length of each parameter and coding.
- (2) generating initial population made up of n random units.
- (3) decoding respective parameter value from the population and then using this parameter value to get cost function value and fitting function value.
- (4) using copy, crossover and mutation operators to operate population to produce a new population.
- (5) operating procedures (3) and (4) till parameter convergence or achieving desired indicator.

Adopting binary coding is to avoid using the length of 10 bits binary coding string to mark three decision variables k_p , k_i and k_d . The equation (24) is used to select the most optimized indicator. The number of samples used in GA is Size=30. Crossover probability and mutation probability are $P_c=0.60$, $P_m=0.001$ [1:1:Size]x0.001/Size. The value span of parameter k_p is [0.20]. The value span of k_i and k_d is [0.1], $w_1=0.99$, $w_2=0.001$, $w_3=1.0$, $w_4=100$.

The simulation result is shown in Figure 8.

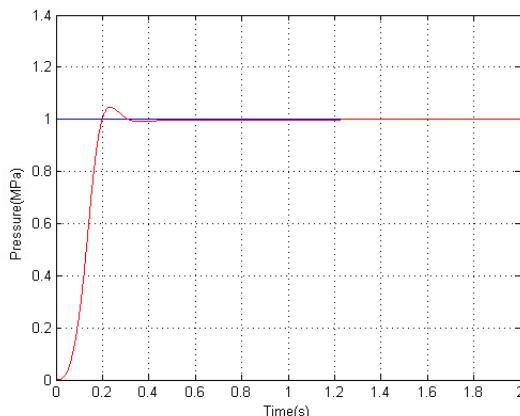


Figure 8. The result of PID controller on GA.

From above Figures, traditional control algorithm needs longer tuning time, rise time and larger overshoot. The controller employing soft computing technique to realize can obtain better tuning performance for the system tuning. The specific comparison is shown in tableI.

TABLE I.
COMPARISON OF STEADY STATE RESPONSES

	Peak over shoot (%)	Settling time (sec)	Rise time (sec)
Con.	39.6	2.52	0.362
FS	12.7	0.26	0.067
NN	8.2	1.7	0.261
GA	6.7	1.23	0.236

V. CONCLUSION

The PID control algorithm is the basis for many advanced control algorithms and strategies in the process industries. In order to use a controller, it must first be tuned to the system. This tuning synchronizes the controller with the controlled variable, thus allowing the process to be kept at its desired operating condition. Conventional PID control method can be used only when the process mode is of a certain type. The study on PID controller using soft computing technique embodies its excellence. Soft computing has the features of intelligence and self-adaptation, so there is no need for the controlled objects constructing precise model. The system has stronger capacity of resisting disturbance. For complicated controlled objects, the system can obtain higher control quality, hence robustness.

While solving complicated actual problems, the system may combine knowledge, techniques and methods of various sources. Various computing models of soft computing technique aren't mutually exclusive and teamwork will own bigger excellence. The systems employing different computing techniques are regarded as complementary mixed intelligent systems, which will be a further study content.

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support of The Scientific and Technological Key Support Projects with the project number 20090307.

REFERENCES

- [1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.
- [2] J.Astrom and T.H.agglund. PID Controllers: Theory, design and tuning. Instrument society of America, 2nd edition, 1995.
- [3] Kiam Heong Ang, Gregory Chong: "PID Control System Analysis, Design, and Technology" IEEE Transactions on

- Control Systems Techonology, Vol.13, No.4, July 2005 pp. 559-576.
- [4] Honghua Xu, Guangcui Cui, Ji Li. FuzzyPID Control Technology Application in Rock-testing System. Journal of changchun University of Science and technology. 2005,28(3), pp. 42-45.
- [5] Neuron Chip Distributed Communication and Control Processors. Motolora Inc., 1994.
- [6] Qing-Guo Wang,Tong-Heng Lee: "PID Tuning for Improved Performance" IEEE Transactions on Controll Systems Technology, Vol.7, No.4 July 1999, pp. 457-465.
- [7] Gravel, A. and H. Mackenberg, Mathematical analysis of the Sugeno controller leading to general design rules. Fuzzy Sets and Systems, 85, pp. 165-175, 1995.
- [8] Kadmiry B.,Bergsten P. "Robust Fuzzy Gain Scheduled visual-servoing with Sampling Time Uncertainties" IEEE Int. Symposium on Intelligent Control (CSS / ISIC), Taiwan 2004.
- [9] Z. Y. Zhao, M. Tomizuka, and S. Isaka, .Fuzzy gain scheduling of PID controllers,. IEEE Transactions on Systems, Man, and Cybernetics, vol. 23, no. 5, pp. 1392.1398, 1993.
- [10] Honghua Xu, Guang Dong, Weili Shi. "An Intelligent PID Controller Based on Fuzzy System", IEEE Conf. ICEEE2010, vol. 5, pp. 3117-3121, 2010.
- [11] Di Nola, A., S. Sessa, W. Pedrycz, and E. Sanchez, Fuzzy Relation Equations and Their Applications to Knowledge Engineering, Kluwer, Boston, 1989.
- [12] Honghua Xu, Yinghua Tian. "An Intelligent PID Controller on RBFNN", IEEE Conf. Electrical Engineering and Automation Control, vol. 4, pp. 1256-1259, 2011.
- [13] Cheeseman, P., "An inquiry into computer understanding," (with discussions) Computational Intelligence, 4., pp.58-142, 1988.
- [14] Bonadeo N H, Erland J, Gammon D, et al. Coherent Optical control of the Quantum State of a Single Quantum Dot. Science, 282: pp.1473-1475, 1998.
- [15] Karayiannis, N. B. Reformulated radial basis neural networks trained by gradient descent. IEEE Trans. on Neural Networks. 10(3): pp. 657-671, 1999.
- [16] Jukka Lieslehto "PID controller tuning using Evolutionary programming" American Control Conference, VA June pp. 25-27, 2001
- [17] Hang, C. C. and K. K. Sin, "A comparative performance study of PID auto-tuners", IEEE Control Syst. Msg., vol. 11, pp. 41-47, 1991.
- [18] T. Ota and S. Omatsu, "Tuning of the PID control gains by GA," in Proc. IEEE Conf. Emerging Technol. Factory Automation, Kauai, HI, Nov., pp. 272-274, 1996.



Honghua Xu graduated from Changchun University of Science and Technology in 2004. He is a lecturer at Changchun University of Science and Technology and mainly engaged in teaching and research work. He has participated in some provincial and ministerial research projects. He has presented many papers in journals and international conferences. Main areas of research are computer simulation and artificial intelligence.



Xiaoqiang Di graduated from Changchun University of Science and Technology in 2002. He is a lecturer at Changchun University of Science and Technology and mainly engaged in teaching and research work. He has participated in some provincial and ministerial research projects. He has presented many papers in journals and international conferences. Main areas of research are image processing and artificial intelligence.



Huamin Yang graduated from Dalian University of Technology in 1985. He is a professor and doctoral tutor at Changchun University of Science and Technology and mainly engaged in teaching and research work. He is a member of the Computer Federation of Province. He has chaired many provincial and national research projects, including some of projects received awards. He has presented more than 70 papers in journals and international conferences. Main areas of research are computer simulation, image processing and artificial intelligence.



Guangcui Cui graduated from Beijing Institute of Technology in 1986. He is a professor and doctoral tutor at Changchun University of Science and Technology and mainly engaged in teaching and research work. He is a member of the Computer Federation of Province. He has chaired several provincial research projects, including some of projects received awards. He has presented more than 40 papers in journals and international conferences. Main areas of research are artificial intelligence and Chinese information processing.

An Approach to Interactive Affective Learning Algorithms

Chong Su

College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China
Email: suchong@mail.buct.edu.cn

Hongguang Li

College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China
Email: lihg@mail.buct.edu.cn

Abstract—To solve multi-objective decision-making problems without explicit mathematical description for objective functions, traditional interactive evolutionary computing approaches are usually limited in searching ability and vulnerable to human's subjectivity. Motivated by this observation, a novel affective computing and learning solution adapted to human-computer interaction mechanism is explicitly proposed. Therein, a kind of stimulating response based affective computing models (STAM) is constructed, along with quantitative relations between affective space and human's subjective preferences. Thereafter, affective learning strategies based on genetic algorithms are introduced which are responsible for gradually grasping essentials in human's subjective judgments in decision-making, reducing human's subjective fatigue as well as making the decisions more objective and scientific. To exemplify applications of the proposed methods, test functions are suggested to case studies, giving rise to satisfied results and showing validity of the contribution.

Index Terms—Interactive evolutionary computing (IEC), affective computing, affective learning

I. INTRODUCTION

Interactive evolutionary computation (IEC)^[1, 2] has been recognized as an effective approach to accommodate complex decision-making relevance, where the objectives involved are usually far from completely structured and quantified, or implicitly expressed due to the uncertainty associated with decision makers' preferences. As a kind of evolutionary algorithms demanding human direct participation, its prominent features lie in the fact that human may influence the evolutions by directly evaluating individual performances. For example, Lai& Chang (2009)^[3] put forward a method to solve image segmentation problems using interactive evolutionary computation under genetic algorithms frameworks. John etc. (2010)^[4] introduced hierarchical

concepts into interactive evolutionary computing, separating global and local searching thereby solving regional nonlinear optimization problems involved in video working environments. Nonetheless, IEC suffers human's limitations in discriminating abilities that account for slow convergence, even if the individuals are in the vicinity of the optimum solutions. In other terms, IEC usually reveals weakness in local searching capability as well as in overcoming human's subjective deviation for quantitatively evaluating excellent solutions.

Alternatively, affective computing becomes a new emerging research issue that targets the emotion of computer agents to improve their autonomy, adaptability, and social interactive ability. Currently, majority of related researches are concerned with theoretical relevance and simulation issues. Specifically, hidden Markov chain models (HMM) (Picard, 1997)^[9] have been adapted to represent human's spontaneous transitions with certain external stimuli added. Therein, adjustable parameters could be modified to manage the speed and amplitude associated with transitions. Additionally, approaches to assign affective matrix are available for personality based OCC models (Onony et al, 1998)^[10], which could effectively express human's affective changes when externally stimulated. Besides, conceptual affective entropy, energy, strength and threshold were provided by some researchers (Wang and Zhao, 2001)^[11] as well. Whereas, it turns out that the above-cited models are confined by specific models in terms of applications. For instance, hidden Markov chain models are adapted to the spontaneous transitions driven by singular external stimulus only, which inevitably discourages the applications in human-computer interactive environments. It is conceivable that the problem of how quantitatively describe affective transitions impacted by persistent external stimuli still represents a challenge.

In contrast to the achievement of affective computing, relatively few attentions have been paid on affective learning ever since. In the descriptions of virtual affective models, Aard et al (2000)^[12] addressed affective learning problems but rarely covered them in detail. Ishihara

Corresponding author: Hongguang Li; Tel.: 86-10-64434797; Fax: 86-10-64442932.

&Fukuda (2000)^[13] employed affective algorithms to deal with traffic signal systems. Owing to simple affective models employed, they only described affective responses to a certain external stimulus, rarely taking into account of shifted external stimuli. Hongguang li & Chong Su (2011)^[16] introduced a method of affective computing and gave an approach to multi-objective fitness index computing, but they didn't introduce how to make computer learn from human's affective rule in interactive multi-objective decision-making.

Motivated by getting rid of human's participation in IEC thereby reducing the impacts of subjective judgments on decision-making, this paper proposes an approach to affective computing adapted to human-computer interactive mechanism. Successively, a kind of interactive affective learning models is presented, along with enabling algorithms, aiming at grasping essentials in human's affective preferences towards multi-objective decision-making. The proposed methodologies are applied to numerical examples, leading to satisfactory results.

The remainder of this paper is organized as follows. Section 2 presents a method of affective computing with continuous external stimuli. This is followed in Section 3 by an in-depth investigation on interactive affective learning philosophies as well as enabling algorithms. Section 4 provides case studies consisting in test functions. Section 5 concludes the article and assesses the future perspectives.

The remainder of this paper is organized as follows. Section 2 presents a method of affective computing with continuous external stimuli. This is followed in Section 3 by an in-depth investigation on interactive affective learning philosophies as well as enabling algorithms. Section 4 provides case studies consisting in test functions. Section 5 concludes the article and assesses the future perspectives.

II. AFFECTIVE COMPUTING MODELS

In order to quantitatively compute the affective transitions driven by affective stimulus persistently, an improved affective computing model, STAM (stimulated transferring affect model), is suggested as follows.

Definition 1(Affective stimulus):

A shifted external environment which human could feel and response emotionally refers to an affective stimulus.

Step 1: According to [9, 10、11], Picard^[9] mainly did much research in human's negative affect, such as "negative、frustration、boredom、confusion". Here, we pay attention to human's common affect, such as "happiness", "calmness" and "sadness".

specify an affective space as

$$\Phi_s^n = \{s_1^n, s_2^n, s_3^n\} \quad (1)$$

where, the components, s_i^n ($i = 1, 2, 3$), represent the affective strengths associated with "happiness", "calmness" and "sadness", respectively, constrained by $(s_1^n + s_2^n + s_3^n = 0)$ similar to that of HMM. It is noted

that, at any time, only the affective component with positive and largest absolute value takes effect.

Step 2: The affective changing quantity driven by affective stimuli is characterized by:

$$\Delta\Phi_s^n = \Phi_s^n p_s \quad (2)$$

where $p_{s(3\times3)}$ denotes the affective stimulating matrix. Thus, the changes of affective space driven by affective stimuli are described as:

$$\Phi_s^{n+1} = \Phi_s^n + \Delta\Phi_s^n \quad (3)$$

Step 3: Consider an artificial psychological stress model $y = A\varpi^x$ ^[9] as the affective stimulus description of STAM, whose output, y , changes with affective stimuli.

Step 4: In order to measure the affective changes, ϖ is suggested as the standard quantitative parameter of shifted affective stimuli, i.e. $\varpi = \partial_n - \partial_{n-1}$, where ∂_n corresponds to the affective stimulus at timescale n ($n=1, 2 \dots$). Therefore, we have $y = A(\partial_n - \partial_{n-1})^x$.

Step 5: According to Eq. (3), we get the affective space expression as follows.

$$\begin{aligned} \Phi_s^{n+1} &= \Phi_s^n + y\Phi_s^n = \Phi_s^n + AB^x\Phi_s^n \\ &= \Phi_s^n + A(\partial_n - \partial_{n-1})^x\Phi_s^n \end{aligned} \quad (4)$$

The affective components correspond to

$$s_i^{n+1} = s_i^n + AB^x s_i^n = s_i^n + A(\partial_n - \partial_{n-1})^x \cdot s_i^n \quad (5)$$

where A and x are used as adjustable parameters responsible for human's different personalities in response to affective stimuli.

In what follows, we present two fundamental properties associated with STAM.

Property 1(Uniqueness):

Only one component of the affective space accounts for the terminal state of affective transitions driven by a certain affective stimulus.

Property 2(Boundness):

Components of the affective space should take values in the range $[-1, 1]$, i.e. $s_k^n \in [-1, 1]$.

III. AFFECTIVE LEARNING ALGORITHMS

A. Affective Learning

Affective learning is aiming at gradually imitating human's affective preferences in decision-making by means of adjusting parameters of corresponding affective computing models. Consequently, during the period of human-computer interactive decision-making, computer could gradually replace the human's subjective participation, lessening human's subjective fatigue and promoting more scientific implementations.

Definition 2(Affective preferences):

Human's affective satisfaction degrees for achieved objectives in decision-making refer to affective preferences.

Property 3:

In multi-objective-attribute decision-making, changes of objective-attributes may create affective stimuli, leading to a shifted affective space, where the strengths of "happiness" (S_1^n , which is defined in formula (1)) in the terminal response could be mapped into affective preferential membership degrees.

In the presence of that $S_1(x) \in [a, b]$ and $\mu(x) \in [0, 1]$ are specified as the strengths of "happiness" and affective preferential membership degrees with respect to objective-attribute x , respectively,

the linear mapping $\frac{S_1(x) - a}{b - a} = \mu(x)$ could be employed to convert affective space into affective preferential membership.

According to the affective computing models, affective learning problems can be formulated as a constrained nonlinear programming, described as follows.

$$\begin{aligned} \min \sum_{i=1}^n [\mu(x_i) - \mu'(x_i)]^2 & (i = 1, \dots, n) \\ \text{s.t. } |S_{\text{happy}}^n| & \leq 1 \end{aligned} \quad (6)$$

where, for objective-attribute x_i , $g(f(x_i))$ corresponds to the mapping relationship from strength of "happiness" into membership degrees of affective preferences, $\mu(x_i)$ corresponds to the desired preferential membership functions. Genetic algorithms can be invoked to solve this optimization problem

B. Multi-objective Assessment

Initially, the relative importance of an objective-attribute needs to be particularly considered. We postulate that a set of objective-attributes is expressed as

$$\text{Property - Set} = \{p_1, p_2, \dots, p_i\} (i = 1, 2, \dots, n) \quad (7)$$

where, $p_i (i = 1, 2, \dots, n)$ signify the values of objective-attributes, n is the number of the objective-attributes.

Definition 3(Multi-objective fitness index):

Specify a multi-objective fitness index as:

$$k = \sum_{i=1}^n \lambda_i \varphi_i \quad (8)$$

where λ_i is defined as the relative importance degrees of attribute p_i which are constrained by $\sum_{i=1}^n \lambda_i = 1$.

Obviously, $k \leq 1$ implies that the objective is perfectly achieved.

Definition 4(Fitness of objective attribute):

A parameter φ_i is introduced to measure the fitness of attribute p_i , which is defined as

$$\varphi_i = \frac{p_i'}{\hat{p}_i} \times 100\%, \quad (9)$$

where p_i' and \hat{p}_i are the actual and admissible values of p_i at the current time period, respectively.

Algorithm 1(Multi-objective attribute weight):

According to fuzzy AHP [15], the relative importance of each objective-attribute can be identified as follows:

Step1: Establish priority relation matrix $F = (f_{ij})_{n \times n}$, where,

$$f_{ij} = \begin{cases} 0.5, & s(i) = s(j) \\ 1.0, & s(i) > s(j) \\ 0.0, & s(i) < s(j) \end{cases} \quad (10)$$

$s(i)$ and $s(j)$ indicate the importance degrees of f_i and f_j ($i, j = 1, 2, \dots, n$), respectively.

Step2: Transform the priority relation matrix F as follows:

$$\begin{aligned} r_i &= \sum_{k=1}^n f_{ik}, k = 1, 2, \dots, n, \\ r_{ij} &= \frac{r_i - r_j}{2n} + 0.5 \end{aligned} \quad (11)$$

where $R = (r_{ij})_{n \times n}$ refer to fuzzy consistent matrix.

Step3: Compute importance degrees as follows:

$$l_i = \sum_{j=1}^n r_{ij} - 0.5, i = 1, 2, \dots, n \quad (12)$$

To normalize l_i we eventually obtain the importance degrees as

$$\lambda_i = 2l_i / [n(n-1)], i = 1, 2, \dots, n \quad (13)$$

C. Interactive Evolutionary Computing (IEC)

Step 1: Initiate $t = 0$ and create initial

population $x_l(t)$ of multi-objective decision-making solutions randomly over the global searching space;

Step 2: Specify an importance degree for each objective-attribute, and, in regard to every individual, calculate corresponding objective fitness index K based on multi-objective assessment;

Step 3: Aided by computers, human operators evaluate the excellent individuals of multi-objective decision-making solutions in terms of objective fitness index K. At the same time, computers perform affective computing and learning algorithms, generating affective evaluations of individuals for human references;

Step 4: Select excellent individuals based on human-computer interaction;

Step 5: Perform crossover and mutation operations to generate the offspring;

Step 6: Decode and return to Step 2.

Figure 1 shows the implementing steps of affective learning algorithms.

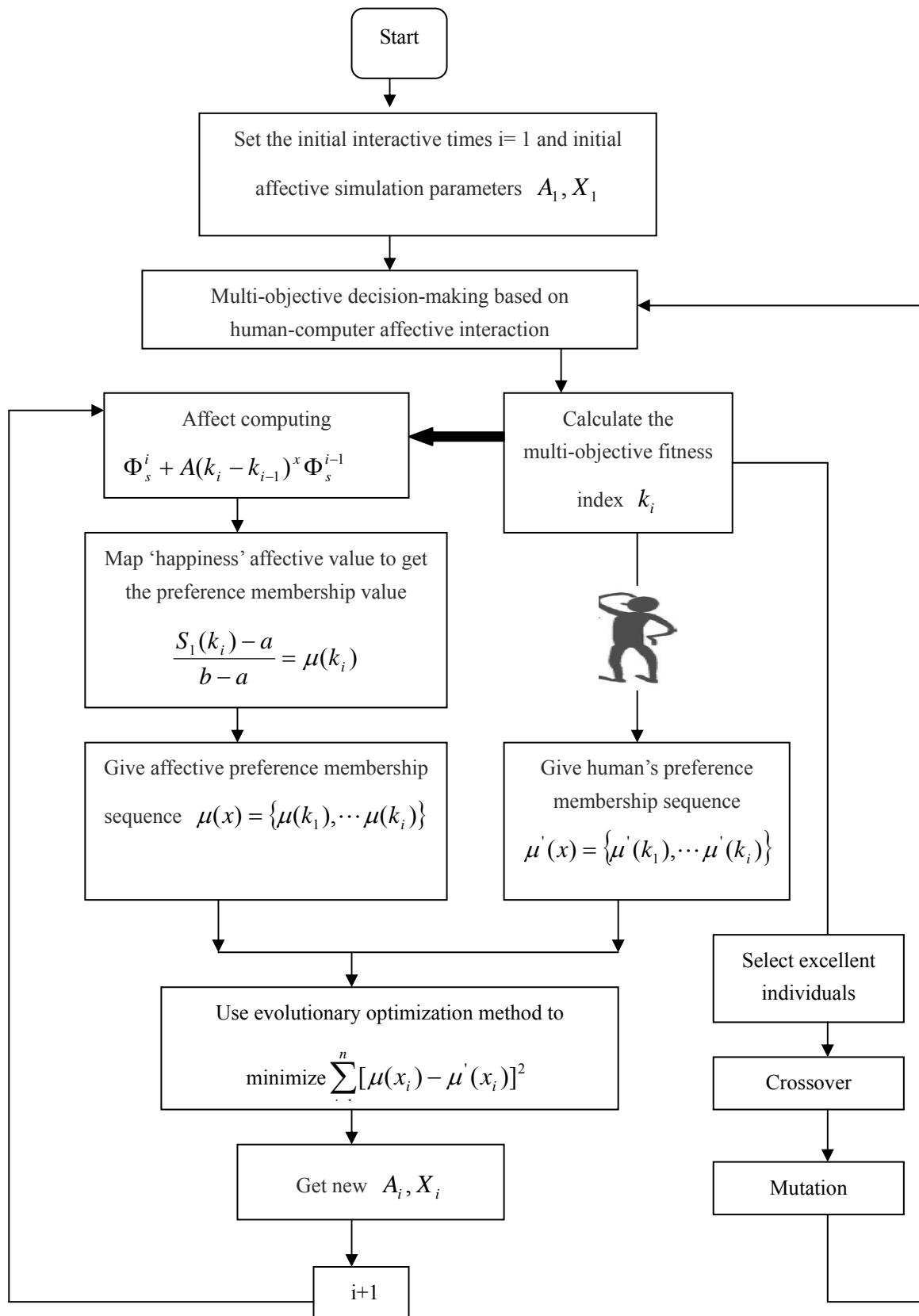


Figure 1 .Affective learning algorithms based on IEC frame

IV. CASE STUDIES

Consider a multi-objective-attribute decision-making problem consisting in three test functions characterized by following equations.

$$\begin{cases} y_1 = -a^{x+b} + c \\ y_2 = \frac{b}{x+a} + d \\ y_3 = a \sin(bx + e) \end{cases} \quad (14)$$

Specifically, decision-makers would like to attain a largest y_1 , a smallest y_2 and a most steady y_3 within a specific range of x . It is noticed apparently that the three objective-attributes may turn out somewhat conflicting due to the commonly used parameters, a , b , c , d and e .

We establish an objective-attribute^[16] set as

$$\text{Property-Set} = \{\text{high}, \text{low}, \text{slow}\} \quad (15)$$

and an objective fitness index (algorithm 1) as

$$k_n = 0.5 \cdot \varphi_1^n + 0.167 \cdot \varphi_2^n + 0.333 \cdot \varphi_3^n,$$

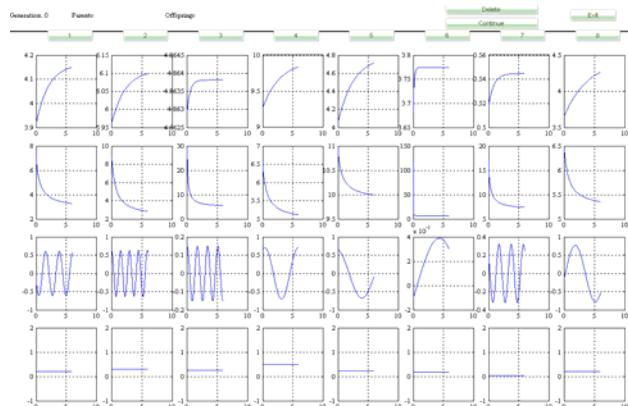


Figure 2 .Objective-attributes of the initial population

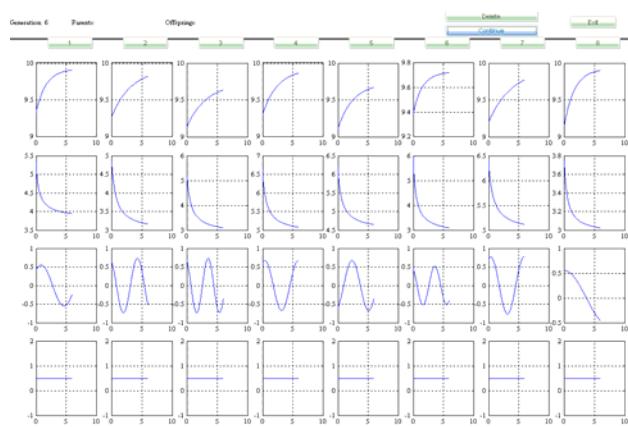


Figure 3. Objective-attributes of the last population

$$\text{Where } , \varphi_1^n = \frac{p_1^n}{\hat{p}_1} , \varphi_2^n = \frac{p_2^n}{\hat{p}_2} \text{ and } \varphi_3^n = \frac{p_3^n}{\hat{p}_3}$$

correspond to the ratios of current values with respect to desired values associated with three objective-attributes, respectively.

Along with affective learning algorithms, human-computer interactive evolutions are being performed. The performances of initial and last generation are shown in Figures 2 and 3 respectively. In addition, Figures 4 and 5 show the profiles of affective space and average fitness K with evolutions going on. To help get access to affective learning metrics, Figures 6 and 7 show approximations and approximating offsets with respect to human's affective preferential membership functions, respectively. Accordingly, Figures 8 and 9 show the evolutions of adjustable parameters associated with affective computing models, A and x

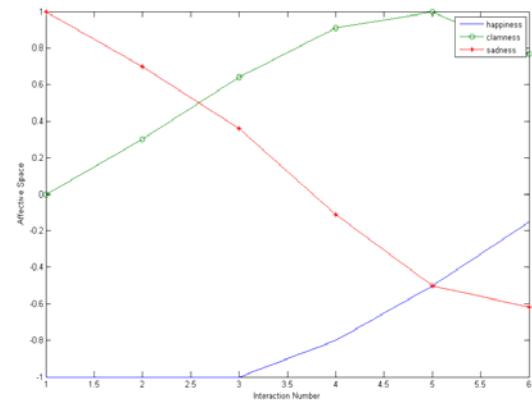


Figure 4. Affective computing curves

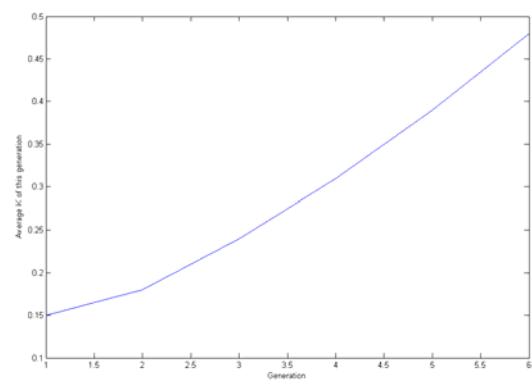


Figure 5. Curves of average K

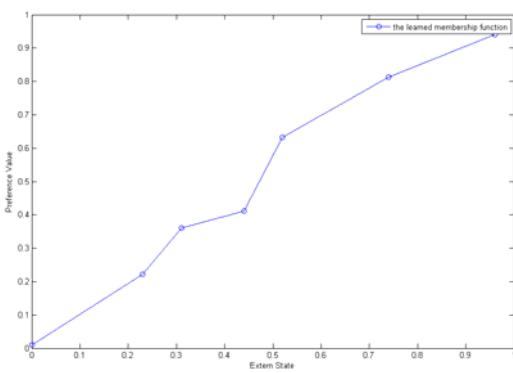


Figure 6 . Approximations of the affective preferences

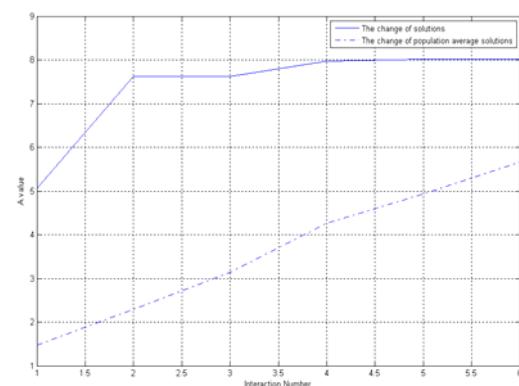


Figure 8. Evolutions of affective parameter A

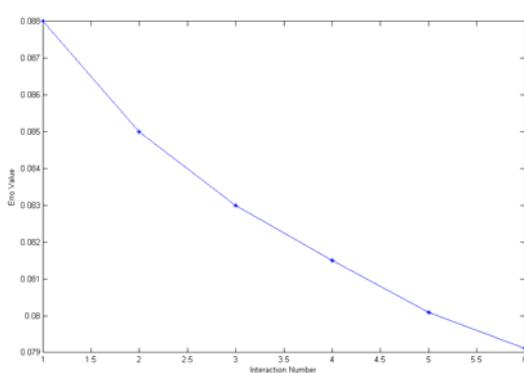


Figure 7. Approximating offsets of the affective preferences

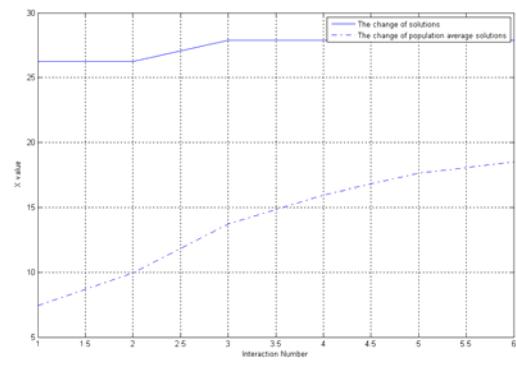


Figure 9. Evolutions of affective parameter x

V. CONCLUSIONS

Under the framework of IEC, compatible affective computing and affective learning philosophies have been extensively investigated along with enabling algorithms. In contrast to traditional IEC, the proposed approaches are recognized able to gradually grasp essentials in human's subjective judgment in decision-making, reducing human's subjective fatigue and making the decisions more objective and scientific. To exemplify their applications in industrial process further, we particularly provided an affective interactive evolutionary solution for test function which is a kind of intractable and time-consuming task. Case studies demonstrate the effectiveness and promising potentials of these contributions. Nonetheless, it should be pointed out that this research remains rather fundamental currently, which is in desperate need of further in-depth investigations on certain key issues. For example, study on problems of how to attain more appropriate subjective preferential structures in evolutions would be a potentially useful avenue for future research.

REFERENCES

- [1] ZITZLER E, THIELE L, LAUMANNS M, et al. Performance assessment of multi-objective optimizers: an

analysis and review. *IEEE Trans on Evolutionary Computation*, 17(2003), pp. 117-132.

- [2] HAO Peng, XIANG Ling. Optimal design approach for the plate-fin heat exchangers using neural networks cooperated with genetic algorithms. *Applied Thermal Engineering*, 28(2008), 28, pp. 642-650.
- [3] CHIH-CHIN LAI, CHUAN-YU CHANG. A hierarchical evolutionary algorithm for automatic medical image segmentation. *Expert Systems with Applications*, 36(2009), pp. 248-259.
- [4] VIJAY JOHN, EMANUELE TRUCCO, SPELA IVEKOVIC. Markerless human articulated tracking using hierarchical particle swarm optimization. *Image and Vision Computing*, 28(2010), pp. 1530-1547.
- [5] J.K. BURGOON, et al. Interactivity in human-computer interactive: a study of credibility, understanding, and influence. *Computers in Human Behavior*, 16(2000), pp. 553-574.
- [6] JAMES A and SEBE N. Multimodal human-computer interactive: A survey. *Computer Vision and Image Understanding*, 108(2005), pp. 116-134.
- [7] SERENKO A. Are interface agents scapegoats? Attributions of responsibility in human-agent interactive. *Interacting with Computers*, 19(2007), pp. 293-303.
- [8] ST. AMANT R, A.R. FREED and F.E. RITTER. Specifying ACT-R models of user interactive with a GOMS language. *Cognitive Systems Research*, 6(2005), pp. 71-88.
- [9] PICARD R W. *Affective Computing*. London, England: MIT Press, 1997.

- [10] ONONY A, CLORE G.L. and COLLINS A. The Cognitive Structure of Emotions. Cambridge, UK: Cambridge University Press, 1988.
- [11] WANG Zhiliang, ZHAO Yanling. An Expert System of Commodity Choose Applied with Artificial Psychology. *IEEE International Conference on Systems, Man and Cybernetics*, 2001, pp. 2326-2330.
- [12] VAN KESTEREN A-J, AKKER O D, POLE M, et al. Simulation of Emotions of agents in Virtual Environments using neural networks. In: *Proceedings of the Twenty Workshop on Language Technology 18*. Enschede, 2000, pp. 137-147.
- [13] ISHIHARA H, FUKUDA T. Traffic signal networks simulator with learning emotional algorithm. *IEEE/RSJ International Conference on Intelligent Robots and Systems*.2000, pp. 2274-2279.
- [14] VENAYAGAMOORTHY G.K., HARLY R.G. Two separate continually online-trained neurocontrollers for excitation and turbine control of a turbo generator. *IEEE Trans: Industry Applications*, 38(2002), pp. 887-893.
- [15] MIKHAILOV L and TSVETINOV P. Evaluation of services using a fuzzy analytic hierarchy process. *Applied Soft Computing*, 5(2004), pp. 23-33.
- [16] LI Hong-guang, SU Chong ,Affective Interactive agents with Applications in Control Performance Assessment ,*Computer Integrated Manufacturing System*,11(2011), pp. 2438-2446..



Chong Su was born in Tianjin, P. R. China in 1983. He is the full time Ph. D. candidate and Lecturer in Beijing University of Chemical Technology. His research interests are intelligent applications, affect computing and human-computer interaction.



Hongguang Li was born in Liaoning Province, P. R. China in 1963. He received the Ph. D. degree in East China University of Science and Technology in 2004. At present, he is professor in Beijing University of Chemical Technology. His research interests are Modeling, control and optimization of chemical process as well as computer based intelligent control for industrial plants.

A New Error-Correcting Transmission Method for Dual Ring Fieldbus in CNC System

Lei Yang

School of Computer Science and Technology,
University of Science and Technology of China, Hefei, China
Email: xenosaga@mail.ustc.edu.cn

Hu Lin

Shenyang Institute of Computing Technology,
Chinese Academy of Sciences, Shenyang, China
Email: linhu@sict.ac.cn

Dongfeng Yue

School of Computer Science and Technology,
University of Science and Technology of China, Hefei, China
Email: yuedongfeng@sict.ac.cn

Tianrong Gao

Graduate School of Chinese Academy of Sciences, Beijing, China
Email: gaotianrong@sict.ac.cn

Abstract—For solving the security problems caused by the errors occurred during the real-time transmission of the fieldbus in the CNC system, a new transmission scheme, which involves the error detection and self-correction, was proposed based on the dual ring fieldbus in the CNC system. This scheme specifies a structure of two-way transmission in fieldbus and a data format with self-correcting function as well as related detection and correction method was designed. The result of testing demonstrated that, comparing to the existed method, this scheme excellently met the real-time requirements of CNC system while effectively reducing the retransmission probability and improving the efficiency.

Index Terms—error correction, dual ring fieldbus, real-time performance, retransmission probability, CNC system

I. INTRODUCTION

As the widely using of CNC system in the manufacturing field, it is evolving to be high-speed, high-precision and high reliable^[1]. The significant factors inside are the real-time performance and the secure data communication^[2]. They are also the demands for the CNC system fieldbus in which real-time performance and transmission security are essentially foundational. The real-time performance guarantees data's arrival to stations in time and data security ensures safety of system by avoiding accidents. Therefore, the improvement on these two parts is quite significant^[3].

In the past, the key point of transmission security research is focused on optimization of fieldbus structure, error detection, the enhancement of retransmission efficiency and so on. However, retransmission would result in delay in the data exchanging which would badly, if lacking an efficient self-correcting method, influent the system real-time performance.

For solving these problems, this article designs a dual ring safe transmission scheme, which has implemented self detection and correction in a single cycle time that reduces the retransmission probability on the condition that satisfied the requirement of real-time performance, while communicating between master and secondary stations.

II. RESEARCH BACKGROUND

At present the safety research of fieldbus in CNC system is concentrated on fault protection protocols, which are committed to error detecting and feedback transmitting, just like FF, PROFIBUS and CANBUS^[4] does. Those fieldbus protocols support secure communication within fault detection such as Profisafe^[5], in which data are encrypted, and CANopen safety with Safeguard Cycle Time^[6].

In Profisafe, the data to be sent will be preprocessed in the master station. Before it is formatted, the encryption algorithm picks up the information of both bit streams and current situation variables to generate secure keys. The master station sends the keys with the encrypted data to the secondary stations. While secondary stations receive the encrypted data, they first fetch the keys, and

This work is supported by the Chinese National Projects for Science and Technology Development #2011ZX04016-071 to H. Lin.

then use the keys to verify the data. If the data is transmitted accurately, they will be decrypted and accepted by the stations. But if there are errors, the decryption will fail, and then the secondary stations report to the master and request retransmission.

The CANopen protocol provides another safe-guaranteed method^[7]. In the station a timer is set to count the transmission time. It has a standard safe time in every initialization and after the data is sent, the timer launches^[8]. If the master station receives the feedback of secondary stations in time, it means that the transmission is correctly working. If the transmission is time-out or even the feedback has not been received, which means one or more stations may interrupt, the circuit will reform immediately and the data will retransmitted^[9].

In general, both the protocols focus on protecting connectivity of message exchanging among stations and reliability of feedback that transmitting to ensure whenever error occurred the master is able to keep connection with secondary stations to receive feedbacks in order to start retransmission. On the one hand, it guarantees errors could be found and then corrected, however, on the other hand, retransmitting frequently will increase the workload of master station and also waste extra communicating rounds in sending repetitious data which would lead to delays, even interruptions in the working duration^[10].

Fig. 1 shows the two-channel transmission structure. It is widely used in the industrial fieldbus^[11].

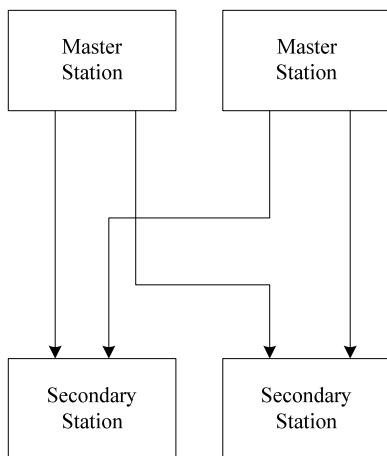


Figure 1. The two-channel transmission structure

This structure has two or more master stations working in the system. In each cycle time the same data will be sent to a station in separated circuits^[12]. One may lost during the transmission while another one will arrive to the destination successfully. It keeps double redundancies in transmission and enhances the security^[13].

Further more, a multi-fieldbus scheme with sorts of transmitting technologies is defined^[14]. To meet the demand of respective real-time performance, it differentiates high speed and low speed stations in different fieldbus and adopts kinds of safe communication methods as shown below.

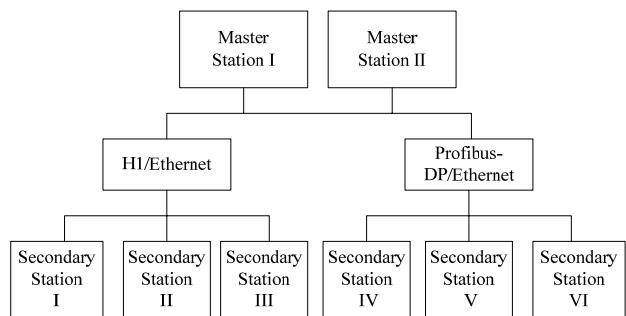


Figure 2. The structure of multi-fieldbus scheme.

The multi-fieldbus scheme separates a mount of stations by disparate fieldbus, minimizing the station number in a single ring^[15]. It shortens communication cycle time and reduces the badly influence of retransmission. Nevertheless, integrating various fieldbus protocols into a system is quite a complicated scheduling for master station, especially in which data has to be formatted specifically in each communication cycle time^[10]. Even more, when the number of stations isn't so gigantic, the multi-fieldbus scheme rarely enhances the systematic performance. For this reason, problems still exist.

III. THE DUAL RING SELF-CORRECTING TRANSMISSION SCHEME

Considering the shortages of existing schemes about data transmission in fieldbus, this article proposes a self-correcting transmission method based on dual ring structure of fieldbus. It allows two-way transmitting between master and slaves and carries out both error detection and self correction in a single cycle time without retransmission so frequently.

A. Dual Ring Topological Structure

As Fig. 3 shows, communication is carried on two rings. After the initialization in each cycle time, the master station sends a data packet through Ring 1 and another through Ring 2. The two packets are transmitted at the meantime. When a secondary station receives a packet, first it verifies which ring the packet comes from. Then it fetches the data and sends the packet to its neighbor^[16]. If there are errors detected, it waits for the packet of the other ring. When the other packet arrives, the station could fetch the data again without requesting retransmission. After two packets return to the master station, the master shall check the packets at once, if there's no request for retransmission, the current cycle ends and the next one starts. The advantage of this structure is that only one master station is necessary and the running efficiency has increased. In future, it could be expanded to multi-ring structure.

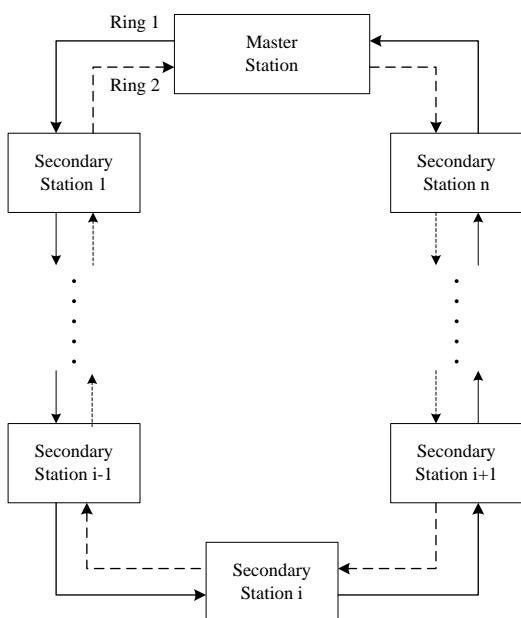


Figure 3. The topological structure of dual ring fieldbus.

The next figure shows the abstract communication structure of dual ring fieldbus in CNC system. Both master and secondary stations have two ports which form a ring connectively one by one. The master sends Ring 1's packet in port 1 and the other in port 2. If two packets achieve the same secondary station, they will be handled in the manner of FIFO^[17].

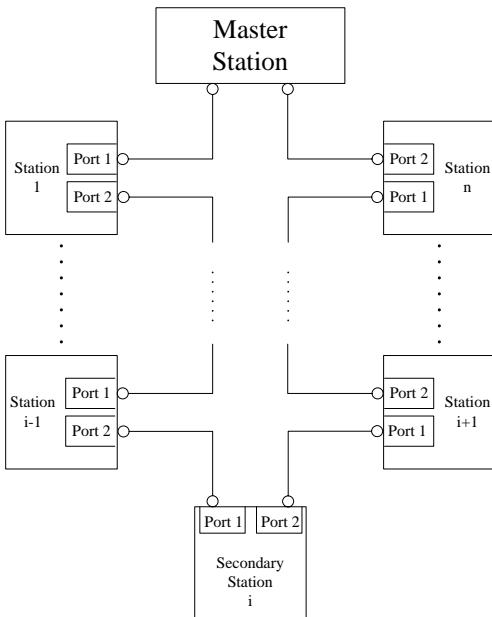


Figure 4. The communication structure of dual ring field bus.

B. Self Correction Method

In the communication, the data are not sent directly, instead they are packed to the fixed format^[18] in order to be get and set correctly.

The packet format is defined as below:

Protocol Header (12 Byte)	Ring Num (2 bit)	Packet Capacity (4 Bit)	Sequence Num (1 Byte)	Packet Payload
------------------------------	---------------------	----------------------------	--------------------------	----------------

Figure 5. The format of packets in the rings.

Protocol Header: the header information contained in an packet for the current transmitting protocol of fieldbus.

Ring Num: The distinctive sign of rings to the packets.

Packet Capacity: the number of data contained in Packet Payload.

Sequence Num: the number generated by master station every time a new cycle starts to maintain consistency with secondary stations.

Packet Payload: the field consists of data ordering by Station Num from which secondary stations would fetch the data.

As seen above, the *Protocol Header*, *Ring Num* and *Sequence Num* are the control field^[19]. They are used to be recognized by the secondary stations for which ring it comes from and which cycle it belongs to. All the information would be saved by the secondary stations to keep synchronous with the master. The *Packet Capacity* is an important part of the self-correction operation. Stations use it as the base number of the selection of related data.

The *Packet Payload* is the data field of the packet^[20]. It has the data sending to the secondary stations and the data in the *Packet Payload* also have been formatted by the master station as below:

Station Num (4 bit)	Data	CRC (2 Byte)
------------------------	------	-----------------

Figure 6. The data format in Packet Payload.

Station Num: the mark of the secondary station it's sending to.

Data Field: the content of Packet Payload.

CRC: Checking for the errors occurred in transmission duration.

The *Station Num* of the data is used to be distinguished by the station. It is also made use of the self correction as the label of each data. *CRC* checking is a part of the self detection and it is always invoked before stations fetch the instruction from the *Data Field* to judge if there are errors in the data^[21].

The self-correction method is used to generate self-correcting data by doing XOR operations with original data in the master station. Each self-correcting data is useless, but by doing the XOR operations the original data could be regained with two or three self-correcting data. After generation, these data are put into Data Field of the self-correction packet which has the same format as original packet. The original packet will be sent to Ring 1 through Port 1 while the self correction packet will be sent to Ring 2 through Port 2. The execution procedure of self-correction method is shown as blow:

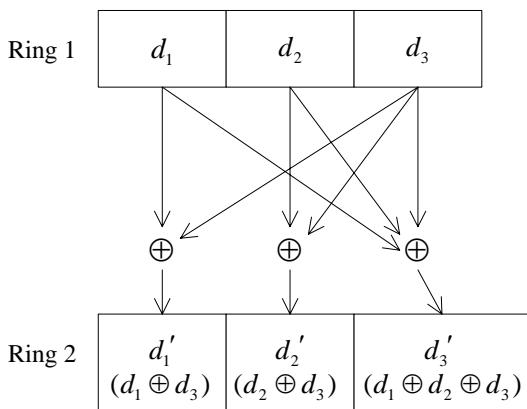


Figure 7. The generation of self-correcting data.

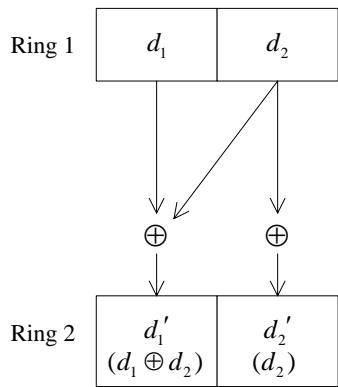


Figure 8. The generation of the last two self-correcting data.

Fig. 7 has overleaped the Sequence Num, Station Num and other unnecessary descriptions in order to pay attention to the executing procedure. In Ring 1 it's the original data in the payload and the results are put into the payloads in Ring 2. Especially if the number of secondary stations is not multiples of 3, some extra operations are needed. If only one more, directly put its original data into the payload of self correction packet without XOR operation. If two more, just do XOR once with the two original data, and then put the result and final original data into the last positions of the payload.

Finally, data in the payload of Ring 2 should be performed as (1) and (2). The variable with quote stands for the self-correction data in Ring 2 and others signify the original data in Ring 1. The subscript i means Station Num which the data belongs to. In an integrated cycle time, it does $m + m/3$ times XOR operations at most, so time complexity is low to $O(m)$ which means all the operations could be accomplished in constant time and has no bad effects on real-time performance.

$$d'_i = \begin{cases} d_i \oplus d_{i+2}, & i=3k-2; \\ d_i \oplus d_{i+1}, & i=3k-1; \\ d_{i-2} \oplus d_{i-1} \oplus d_i, & i=3k; \end{cases} \quad k \in [0, \lfloor \frac{m}{3} \rfloor] \quad (1)$$

$$d'_i = \begin{cases} d_i \oplus d_{i+1}, & \text{if } i \text{ is the last but one;} \\ d_i, & \text{if } i \text{ is the last one;} \end{cases} \quad (2)$$

A whole working process of the master station is shown as below:

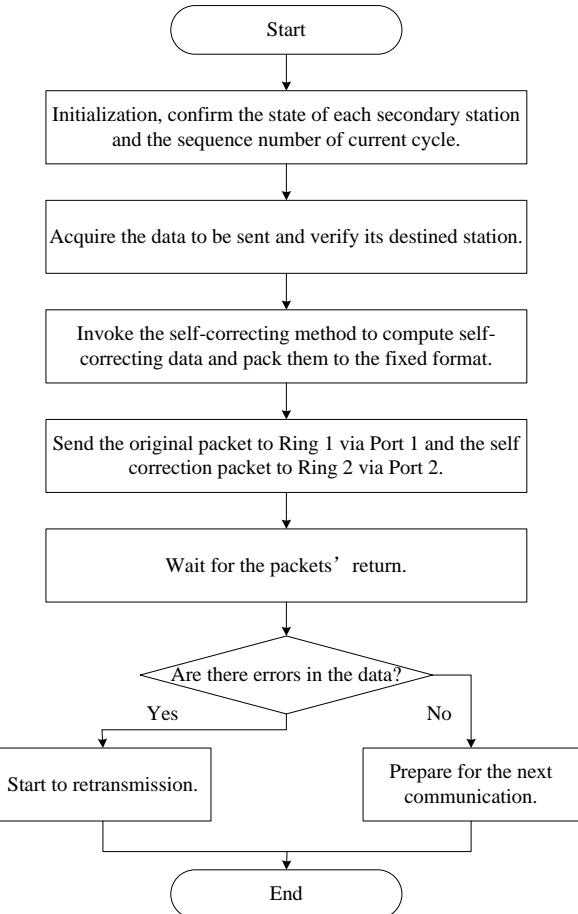


Figure 9. The flow chart of the master station

C. Self-Correcting Transmission Scheme

After both packets of two rings have been generated, master station sends the original packets to Ring 1 through Port 1 while the other to Ring 2 through Port 2. A secondary station receives the packet from its neighbor and fetches the data belonging by comparing Station Num and Port Num.

Each secondary station has a status bit used to indicate if there's error or data loss in fetched data. The status bit is set to zero initially, if the station finds that data has lost or CRC checking fails after the packet's arrival from Ring 1, it sets the bit to 1. When the self-correction packet gets to the station via Ring 2, the station checks the status bit first. If the bit is zero, which means the original data has been received correctly, and then the station sends the packet to next station directly without any operation. If the bit is 1, it means something goes wrong in the transmission of original data and needs repairs in order to send the right data to the station.

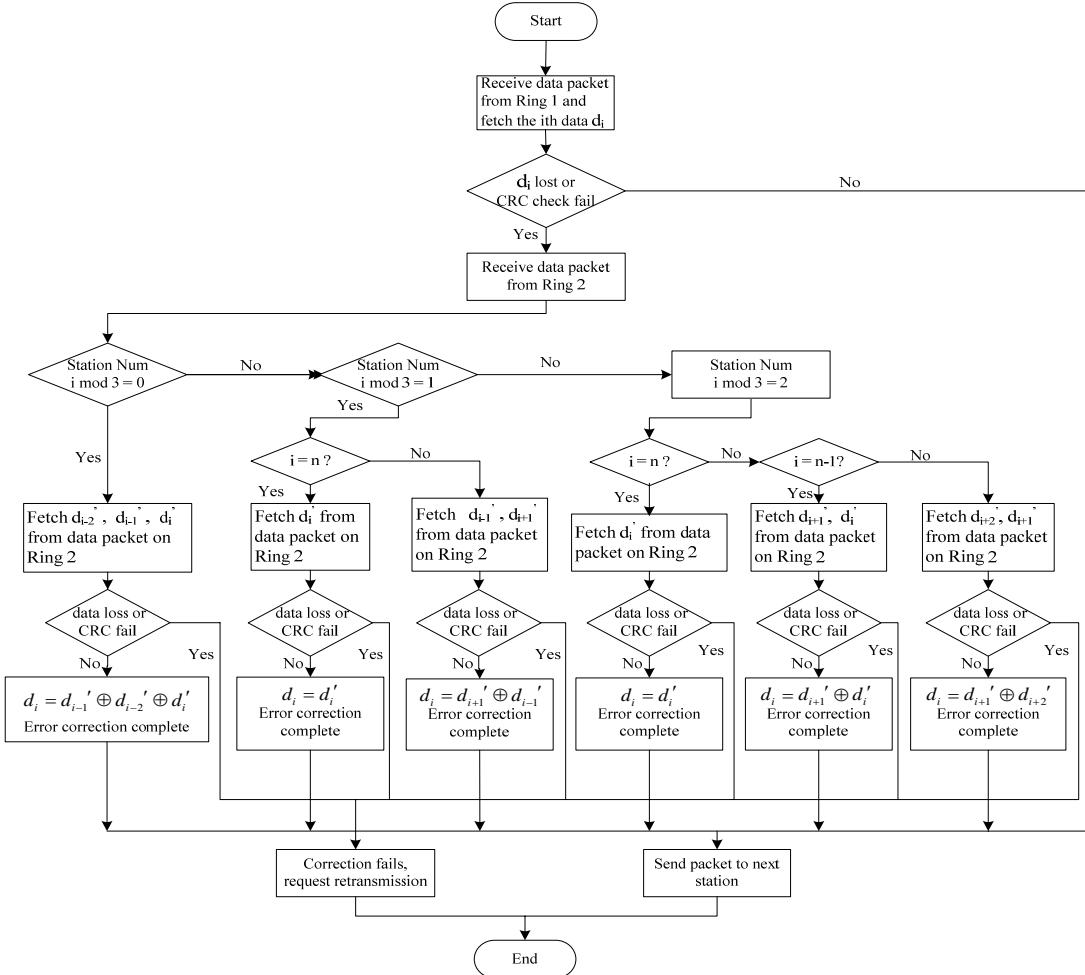


Figure10. Note how the caption is centered in the column.

At the beginning of correction, it does the modulus operations for the next step due to the remainder. Supposing SN is the Station Num of current station and R is the remainder, then $R = SN \bmod 3$.

$$d_{SN} = \begin{cases} d_{SN+1}' \oplus d_{SN+2}', & R=1; \\ d_{SN}' \oplus d_{SN+2}', & R=2; \\ d_{SN}' \oplus d_{SN+1}' \oplus d_{SN+2}', & R=0; \end{cases} \quad (3)$$

$$d_{SN} = \begin{cases} d_{SN}' \oplus d_{SN+1}', & \text{if } SN \text{ is the last but one;} \\ d_{SN}, & \text{if } SN \text{ is the last one;} \end{cases} \quad (4)$$

Equation (3) and (4) shows the specific operations. d is the original data needed to be restored and d' on the right is self-correction data fetched from the packet in Ring 2. According to do XOR operation to related data, secondary station will get original data back, then set status bit to zero again.

Specifically if self correction data has lost or CRC checking fails, correction is infeasible, secondary station must request retransmission. Fig. 10 detailed describes the self-correcting process in a secondary station which Station Num is i and Packet Capacity (the number of all the secondary station) is n .

Because of the status bit stations can acquire the state of transmission, if no errors, the packet can be passed to

next station without delay. It guarantees the real-time performance in comparison to other existed methods. Otherwise, the self correction data is separated from original data, so it is able to restore sequential data errors. Only if both the original data and self correction data go wrong, the station just requests retransmission. The retransmission probability has been brought down greatly in this new method.

IV. PERFORMANCE ANALYSIS

Experimental facilities include one master station and five secondary stations. Configuration of master station is 800 MHz Loongson CPU and 512 MB RAM. Transmission Medium is 100Mb/s fieldbus. The size of packet is 20 Bytes, cycle time is set to 1 ms.

A. Real-Time Performance Test

Real-time performance is the basic requirement, if in a transmitting process the information data could not be sent to stations in time, as a result system may terminate accidentally. For this reason, the real-time performance is one of the important terms in the transmission scheme of fieldbus. In this test, the performance in the master and secondary station are recorded separately.

Theoretically in the self-correcting transmission scheme, the time complexity is constantly $O(m)$ in dual-ring transmission without additionally delay. In the test, the delaying time in every station and the cycle time, of

both two channel transmission scheme without self correction and self-correcting transmission scheme, will be recorded. By comparing the proportion that delaying time to the cycle time in these two schemes, the influence of self correction method to the real-time performance will be shown.

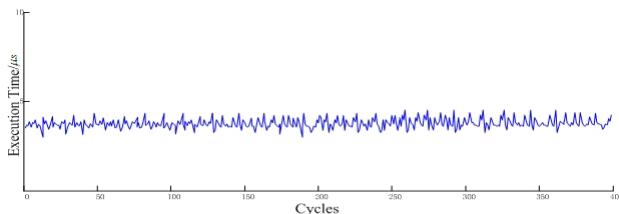


Figure 11. The delay of secondary stations in two-channel transmission.

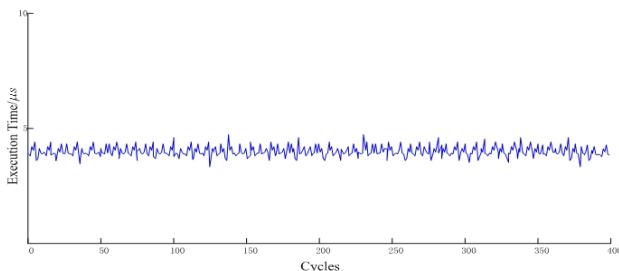


Figure 12. The delay of secondary stations in dual ring self-correcting transmission.

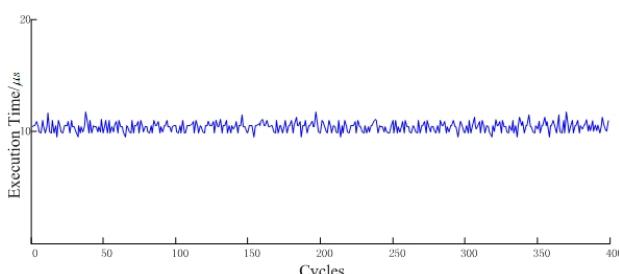


Figure 13. The delay of master station in two-channel transmission.

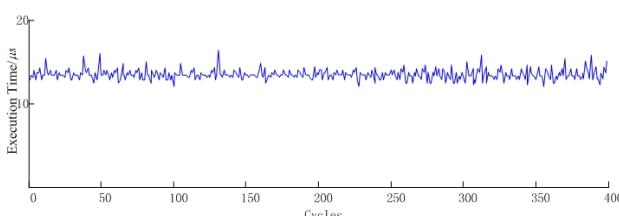


Figure 14. The delay of master station in dual ring self-correcting transmission.

The test shows that the self correction scheme is not so far different from two channels transmission scheme on the average executing time in stations. In master station, it takes some time to do the exclusive OR operations. The delay in each secondary station is less than 2%, so it can well meet the real-time requirement.

TABLE I.
THE AVERAGE DELAY OF STATIONS AND CYCLES

Average delay(μs)	Secondary	Master	Cycle time
Two channel	3.6645	10.6046	182.5637
Dual-ring	3.8422	13.8542	203.9815

B. Recovery Capability Test

While testing the restoring efficiency of self-correction method, master station will send packets within determinate loss rate. After transmissions accomplish, the restored loss rate will be recorded. By comparing loss rates in different loss rates and cycle numbers before and after, the validity of self correction method will be shown.

TABLE II.
THE REPAIR RATES OF SELF-CORRECTING METHOD

Test Sequence	Cycles	Loss rate Determinate (%)	Loss rate After correction (%)
1	10	30	<5
2	10	20	<3
3	20	30	<2
4	20	20	<1

As can be seen from the table above, the effect of self correction method shall enhance as increasing of cycles and the method is able to restore most data losses effectively in different loss rates. The recovery capability of the method is admirably testified.

C. Consequence

From the tests, the self-correcting method in dual ring fieldbus is proved to be able to meet the requirement in the efficiency of transmission, but also the demand in the self correction.

Comparing to the two-channel transmission, the self-correcting method has no significant differences in the execution time of the secondary stations. But it spends much more time on master station. That's because of the additional XOR operations and accompanying data search and access. Each XOR operation does at least two data search. If the data are not sequential, even more time will be wasted on finding these related data.

The self correction method is able to restore most errors on the condition that the self-correcting data are unbroken. If that, the correction will be completed. But if the original data and the self-correcting data are both lost, the correction isn't able to continue which is the shortage should be covered.

V. CONCLUSION

In this article a self-correcting transmission scheme is proposed based on dual ring fieldbus in CNC system to meet the requirements of both real-time performance and

security communication. In the scheme the dual ring structure is used to transmit information in the fieldbus, a self correction method is designed to generate correction data by processing the original data. The self-correcting data are able to restore the errors independently. These two data are packed to the fixed format with control field and it can be accessed by the secondary stations. The packets are sent in two rings through different ports. Both error detection and correction can be accomplished in a single cycle time. The tests prove that, with tolerable increasing on the master stations' workload, the scheme meets the real-time demand of CNC system. It decreases the retransmission probability effectively as well. The follow-up will be focused on the research of safety communication protocols and error detection and correction method to meet the security requirement of fieldbus in Numerical Control system.

ACKNOWLEDGMENT

We would like to thank Shenyang Golding NC Tech. Co., Ltd. Funding for this research was provided by the Chinese National Projects for Science and Technology Development 2011ZX04016-071.

REFERENCES

- [1] MIGUEL G. Safety over EtherCAT: the safety solution for EtherCAT[EB/OL].[2010-01-22]. http://www.ethernetcat.org/pdf/english/pcc0107_safety_over_ethercat_e.pdf.
- [2] JARMO A, MARITA H, TIMO M. "Safety of digital communications in machines" [EB/OL]. [2010-01-22]. <http://www.vtt.fi/inf/pdf/tiedotteet/2004/T2265.pdf>.
- [3] WU Qin, QIN Qiang. "Analysis of implementation of the safety function in CC-Link safety net work". *Automation Panorama*, 2007, 24(2): 36-38.
- [4] REXROTH B, HUMPHREY D W. "Safe motion technologies: competitive advantages for early adopters" [EB/OL]. [2010-01-20]. http://www.plantservices.com/Media/MediaManager/rex_r0th_safe_motion.pdf.
- [5] WU Chengyan. "The research and analysis of safety bus protocol-profisafe". Beijing: Beijing Jiaotong University, 2006.
- [6] RAHMANI M, HINTERMAIER W, RATHGEBER B M. "Error detection capabilities of automotive network technologies and Ethernet-a comparative study". *IEEE Intelligent Vehicles Symposium*, 2007, 13(15): 674-679.
- [7] DZUNG D, NAEDELE M, HOFF T P V, et al. "Security for industrial communication systems". *Proceedings of the IEEE*, 2005, 93(6): 1152-1177.
- [8] Dempsey B J, Weaver A C. "On retransmission-based error control for continuous media traffic in packet-switching network". *Computer Networks and ISDN Systems*, 1996, 28(5): 719~736.
- [9] ITU-T Recommendation H. 323. "Packet based multimedia communications systems". 1998. http://www.databeam.com/Standards/H323/H323_Primer.html.
- [10] C. Schwaiger and A. Treytl, "Smart Card Based Security for Fieldbus Systems", *2003 IEEE Conference on Emerging Technologies and Factory Automation*, Lisbon, 16-19. Sep. 2003, pp. 398-406, 2003.
- [11] Ch. Schwaiger and T. Sauter, "Security Strategies for Field Area Networks", *28th Annual Conference of the IEEE Industrial Electronics Society (IECON)*, Sevilla, 5.-8. Nov.2002, pp.2915-2920.
- [12] T.Sauter and Ch. Schwaiger, "Achievement of secure Internet access to fieldbus systems", *Microprocessors and Microsystems*, 26(2002), pp. 331-339.
- [13] Boterenbrood H. CANopen: High Level Protocol for CAN-bus[Z]. 2000
- [14] CAN in Automation, CANopen Specification.
- [15] R.Bosch GmbH, Controller Area Network Specification: Version2. 1991
- [16] CAN in Automation, CiA DSP 304-V1.0.1: Framework for safety-relevant communication, 2004
- [17] Jungandreas, F., "The CANopen Safety Chip", *CAN Newsletter*, 4/2003.
- [18] Liu Cuiqing, Ping Xijian and Zhang Tao, "A Research on Steganography Method Based on Error-correcting codes", *Proceeding of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IHI-MSP'06)*, 2006.
- [19] Wang Weixiang, Liu Yujun and Li Wenxiang, "Information Hiding Algorithm in Channel Coding Based on m-Sequence", *Computer Engineering*, 2007,(3): 118-122.
- [20] R. Ahlswede, B. Balkenhol, and N. Cai, "Parallel Error Correction Codes", *IEEE Trans. Inf. Theory*, vol. 48, no.4, pp.959-962, Apr. 2002.
- [21] F. J. McWilliams and N. J. A. Sloane, "The Theory of Error Correcting Codes", Amsterdam, The Netherlands: North-Holland, 1986.

Lei Yang received his Bachelor degree in Computer Science from School of Computer Science and Technology, University of Science and Technology of China, Hefei, China, in 2009. He is currently a Ph.D. candidate at the School of Computer Science and Technology at University of Science and Technology of China. His current research interest is in the area of CNC technology, safety fieldbus and security control system.

Hu Lin(Ph.D.) is a professor in Chinese Academy of Sciences and director of the doctoral school of University of Science and Technology of China. His research interests include CNC technology and real-time system.

Dongfeng Yue received his Bachelor degree in Computer Science from School of Computer Science and Technology, University of Science and Technology of China, Hefei, China, in 2007. He is currently a Ph.D. candidate at the School of Computer Science and Technology at University of Science and Technology of China. His current research interest is in the area of CNC technology, safety fieldbus and security control system.

Tianrong Gao received her M.E. degree in Graduate School of Chinese Academy of Sciences, Beijing, China, in 2010. She is currently a Ph.D. candidate at Graduate School of Chinese Academy of Sciences. Her current research interest is in the area of CNC technology, fault diagnosis and networked CNC system.

Numerical Simulation of Heavy Rail Quenching Process

Gongfa Li

College of Machinery and Automation, Wuhan University of Science and Technology, Wuhan, China
 Guangxi Key Laboratory of Manufacturing System & Advanced Manufacturing Technology, Guangxi, China
 Wuhan University of Science and Technology City College, Wuhan, China
 Email:ligongfa@yahoo.com.cn

Yuesheng Gu, Jianyi Kong, Guozhang Jiang , Jintang Yang, Liangxi Xie , Zehao Wu , Siqiang Xu ,
 Yaping Zhao and Gang Zhao

Department of Computer Science, Henan Institute of Science and Technology, Xinxiang 453003, Henan, China
 College of Machinery and Automation, Wuhan University of Science and Technology, Wuhan, China

Email: hz34567@126.com, jykong_1961@yahoo.com.cn, whjgz@wust.edu.cn, yjintang163@163.com,
 lx_tse@163.com, zehaowu@163.com, xusiqiang@yahoo.com.cn, zhaoyaping1975@126.com, snowcampus@163.com

Abstract—Heat-treatment is emphasized for its important role to qualify the heavy rail. The temperature and stress field of U71Mn heavy rail during the quenching process were simulated by the FEM software and various kinds of factors that may influence temperature and stress field distribution were investigated. The result shows that the main factors influencing the temperature and stress field distribution were the heating and holding time, as well as the pressure of air blast during the cooling, whose values would directly influence the final performance of the heavy rail after the heat treatment. It is significantly valuable for the choice of relevant parameters for the heat treatment of U71Mn heavy rail.

Index Terms—Heavy Rail, Quenching, Temperature Field, Stress Field, Numerical Simulation

I. INTRODUCTION

With continuous increasing of speed of trains on the mainline railway, impact that the train's hub exerts on the rail-end was larger and larger and the impact frequency is higher and higher. U71Mn heavy rails are discarded since phenomena, such as spalling and crack, occur on their ends after trains move on them for a period time. Frequently changing rails affects train schedules and diminishes the railway efficiency. Moreover, subtle crack is a potential hazard to safe movement for trains. The fact that spalling and crack appear on rails results mainly from that the thermal stress and structural stress (together called the residual stress) generated during the rail-end's quenching process were not entirely eliminated and then

superimposition of rails' residual stresses and impact stresses, which continuous hitting of hubs against rails introduces, exceeds the rail strength. Therefore, performances which are required after the rail's heat treatment according to train safety are enough fatigue resistance, as well as better capability of resistance to compression, abrasion and atmospheric corrosion. In order to produce the rail satisfying these performance requirements, all steel manufacturers need to improve the existing heat-treatment technology. The temperature and stress field of U71Mn heavy rail (hereafter referred to as heavy rail) widely used in the mainline railway during the heat treatment process were calculated in numerical simulation and main factors that influence the temperature and stress field to be distributed were identified. A guiding and referencing significance was provided for steel manufacturers to improve the existing heavy rail heat-treatment technology.

II. MATHEMATICAL MODEL

The temperature and stress distribution were closely related to heat conduction during heavy rail's quenching. In order to research and simulate the heavy rail's quenching process, it was essential to establish a mathematical model of heavy rail's heat conduction and determine the constitutive relation of the conducting process. In addition, establishment of the constitutive relation needed determine definite conditions, such as quenching workpiece's geometric conditions, thermal physical parameters for the heavy rail, initial temperature distribution (initial conditions) of the quenching workpiece and medium, and heat transfer (boundary conditions) between the quenching workpiece outer surface and the quenching medium.

The mathematical model of heat conduction principally included the controlling equation of heat conduction and definite conditions of prescribed problems.

This research reported in the paper is supported by Guangxi Key Laboratory of Manufacturing System & Advanced Manufacturing Technology, Research Center of Green manufacturing and Energy-Saving &Emission Reduction Technology in Wuhan University of Science and Technology (B1004) and National Natural Science Foundation of China (70971102).

Corresponding author: ligongfa@yahoo.com.cn.

The controlling equation of heat conduction was obtained on the basis of the Energy Conservation Law and the Fourier Law. The heat conduction equation for the heavy rail was the Fourier's heat-conduction differential equation which took body temperature varying with time (non-steady problems) into consideration when there was an inner thermal source. It was obtained by applying the Energy Conservation Principle according to the Fourier Law, as shown concretely in equation (1):

$$\lambda \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \right) + q' = \rho c_p \frac{\partial T}{\partial t} \quad (1)$$

Where λ is heat conduction coefficient of the heavy rail; ρ is density of the heavy rail; T is instantaneous temperature of the heavy rail; x , y and z are the three-dimensional coordinate; q' is heat flux density of the inner heat source in the heavy rail, which is derived from latent heat during quenching process of the heavy rail; t is the time variable; c_p is specific heat capacity of constant pressure;

q' is defined as follow:

$$q' = \Delta H \frac{\partial V}{\partial t} \quad (2)$$

Where V is the volume; ΔH is the latent heat per volume.

Relevant boundary conditions and initial conditions were respectively:

$$\begin{aligned} -\lambda \frac{\partial T}{\partial n} \Big|_s &= H_k (T_\omega - T_c) + \sigma \varepsilon (T_\omega^4 - T_c^4) \\ &= H_k (T_\omega - T_c) + H_s (T_\omega - T_c) \\ &= H (T_\omega - T_c) \end{aligned} \quad (3)$$

$$T \Big|_{t=0} = T_0 (x, y, z) \quad (4)$$

Where H is total coefficient of heat transfer:

$$H = H_k + H_s$$

H_s is radiation heat transfer coefficient:

$$H_s = \sigma \varepsilon (T_\omega^2 + T_c^2) (T_\omega + T_c)$$

σ is Stefan-Boltzmann constant parameter, $\sigma = 5.678 \times 10^{-8} W / (m^2 \cdot K^4)$; ε is radiation ratio of heavy rail surface. $T_0 (x, y, z)$ is described as temperature function, where the temperature value should be denoted with absolute temperature.

III. TEMPERATURE AND STRESS DISTRIBUTION IN QUENCHING PROCESS

A. Technology of Rail-end Quenching

The chemical composition of materials used for U71Mn heavy rails was shown in Table 1. The heavy rail's heat-

treatment technology involved as follow: firstly the section 200mm apart from the rail-end was heated to 910°C for 40s in the electromagnetic field, then held for 5s, and cooled for 25s by intensive convection of cold air, at last air cooled to room temperature. The wind-cooling apparatus was shown in Fig.1.

TABLE 1.

CHEMICAL COMPOSITION OF MATERIALS USED FOR U71MN HEAVY RAILS (WT%)

Brand name	Chemical composition %				
	C	Si	Mn	P	S
U71Mn	0.65~0.7 7	0.15~0.3 5	1.10~1.5 0	<0.04 0	<0.04 0

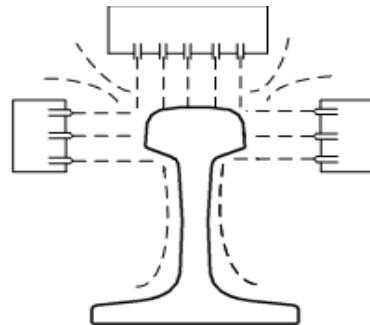


Figure 1. Sketch of the wind-cooling apparatus

B Establishment and Solution of Model

The heavy-rail's 3D model, which was plotted according to the size regulated by YB (T) 68-1987 standards for the heavy rail of 60kg/m with the three-dimensional software, was loaded into the finite element analysis software to be swept with the element SOLID5 to obtain the FEM shown in Fig.3, as presented in Fig.2.

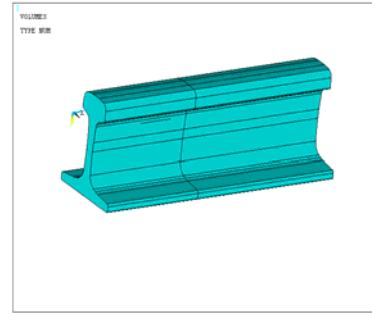


Figure 2. Model of heavy rail



Figure 3. FEM of heavy rail

Referring to the Materials Handbook, the rail-end density was 7920kg/m^3 , besides, the rest of main physical properties were shown in Table 2. Heavy-rail's thermal physical parameters, such as relative permeability, specific heat, resistivity and thermal conductivity, varied with temperature from the Table 2. They were applied to load in tabular method in application.

TABLE 2.

PARAMETERS OF MATERIALS USED FOR RAIL-ENDS

T[°C]	25	100	200	300	400	500
Relative Permeability	200	194.5	187.6	181	169.8	157.3
Specific heat J/[g · °C]	472	480	498	524	560	615
Resistivity [Ω]	1.84e -007	2.54e -007	3.39e -007	4.35e -007	5.41e -007	6.56e -007
Enthalpy[J/m ³]	9.16e +007	3.56e +008	7.5314e +008	1.16e +009	1.63e +009	2.12e +009
Thermal conductivity [W/(m · °C)]	93.23	87.68	83.53	80.44	78.13	76.02
T[°C]	600	700	800	900	1000	1100
Relative Permeability	140.8	100.3	6	1	1	1
Specific heat J/[g · °C]	700	1000	806	637	602	580
Resistivity [Ω]	7.9e -007	9.49e -007	1.08e -006	1.16e -006	1.20e -006	1.23e -006
Enthalpy[J/m ³]	2.65e +009	3.19e +009	3.72e +009	4.22e +009	4.52e +009	5.14e +009
Thermal conductivity [W/(m · °C)]	74.16	71.98	68.66	66.49	65.92	64.02

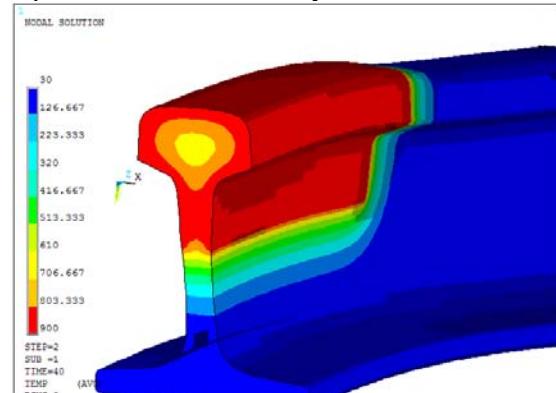
The direction of magnetic induction lines was guaranteed parallel to the rail-end boundary by setting boundary conditions for the electromagnetic field; distribution for induction current of the magnetic field was calculated by loading into coils alternating current, whose magnitude, frequency and time were $1.12\text{e}6(\text{A}/\text{m}^2)$, 1000Hz and 40s, respectively. Then the magnetic field was loaded into the temperature field as an initial condition by the sequential coupling approach. Meanwhile, initial temperature was set to 25 °C and temperature distribution was attained at this load. At last the temperature histories were loaded into the FEM to solve for stresses as initial conditions. Stress distribution in the rail-end was also obtained.

C Analysis of Factors Influencing Distribution for Temperature Field in the Heavy-rail

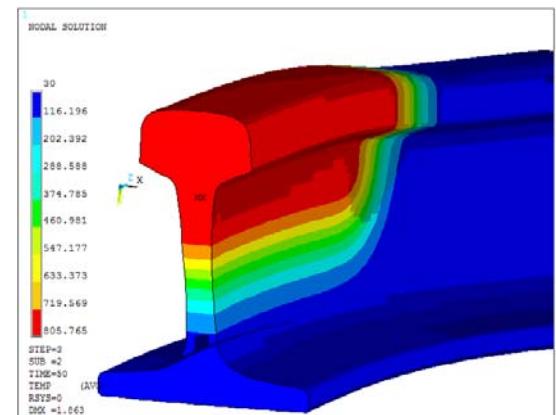
Through researching and speculating the heavy rail's quenching process, factors influencing distribution of heavy rail's temperature field were mainly the heating time, holding time and pressure of air blast during wind cooling.

(a) When heavy rail was heated to 700 °C , it was cooled by air blast at the pressure of 0.4Mpa and 0.8Mpa,

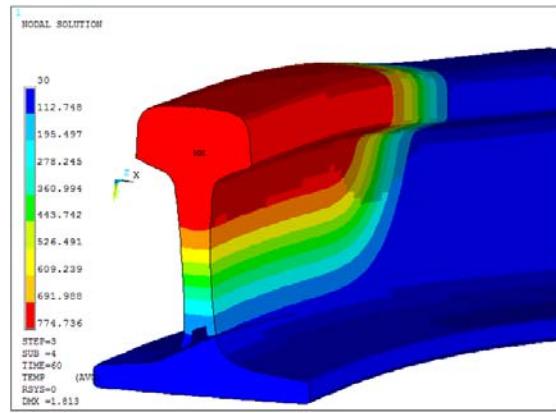
respectively. Temperature fields of the heavy rail at 40s, 50s and 60s were shown in Figure 4 and 5 respectively, from which it was known that cooling capability of air blast at the pressure of 0.4Mpa was better than that at the pressure of 0.8Mpa. The main reason why the former temperature declined more quickly on condition of the same heating, holding and cooling time was that heat convection between air at the pressure of 0.4Mpa and the surface of heavy rail was more.



(1)t=40S

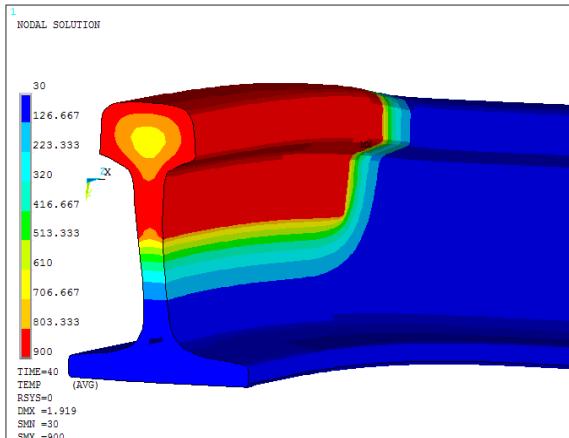


(2)t=50S

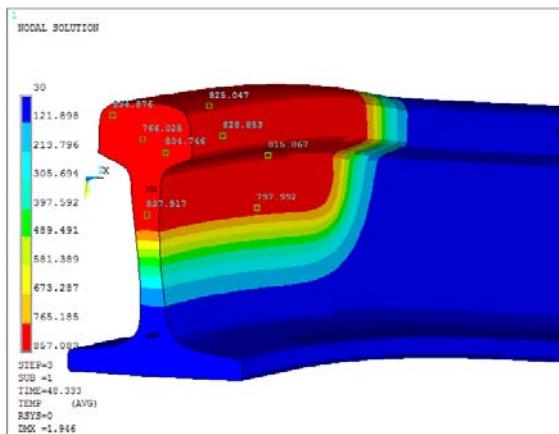


(3)t=60S

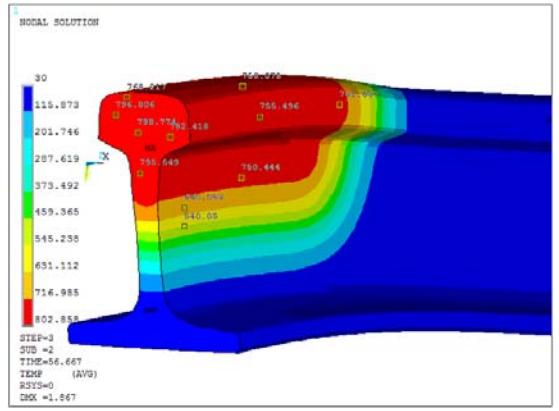
Figure4. Temperature field of heavy rail at different time when p=0.4Mpa (when heating 30s)



(1)t=40S



(2)t=50S

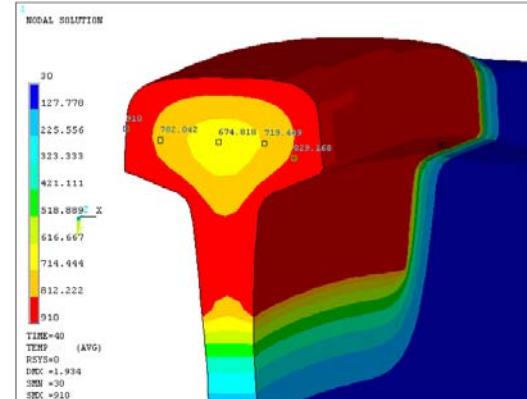


(3)t=60S

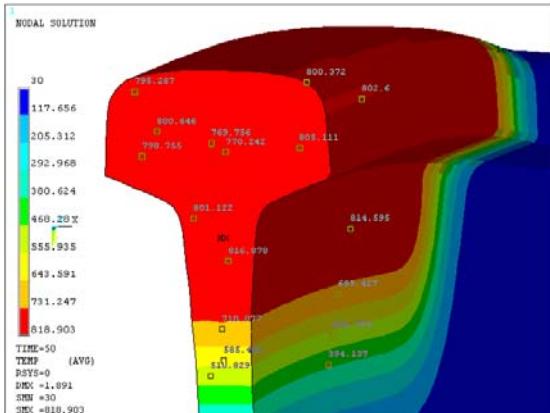
Fig. 5. Temperature field of heavy rail at different time when $p=0.8\text{Mpa}$ (when heating 30s)

(b) Figure 4 showed heavy rail's temperature field at heating time of 30s, holding time of 5s and cooling time of 25s at the pressure of 0.4Mpa. Figure 6 showed the heavy rail's temperature field at heating time of 35s, holding time of 5s and cooling time of 25s at the pressure of 0.4Mpa. It was known through comparing two temperature fields corresponding to different heating time that inner temperature of heavy rail at heating time of 35s (shown in Figure 6(1)) was higher than that at heating time of 30s. The Ministry of Railways strictly ruled that there was pearlite at least on the surface 20mm thick after

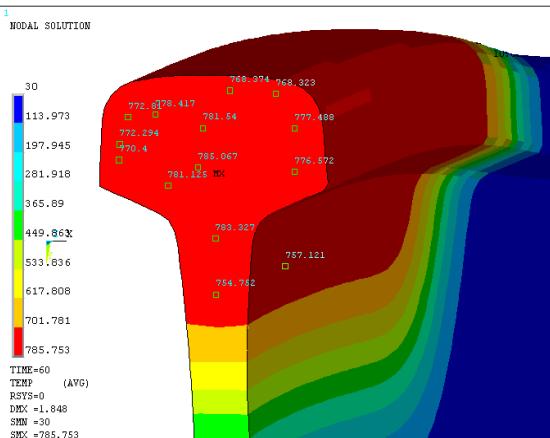
heavy rail was quenched. It meant that temperature of the heavy rail's surface 20mm thick was ensured to be heated to about 800°C . So quality of the temperature field at heating time of 35s was better than that at heating time of 30s.



(1)t=40S



(2)t=50S



(3)t=60S

Fig. 6. Temperature field of heavy rail at different time when $p=0.4\text{Mpa}$ (when heating 35s)

D Analysis of Factors Influencing Distribution for Thermal Stress Field in the Heavy-rail l

Factors influencing distributions of thermal stress fields in the heavy-rails were chiefly heating, holding time and pressure of air blast during wind cooling

through researching and discovering the heavy-rail's quenching process.

(a) Stress distributions in the rail-ends heated for 30s, heated for 30s and then held for 5s, heated for 30s, then held for 5s and at last cooled for 25s by air blast were shown in 7(1), 7(2) and 7(3), respectively; stress distributions in the rail-ends heated for 35s, heated for 35s and then held for 10s, heated for 35s, then held for 10s and at last cooled for 25s by air blast were shown in 8(1), 8(2) and 8(3), respectively.

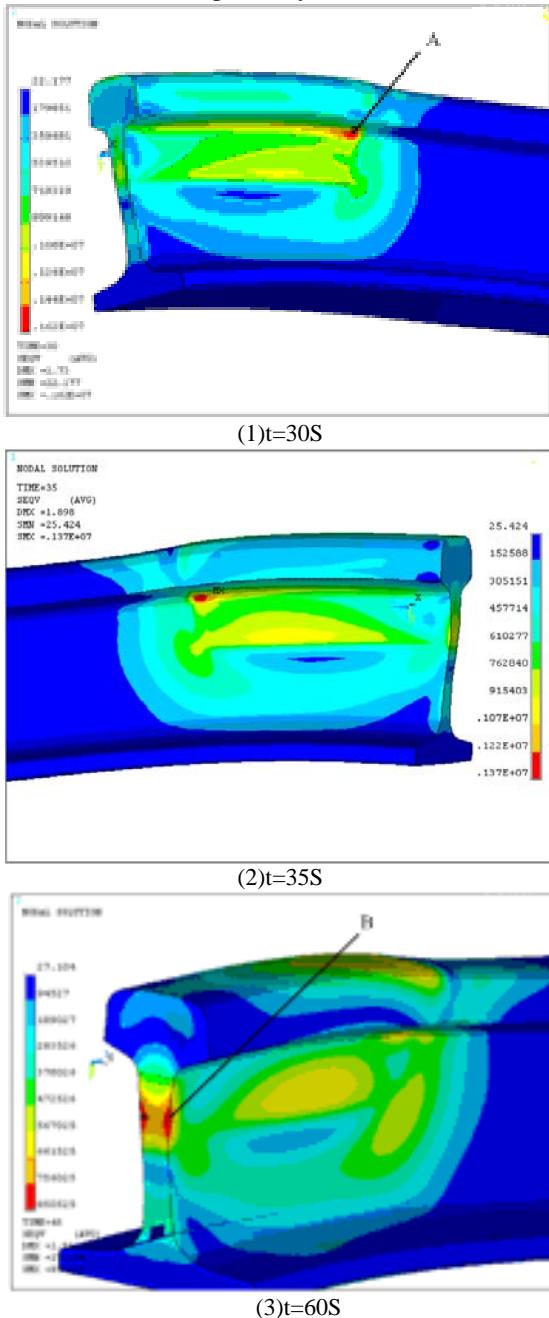


Fig. 7. Stress distributions of heavy rails at different time when $p=0.4\text{Mpa}$ (heated for 30s)

It was found through analyzing stress distribution contours that locations of the maximum stress during heating in two heating methods were at A zone which was shown in Figure 7(1) and 8(1), respectively. And

maximum stress values were 1.62Mpa and 1.35Mpa, respectively. It occurred because A zone was located at the transition region between the heated region and unheated region. A zone was compressed and subjected to compressive stress since volume changes in two regions were nonuniform.

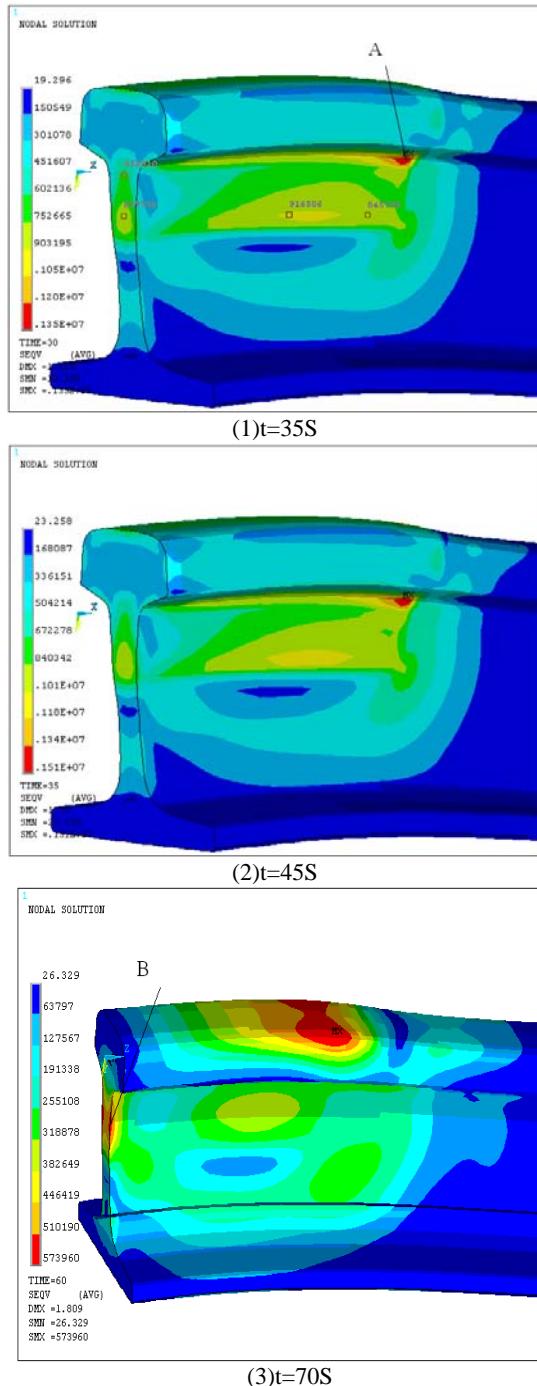


Fig. 8. Stress distributions of heavy rails at different time when $p=0.4\text{Mpa}$ (heated for 35s).

In the meantime, it was also found that locations of the maximum stress at the end of wind cooling were at B zone which was shown in Figure 7(3) and 8(3), respectively. And maximum stress values were 0.8Mpa and 0.5Mpa, respectively. The main reason was that the radiating area at the position of the B zone was less due

to the cross section vertical to the rail-end's longitudinal section during wind cooling. The thermal stress at the location was the largest.

It was discovered with numerous comparative analyses that heating time of rail-ends could not be less than 30s but redundant heating time would decrease production efficiency of heavy-rails. Heating time could be set to 40 \pm 1s after both them were taken into consideration comprehensively.

(b) Stress field distributions in the rail-ends cooled by air blast for 40s, 50s and 60s at the pressure of 0.4Mpa were shown as Fig.9(1), Fig.9(2) and Fig.9(3), respectively; stress field distributions in the rail-ends cooled by air blast for 40s, 50s and 60s at the pressure of 0.8Mpa were shown as Fig.10(1), Fig.10(2) and Fig.10(3), respectively.

Analysis of Figure 8 and 9 performed, it was obtained that concentration regions of the stress, whose values were 1.28Mpa uniformly, were at A zone shown in Figure 9(1) and 10(1) after 40s during wind cooling; concentration regions of the residual stress, whose values were 0.59Mpa and 0.56Mpa, respectively, was at B zone shown in Figure 9(3) and 10(3) at the end of cooling.

Comparison between the residual stress value in Figure 9(3) and that in Figure 10(3) was achieved. It indicated that residual thermal stress of rail-ends was less when they were cooled by compressed air at the pressure of 0.8Mpa. This was because compressed air at the pressure of 0.8Mpa possessed the better behavior for convective heat-transfer and could diminish temperature more rapidly in the heated regions. As a result, it shortened time of compressing B zone and decreased thermal stress.

From the above analysis, pressure of compressed air was set at 0.8Mpa in practical manufacturing in order to diminish residual stress and consider downsizing cost simultaneously.

V. CONCLUSIONS

The common action between thermal and structural stress during the rail-end quenching made distribution of internal stress in the workpiece extraordinarily complex. In generally, stress and deformation in the rail-end during heat treatment were assessed indirectly through gauging final stress and deformation after heat treatment. It was of lagging and apparently had a bad effect on production for heavy-rails. In this paper, the computer model of U71Mn heavy rail's temperature and stress field during quenching process was established by using the three-dimensional nonlinear FEM, and main factors which affected the temperature and stress field were compared and analyzed with various loads. Based on considering cost, the method of the residual stress relief, that the air blast pressure was set at 0.8Mpa, and then heating and holding time of the rail-end was set to 40 \pm 1s and 5s, respectively, and at last it was air cooled for 1850s after it was cooled for 40s by the air blast, was proposed. It could served as a trial tool of simulating practical quenching of heavy rail to properly select heavy rail's quenching time and air blast pressure, and conducted some primary tries for improvement of quenching technology of the heavy rail.

ACKNOWLEDGMENT

This research reported in the paper is supported by Guangxi Key Laboratory of Manufacturing System & Advanced Manufacturing Technology ,Research Center of Green manufacturing and Energy-Saving &Emission Reduction Technology in Wuhan University of Science and Technology (B1004) and National Natural Science Foundation of China (70971102). This support is greatly acknowledged.

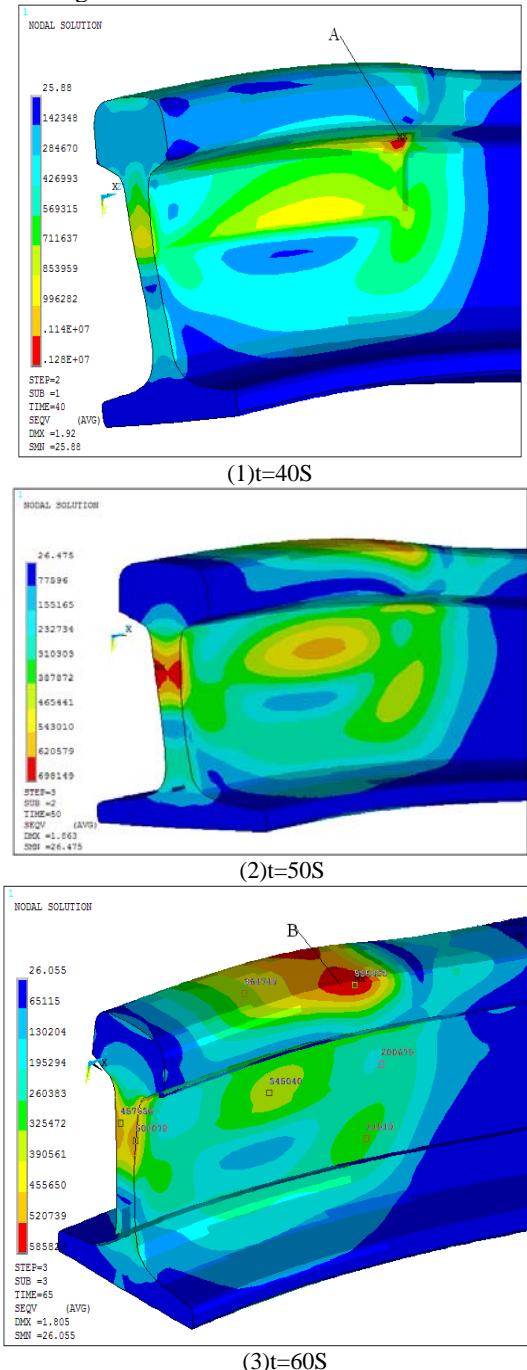


Figure 9. Stress distributions of heavy rails at different time when $p=0.4\text{Mpa}$ (when heating 30s)

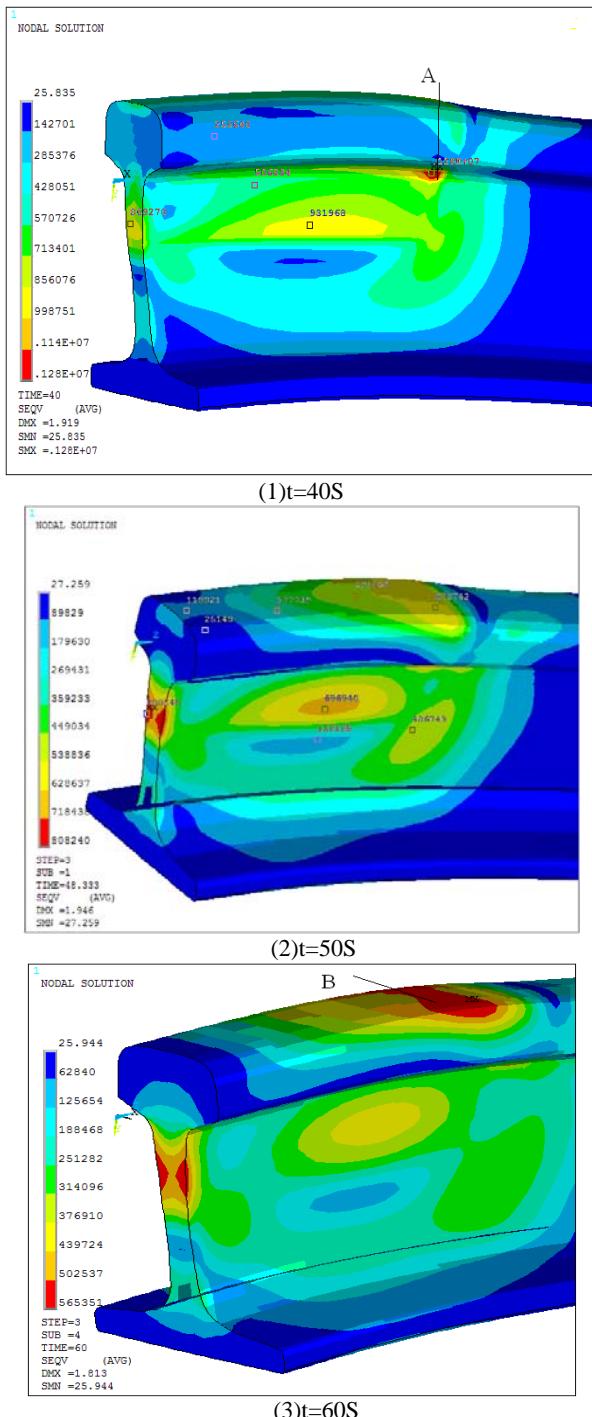
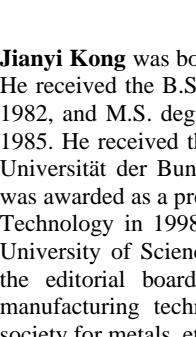
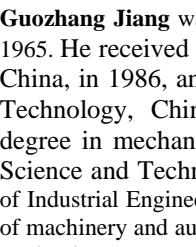


Figure 10. Stress distributions of heavy rails at different time when $p=0.8\text{Mpa}$ (when heating 30s).

REFERENCES

- [1] Jianyi Kong,Siqiang Xu,Gongfa Li,Jintang Yang,Hegen Xiong,Guozhang Jiang, "Temperature Field and Its Influencing Factors of Heavy Rail During Quenching," *Advanced Materials Research*, 2011,156-157, pp. 529–551.
 - [2] Gongfa Li, Jianyi Kong,Guozhang Jiang and Jintang Yang, "Temperature Field and Its Influencing Factors of U71Mn Heavy Rail during Quenching," *Modern Manufacturing Engineering*, 2010,(11), pp. 65–69, in Chinese.
 - [3] Gongfa Li, Jianyi Kong,Guozhang Jiang and Jintang Yang, "Stress Field and Its Influencing Factors of U71Mn Heavy Rail during Quenching," *Modern Manufacturing Engineering*, 2010,(4), pp. 91–94, in Chinese.
 - [4] Hua Long,Jintang Yang and Gongfa Li, "Simulation on Temperature Field of Heavy Rail During Quenching," *Material and Heat Treatment* 2009,37(4), pp.77–79, in Chinese.
 - [5] Hua Long,Jintang Yang and Gongfa Li., "Study on Temperature Field of Heavy Rail during Quenching Based on ANSYS," *Machinery Design and Manufacture*, 2009,(4), pp. 132–134, in Chinese.
- 

Gongfa Li was born in Hubei province, P. R. China, in 1979. He received B.S., M.S. and Ph.D. degrees in mechanical design and theory from Wuhan University of Science and Technology, China, in 2001, 2004 and 2008, respectively. He is an Associate Professor in the College of Machinery and Automation of Wuhan University of Science and Technology. His current research interests include intelligent control, computer aided engineering and optimization design.
- 

Jianyi Kong was born in Jiangxi province, P. R. China, in 1961. He received the B.S. degree in Nanchang University, China, in 1982, and M.S. degree in Xi'an Jiaotong University, China, in 1985. He received the Ph.D. degree in mechanical design from Universität der Bundeswehr Hamburg, Germany, in 1995. He was awarded as a professor of Wuhan University of Science and Technology in 1998. Currently, he is the president of Wuhan University of Science and Technology, China. He services on the editorial boards of the Chinese journal of equipment manufacturing technology. He is a director of the Chinese society for metals, etc. His research interests focus on intelligent machine and controlled mechanism, dynamic design and fault diagnosis in electromechanical systems, mechanical CAD/CAE, intelligent design and control, etc.
- 

Guozhang Jiang was born in Hubei province, P. R. China, in 1965. He received the B.S. degree in Chang'an University, China, in 1986, and M.S. degree in Wuhan University of Technology, China, in 1992. He received the Ph.D. degree in mechanical design from Wuhan University of Science and Technology, China, in 2007. He is a Professor of Industrial Engineering, and the Assistant Dean of the college of machinery and automation, Wuhan University of Science and Technology. Currently, his research interests are computer aided engineering, mechanical CAD/CAE and industrial engineering and management system.

A Redundant FPGA Based Controller for Subsea Blowout Preventer Stack

Zengkai Liu

College of Mechanical and Electronic, China University of Petroleum, Qingdao, China
Email: liuzengk@163.com

Yonghong Liu*, Baoping Cai, Fei Wang, Zhili Chen, Xiaojie Tian, and Yazhou Wang
College of Mechanical and Electronic, China University of Petroleum, Qingdao, China
Email: liuyhupc@126.com, caibaoping987@163.com, wangfei208@sina.com,
Email: chenzhili_218@163.com, tianxj20050101@163.com, sanguoyazhou@163.com

Abstract—A redundant Field Programmable Gate Array (FPGA) based controller for subsea Blowout Preventer (BOP) is presented. Triple modular redundancy technique is used for architecture design, since high reliability is a necessary requirement for subsea BOP control system. A multiprocessor system is developed to enhance reliability of the processors and performance of the system. In addition, the shared memory method is applied to interchange information between the processors. One processor is responsible for communication, while the others run application programs. The proposed system has been implemented using three FPGA development boards, which are connected to each other through RS-232 serial ports. Besides, the voting algorithms for discrete input, analog input and discrete output are proposed. Functional simulation of the output voting is performed by using the Quartus II Simulator software. The results demonstrate that the proposed controller is able to tolerate faults, which means it has extremely high reliability.

Index Terms—FPGA, TMR, multiprocessor system, voting algorithm, subsea blowout preventer

I. INTRODUCTION

Subsea Blowout Preventer (BOP) plays an essential role in providing safety during the subsea drilling activities. Great damage could be caused by the failures of subsea BOP stack, for example, the deep-sea petroleum drilling rig Deepwater Horizon exploded and oil spilt off the coast of Louisiana on April 20, 2010. Eleven workers died in the explosions [1]. Oil gushed out of the damaged well for two months, which is the worst environmental disaster in US history. The Deepwater Horizon accident was considered that on the rig, the subsea BOP did not isolate the well before and after the explosions. The subsea BOP stack might not work before the blowout or it might have been damaged due to the accident [2]. For subsea BOP stack, extreme reliability is necessary.

Reliability is becoming more and more important in supporting next-generation science, engineering and applications, as digital systems become increasingly large and complex. Fault tolerance is a critical approach to improve the system reliability, which is the ability of a

system to operate normally in the presence of faults. It is a significant feature for all kinds of operating environments. This type of reliability is usually obtained through hardware redundancy and Triple Modular Redundancy (TMR) can serve as an example. TMR uses three identical logic blocks performing the same task with corresponding outputs being compared through majority voters [3]. It is a static hardware redundancy scheme for masking a single fault in a digital circuit. Any error in one of the three circuit copies will be masked by the majority voter [4].

Reconfigurable hardware such as Field Programmable Gate Array (FPGA) is widely used in numerous applications such as industrial electronic devices and embedded systems. Because of their high flexibility and throughput, they can achieve multiple requirements such as high performance, low cost and fast turnaround time. Reliable systems implemented into FPGAs can be realized on different design levels. It can be implemented in different ways, including level of separate FPGA units, level of one FPGA and level of functional units [5].

In recent years, various fault-tolerant systems based on FPGAs have been developed. TMR is a fault tolerance method for protecting FPGA designs against single-event upsets (SEUs) caused by radiation. Because of its straightforward implementation and reliable results, TMR is widely used in space applications [6-9]. Ref. [10] presented an algorithm for roadway path extraction and tracking and it was implemented in a FPGA device. FPGA can be used to implement motor controller devices in accordance with the actual core-based design [11]. Ref. [12] have developed a FPGA-based Charge Coupled Device(CCD) data acquisition system, which will be used in the focal plane of Soft X-ray Telescope on India's first multi-wave length astronomy satellite—ASTROSAT. Ref. [13] presented a novel hardware implementation of a disparity estimation scheme targeted to real-time Integral Photography image and video sequence compression. Ref. [14] presented a novel fault-tolerant voter circuit which itself could tolerate a fault and give error free output by improving the overall system's reliability.

In the multiprocessor systems-on-chip field, FPGA-based multiprocessor is a new and increasingly important

trend. Compared with Application Specific Integrated Circuit multiprocessor system, it has advantages of flexibility and reconfiguration, less time-to-market and less cost [15]. Ref. [16] proposed an original design together with an efficient implementation of an authoritative domain name system server on a Virtex 5 FPGA circuit. New definitions for high level performance metrics such as efficiency, scalability and robustness were proposed to overcome some limitations with FPGA-based multiprocessor systems [17]. Ref. [18] proposed an Processor Allocator architecture, which was based on bit map approach and was driven by an Improved First Fit algorithm compared with previously known important techniques.

The subsea BOP control system using Programmable Logic Controller (PLC) has been accomplished by authors, which is able to monitor and control subsea BOP functions. GE Fanuc Genius Modular Redundancy system can provide supervisory control and data acquisition, which consists of a number of modular subsystems [19].

This paper presents a redundant controller based on FPGAs instead of PLCs for subsea blowout preventer. The advantages of FPGAs can reduce the costs, size of the control system and deal with complicated situations. Besides, the flexibility and scalability of control system can be improved. The volume of pressure vessel housing electronic equipments against external hydrostatic pressure can also been reduced, which enhances the reliability and cut down costs of the sealed vessel. The proposed redundant controller is implemented using three FPGA development boards. The paper is structured as follows: Section II describes the architecture of the redundant controller and the functions of the components. Section III describes the implementation of the system, including hardware development and Nios II system implementation. Section IV proposes different voting algorithms for discrete input, analog input and discrete output and performs the functional simulation. Section V summarizes the paper.

II. ARCHITECTURE OF THE REDUNDANT CONTROLLER

Exploitable reservoirs of oil and gas are rare and remote. This leads to increased subsea deepwater well exploration and requires BOPs to remain submerged for a long time in extreme conditions. Therefore, BOP components have grown larger and heavier. However, the space allotted for BOP stacks on existing offshore rigs has not grown accordingly. Thus an important focus on the technological development of subsea BOPs over the past two decades has been decreasing their footprint and weight while simultaneously increasing safe operating capacity. The proposed FPGAs-based system can reduce the volume of pressure vessel housing the electrical equipments and improves safety and reliability of the control system.

Controller is the kernel of BOP control system, which can provide supervisory control and data acquisition. Since the controller is required to have extremely high reliability, TMR technique is employed to develop the controller.

As shown in Fig. 1, TMR is done by implementing

three copies of the same circuit and performing majority voting on the output of the triplicate circuit. Majority voter performs the function of outputting the logic value that corresponds to at least two of its inputs. For example, if two or more of the voter's three inputs are '1', then the output of the voter is '1' [14].

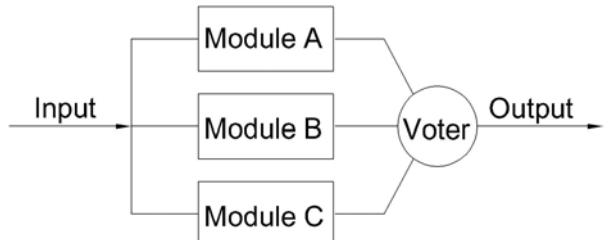


Figure 1. Triple modular redundancy with voter

The block diagram of the proposed redundant controller is shown in Fig. 2. Different redundancy design levels are developed for the controller. Firstly, the fault tolerance system is implemented on the level of separate FPGA boards. This is often implemented in spite of high price, power consumption and size of implementation, but, it has high reliability. FPGA0, FPGA1 and FPGA2 are the three components. And secondly, TMR system is implemented on the level of one FPGA, which is made up of CPU0, CPU1 and CPU2. Reliability of the processors has also been improved. Every FPGA board is connected to the others respectively, transmitting and receiving data. The three FPGAs are designed in exactly the same way. Everyone consists of one input voter, three CPUs, one output voter and one board voter.

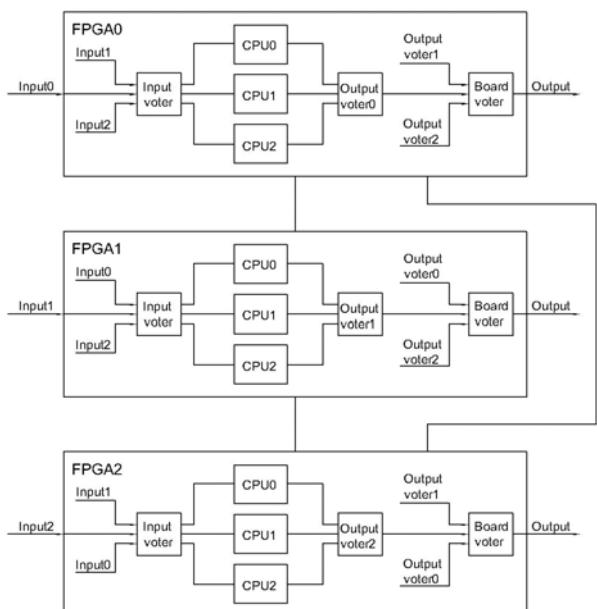


Figure 2. Block diagram of the controller

Input voter in each FPGA executes different voting algorithm for discrete and analog inputs. The input voter in one FPGA can get access to the input values of the other FPGAs through the connecting lines, where all three signals are compared. One processor is responsible for

communication between FPGA boards, which also serves as an input voter. It contains voting logic developed by ANSI C language. After voting, the voted value is obtained by the application programs in CPUs.

Classical TMR redundancy scheme is adopted for masking a single fault in each FPGA. There are three CPUs in each FPGA, which can greatly improve the reliability of processors. The same application programs are performed in all CPUs, including the control logic for subsea BOP. Processors' output signals are voted in the majority voters, output voter0, voter1 and voter2. Any error in one of the three processors will be masked by the majority voters. What's more, if one of them fails, the controller can continue working faultlessly.

The "board voter" is responsible for performing 2 out 3 voting. After the results generated by CPUs are voted in output voters, the three voted results will be voted again in the "board voter" before generating the final output value. It is voted twice altogether, which greatly improves the reliability of the system. Any FPGA can be connected to the controlled device directly or indirectly, such as solenoid valves.

III. IMPLEMENTATION OF THE CONTROLLER

A. Hardware Platform

As shown in Fig. 3, the proposed system is implemented using three Altera Cyclone II FPGA development boards. It is very powerful, relatively inexpensive and adaptable, since its configuration is specified in an abstract hardware description language. This kit includes an EP2C35F484C6N FPGA chip with 33,216 logic elements and 473 Kbits of on-chip RAM. The development board includes a 50 MHz clock, 1MB SRAM, 8 MB flash, 32MB SDRAM and 170 user's IO pins. Software development is performed using Nios II 8.0 IDE and hardware development is performed using Quartus II 8.0.

RS-232 is widely used as a serial interface in a PC computer and communication industries. Because the amount of exchange data is not large and the cable length between the boards is less than 50 feet, this communication method is chosen to connect the FPGA boards. The in-out voltage levels of FPGA are TTL levels, while RS-232 is a standard serial interface, so both the electrical specifications are inconsistent. Therefore, a TTL to RS-232 adaptor is needed. MAX232 chip is used to transform the TTL voltage to the required voltage. The self-made adapter is shown in Fig. 3.

The UART module is an Altera SOPC Builder library component included in the Nios II development kit. It is a common serial interface with variable baud rate, parity, stop and data bits, and optional control signals. The SOPC Builder UART library component has available system options to define device logic and interface signals. It implements simple RS-232 asynchronous transmit and receive logic inside an Altera device.

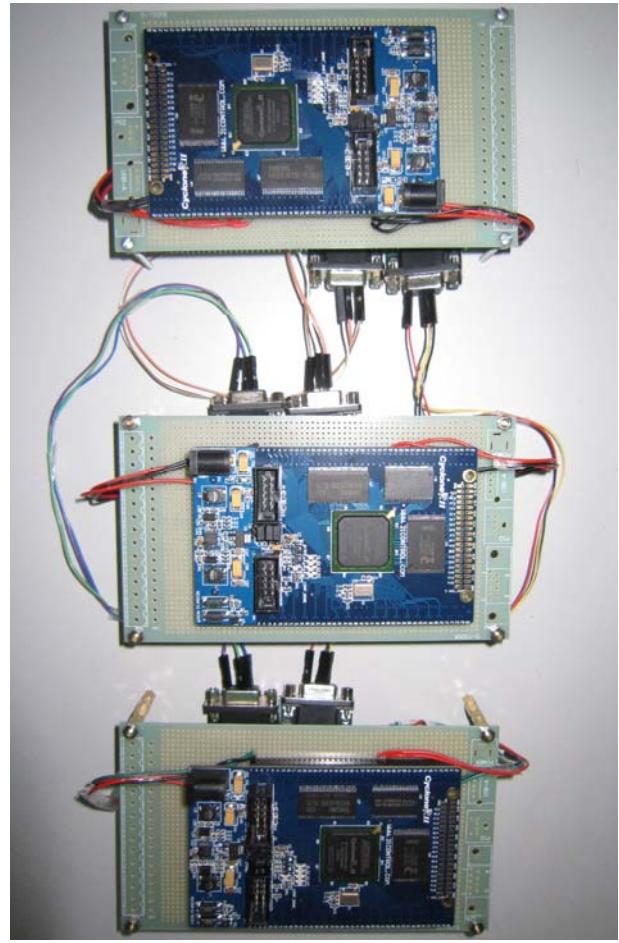
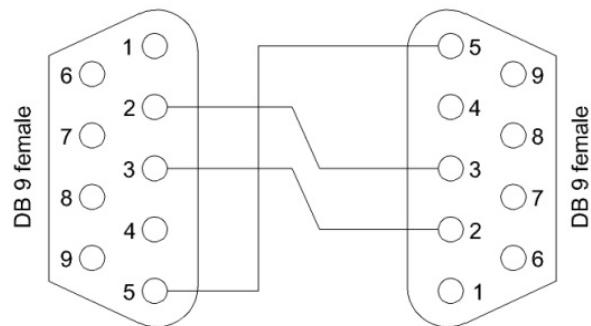


Figure 3. FPGA based controller

As shown in Fig. 4, minimal "3-wire" RS-232 connection method is applied between the serial interfaces, which consists only of transmit data, receive data, and ground. It is commonly used when the full facilities of RS-232 are not required.



Connector 1	Connector 2	Function
2	3	Rx ← Tx
3	2	Tx → Rx
5	5	Signal ground

Figure 4. "3-wire" Connection method

The UART sends and receives serial data over two external pins (Rx and Tx). Software controls and communicates with the UART through five memory-

mapped, 16-bit registers [20]. Its baud rate is set to 115200 bps, with no parity, 8 data bite and 1 stop bite in the SOPC builder interface.

B. Development of Multiprocessor System

Altera's Nios II is a soft processor, defined in a hardware description language, which can be implemented in Altera's FPGA devices by using the Quartus II software. The Nios II processor can be employed with a variety of other components to build a complete system. Altera's Development board contains several components which can be integrated into one Nios II system [21].

Altera's FPGAs allow for the realization of several processors simultaneously. As we know, multiprocessor systems are an effective method to improve system performance and to concentrate processing components in one FPGA. The number of processors can be used in an FPGA's system is only determined by the quantity of device resources allowing designers to implement complex multiprocessors architectures for specific or general purpose systems.

In this paper, multiprocessor system is used for processor redundancy, communication and voting. In fact, there are four Nios II processors in each FPGA, which are embedded-processors designed specifically for the Altera family of FPGAs. As shown in Fig. 5, they are a shared memory multiprocessor system. One processor is responsible for communication between the FPGA boards and serves as an input voter. The rest of the processors run application programs, providing supervisory control and data acquisition. Shared memory method is applied for interchanging information between processors, which is most frequently used. One important reason is that this method can save memory, since FPGAs have a limited amount of on-chip memory. Message passing method is another possible method [15].

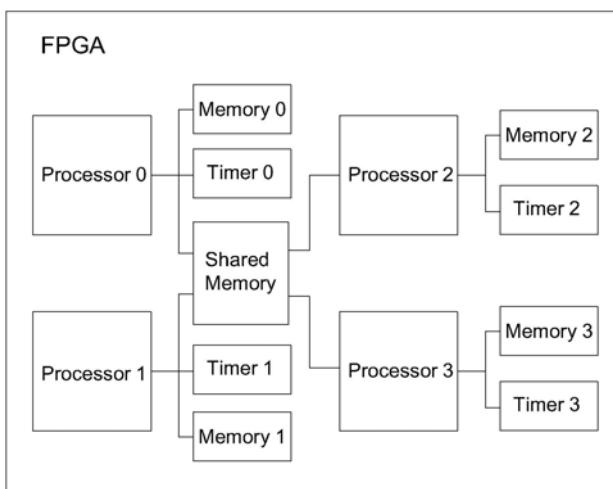


Figure 5. Shared memory multiprocessor system

In order to protect the shared resources, hardware mutex is used, which is a small SOPC Builder component. A mutex allows cooperating processors to have common agreements that one of them should be allowed mutually exclusive access to a hardware resource in the system.

This is effective to protect resources from data corruption. Since if more than one processor tries to use the resource at the same time, data corruption will happen. The mutex core is like a shared resource, providing an atomic test-and-set operation which permits a processor to test if the mutex is available and if so, to get the mutex lock in a single operation. When the processor is finished using the shared resource associated with the mutex, the processor releases the mutex lock. Now another processor can get the mutex lock and use the shared resource. Without the mutex, this sort of function will normally require the processor to perform two set of instructions, test and set, where another one can also make a test for availability and succeed. This situation would make two processors both think they successfully got mutually exclusive access to the shared resource when obviously they did not. There is one important point that the mutex core does not physically prevent resources in the system from being accessed simultaneously by multiple processors. The software running on the processors is responsible for following the rules. The software must be designed [22].

Implementation of the Nios II System

The SOPC builder tool is introduced by Altera and it can be used to create the embedded systems quickly and evaluate them easily. The integration of off-the-shelf intellectual property and a lot of reusable custom elements created by users are realized in a friendly way [23]. It can reduce the required time to set up a System-on-Programmable-Chip and enables to construct and design in hours instead of weeks. The Nios II peripherals and external memories used in the system are shown in Fig. 6. Brief introduction is made below:

- (1) All on-chip peripherals are connected to the Avalon bus. Nios II uses the Avalon switch fabric as the interface to its embedded peripherals. The traditional bus in a processor-based system allows only one bus master access the bus at a time. However, the Avalon switch fabric uses a slave-side arbitration scheme, allows multiple masters operate at the same time [24].
- (2) Processor 0, processor 1 and processor 2 are responsible for running application programs while processor 3 is responsible for communication and input voting.
- (3) Each processor has its own local on-chip memory and timer.
- (4) Shared memory is used for communication between the processors and mutex core can protect the shared memory.
- (5) UART 0 is responsible for communication with host PC, which can receive the orders from it. UART 1 and UART 2 are used to connect to the other FPGA boards through RS-232 serial ports.
- (6) Memory controllers are used to get access to the external memories. Flash can store the program files and relevant data, while SDRAM can run the application program.
- (7) PIO is used to connect to the input and output devices.

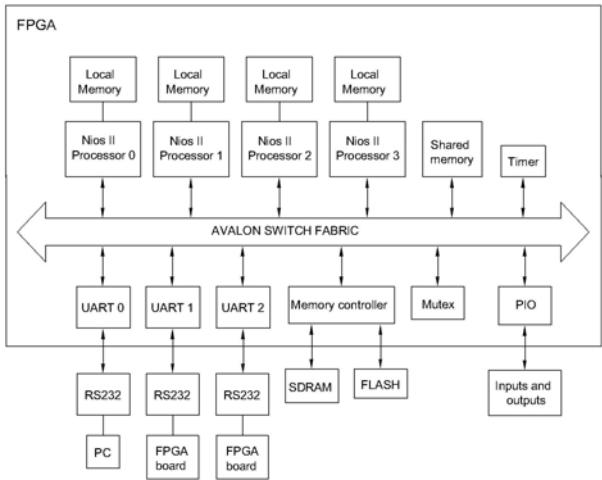


Figure 6. Block diagram of hardware architecture for one FPGA

IV. VOTING ALGORITHMS

The voting algorithms involve discrete input, analog input and discrete output. Discrete and analog input voting is handled by the input voter in each FPGA. One processor serves as an input voter, which contains the voting logic. It compares the three input values from corresponding FPGAs through RS-232 ports. Output voting is accomplished in the hardware voter of each FPGA.

A. Discrete Input Voting

The discrete signals are used to monitor the switch status of BOP components, such as pressure switch signal. Input voting is accomplished by software in the processors, unlike output voting. Besides the input values, the configured Duplex State and Default State may also be used to determine the final voted results. The Duplex State is the substitute value that is used when there are only two available input signals, while the Default State is the value that will be provided directly to the application programs instead of a voted input result. The flow chart of discrete input voting process is shown in Fig. 7.

For three available inputs, the software performs 2 out of 3 voting, with the Duplex State and Default State not used.

If one of the three input values is not available, the software uses the configured Duplex State for performing 2 out of 3 voting, in place of a third actual input. The Duplex State can be set to 0 or 1 according to the actual requirement.

If only one input signal is available or all inputs fail, voting is not performed. The Default State is used instead of the remaining actual input as the final result for the application programs in the processors, which can be set to 0 or 1.

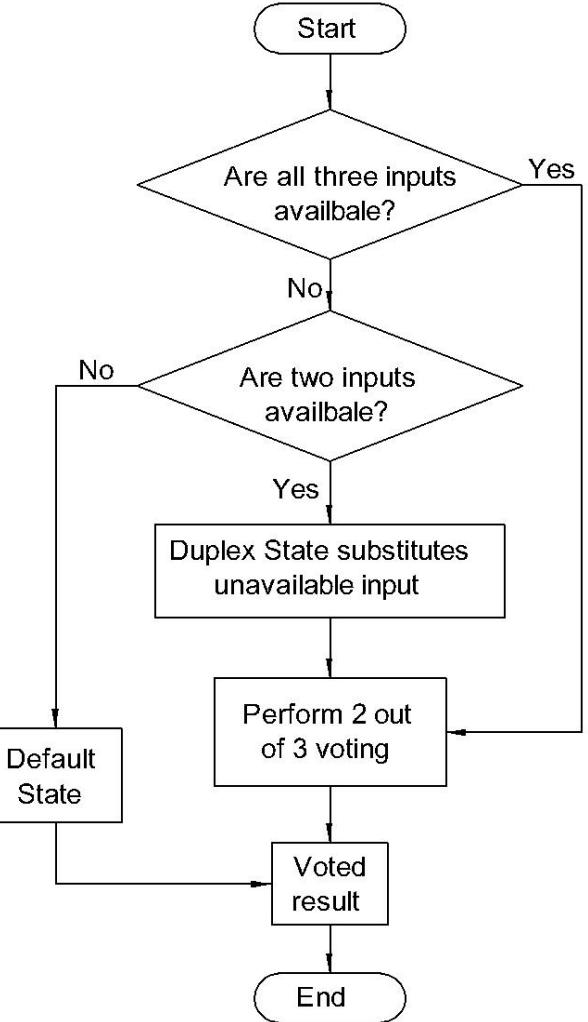


Figure 7. Flow chart of the discrete input voting process

Examples of discrete input voting in all situations are shown in Table 1, where the Duplex State is set to 0, the Default State is set 0 and “DI” stands for discrete input.

TABLE 1.
DISCRETE INPUT VOTING EXAMPLES

<i>DI_A</i>	<i>DI_B</i>	<i>DI_C</i>	Voted results
1	1	1	1
1	1	0	1
1	0	0	0
0	0	0	0
1	1	-	1
1	0	-	0
0	0	-	0
-	-	0	0
-	-	-	0

B. Analog Input Voting

The analog input values are used to monitor temperature and pressure signals. Similar to discrete input voting, analog input voting is also performed by software in the processor. Different voting algorithms are implemented for discrete and analog inputs. The configured Duplex State and Default State may also be used in determining the final voted value in the processor. However, their values are different from those of discrete input voting. The Duplex State may be configured as the

maximum, minimum, or an average of the two actual values. The Default State can be configured as the last input state, or a specific maximum or minimum value. The flow chart of analog input voting process is shown in Fig. 8.

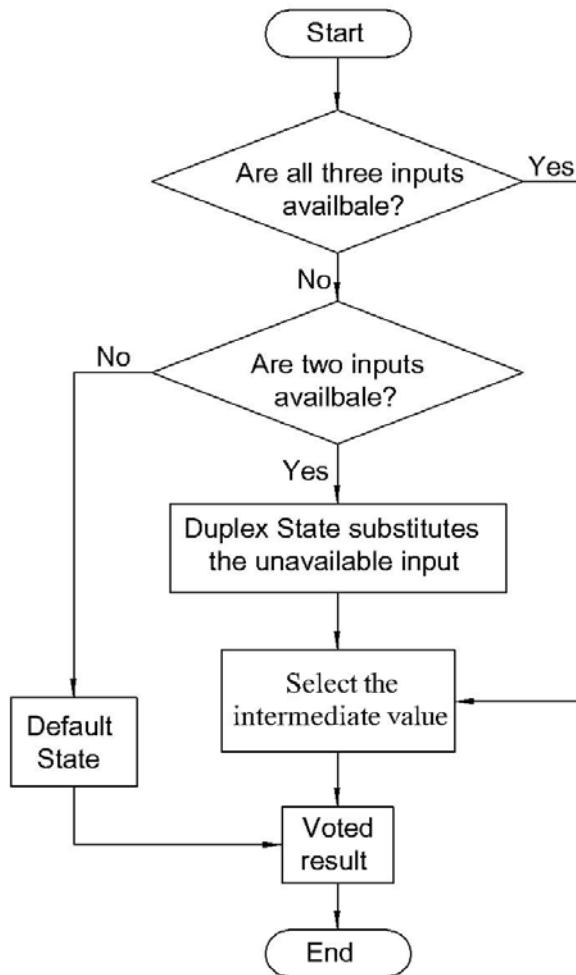


Figure 8. Flow chart of the analog input voting process

For three available inputs, the voter compares three corresponding analog input values, which selects the intermediate value. The Duplex State and Default State are not used.

If one input fails, the software uses the configured Duplex State for voting, in place of a third actual input. The intermediate value will still be selected.

If only one input is available or all inputs fail, voting is not performed. The Default State is used as the only input data.

Voting examples of all situations are shown in Table 2. The Duplex State is set to the average of the two actual values and Default State is set to 8.8. "AI" stands for analog input.

TABLE 2.
ANALOG INPUT VOTING EXAMPLES

AI_A	AI_B	AI_C	Voted results
8	9	10	9
-	9	10	9.5
8	-	10	9

AI_A	AI_B	AI_C	Voted results
-	-	10	8.8
-	9	-	8.8
8	-	-	8.8
-	-	-	8.8

C. Discrete Output Voting and Simulation

Voting of discrete output is performed in hardware voter module, which performs 2 out of 3 voting. Before the final result is voted, the output values will be voted twice. For each FPGA, results generated by processors are voted in its output voter for the first time. Then, all voted outputs from the corresponding FPGAs are voted again in the "board voter". Finally, the final output is produced by each FPGA. Due to the fact that all three FPGAs have the same architecture, anyone can be used to connect to the controlled devices directly or indirectly.

Functional simulation of the output voting is shown in Fig. 9. "FPGA0_cpu0", "FPGA0_cpu1" and "FPGA0_cpu2" denote the output signals of the processors in the first FPGA, while "output_voter0" is the voted result of the three CPUs. The other FPGAs are denoted in the same way and "voted_result" is the final output.

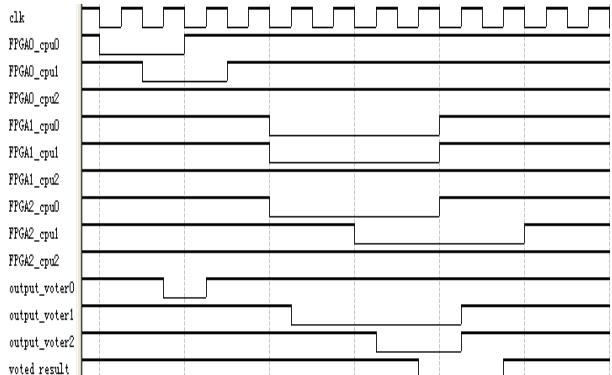


Figure 9. Functional simulation of output voting

The voting process is accomplished at posedge clk. For the signals of CPUs, suppose '1' is the correct value while '0' means an error in the simulation. It can be seen, each FPGA can tolerate one error at a time. When two of the processors fail in one FPGA, the output of this FPGA board will be faulty. However, the fault can still be masked by the board voter, which can continue work normally. It can be seen that when there are two FPGAs fail, the fault can't be masked any more. Failure of one FPGA plus one processor is the maximum tolerance of the redundant controller. The results show that the proposed redundant architecture greatly improves the reliability of the controller. It can be used for the subsea BOP control system.

V. CONCLUSIONS AND FUTURE WORK

A novel redundant controller based on FPGAs for subsea BOP control system has been proposed in this

paper. The contributions of this work can be summarized as:

- (1) The proposed hardware architecture is implemented by using three FPGA development boards and RS-232 serial ports are employed to communicate.
- (2) Nios II System is developed in each FPGA. A multiprocessor system is implemented to improve reliability of the control system and software is designed to use mutex in order to protect the shared memory.
- (3) In order to obtain reliable results, different voting algorithms for discrete input, analog input and discrete output are proposed.
- (4) Functional simulation is accomplished and it demonstrates that the proposed controller is able to tolerate faults with high reliability.

In the future research work, the control logic for subsea BOP function, human machine interface graphical design and redundant databases will be developed.

ACKNOWLEDGMENT

The authors wish to acknowledge the financial support of the National High-Technology Research and Development Program of China (No.2007AA09A101), National Natural Science Foundation of China (No.50874115), Project of Science and Technology Development of Shandong Province, "Taishan Scholar" Construction Project Special Fund Project (TS20110823) and Innovation Foundation of Graduate School of China University of Petroleum (No.CXZC11-13).

REFERENCES

- [1] W. F. Harlow, B. C. Brantley, and R. M. Harlow, "BP initial image repair strategies after the Deepwater Horizon spill," *Publ. Relat. Rev.*, vol. 37, 2011, pp.80-83.
- [2] J. E. Skogdalen, I. B. Utne, and J. E. Vinnem, "Developing safety indicators for preventing offshore oil and gas deepwater drilling blowouts," *Saf. Sci.* vol. 49, 2011, pp.1187-1199, doi:10.1016/j.ssci.2011.03.012
- [3] J. A. Cheatham, J. M. Emmert, and S. Baumgart, "A survey of fault tolerant methodologies for fpgas," *ACM Trans. Des. Autom. Electron. Syst.* vol. 11, 2006, pp. 501-533.
- [4] K. S. Morgan, D. L. McMurtrey, B. H. Pratt, and M. J. Wirthlin, "A comparison of TMR with alternative fault-tolerant design techniques for FPGAs," *IEEE Trans. Nucl. Sci.* vol. 54, 2007, pp. 2065-2072, doi:10.1109/TNS.2007.910871
- [5] M. Straka, J. Kastil, and Z. Kotasek, "Fault Tolerant Structure for SRAM-Based FPGA via Partial Dynamic Reconfiguration Digital System Design," *Proc. - Euromicro Conf. Digit. Syst. Des.: Archit., Methods Tools, DSD*, 2010, pp. 365-372, doi: 10.1109/DSD.2010.12
- [6] X. X. She and K. S. McElvain, "Time Multiplexed Triple Modular Redundancy for Single Event Upset Mitigation," *IEEE Trans. Nucl. Sci.* vol. 56, 2009, pp. 2443-2448, doi: 10.1109/TNS.2009.2021656
- [7] B. Pratt, M. Caffrey, J. F. Carroll, P. Graham, K. Morgan, and et al. "Fine-grain SEU mitigation for FPGAs using partial TMR," *IEEE Trans. Nucl. Sci.* vol. 55, 2008, pp. 2274-2280, doi:10.1109/TNS.2008.2000852
- [8] L. Sterpone and M. Violante, "A new reliability-oriented place and route algorithm for SRAM-based FPGAs," *IEEE Trans. Comput.*, vol. 55, no.6, pp. 732-744, Jun. 2006.
- [9] S. Golshan and E. Bozorgzadeh, "SEU-aware resource binding formodular redundancy based designs on FPGA," in Proc. DATE, Apr.2009, pp. 1124-1129.
- [10] R. Marzotto, P.Zoratti, D. Bagni, A. Colombari, and V.Murino, "A real-time versatile roadway path extraction and tracking on an FPGA platform," *Comput Vision Image Understanding*, vol. 114, 2010, pp. 1164-1179, doi:10.1016/j.cviu.2010.03.015
- [11] A. Astarloa, J. Lázaro, U. Bidarte, J. Jiménez, and A.Zuloaga, "FPGA technology for multi-axis control systems," *Mechatronics*, vol. 19, 2009, pp. 258-268, doi:10.1016/j.mechatronics.2008.09.001
- [12] A. Kothare, I. Mirza, K. P. Singh, and A. F. Abbey, "FPGA-based flexible CCD control system for X-ray astronomy payloads," *Nucl Instrum Methods Phys Res Sect A*, vol. 604, 2009, pp. 747-754, doi:10.1016/j.nima.2009.01.103
- [13] D. Chaikalis, N. Sgouros, D. Maroulis, and P. Papageorgas, "Hardware implementation of a disparity estimation scheme for real-time compression in 3D imaging applications," *J Visual Commun Image Represent*, vol. 19, 2008, pp. 1-11, doi:10.1016/j.jvcir.2007.09.003
- [14] R.V. Kshirsagar and R.M. Patrikar, "Design of a novel fault-tolerant voter circuit for TMR implementation to improve reliability in digital circuits," *Microelectron. Reliab.* vol. 49, 2009, pp. 1573-1577, doi:10.1016/j.microrel.2009.08.001
- [15] T. Dorta, J. Jimenez, J. L. Martin, U. Bidarte, and A. Astarloa, "overview of FPGA-Based Multiprocessor systems," *ReConFig - Int. Conf. ReConfigurable Comput. FPGAs*, 2009, pp. 273-278, doi: 10.1109/ReConFig.2009.15
- [16] R. Sadoun, A. Belouchrani, E. B. Bourennane, and A. Zerguerras, "An FPGA based soft multiprocessor for DNS/DNSSEC authoritative server," *Microprocessors Microsyst*, vol. 35, 2011, pp. 473-483, doi:10.1016/j.micpro.2011.03.004
- [17] M. Beltrán, A. Guzmán, and F. Sevillano, "High level performance metrics for FPGA-based multiprocessor systems," *Perform Eval*, vol. 67, 2010, pp. 417-431, doi:10.1016/j.peva.2009.12.004
- [18] D. Zydek and H. Selvaraj, "Hardware implementation of processor allocation schemes for mesh-based chip multiprocessors," *Microprocessors Microsyst*, vol. 34, 2010, 39-48, doi:10.1016/j.micpro.2009.11.003
- [19] B. P. Cai, Y. H. Liu, Z. K. Liu, F. Wang, X. J. Tian, and et al, "Development of Automatic Subsea Blowout Preventer Stack Control System using PLC based SCADA," unpublished.
- [20] Nios UART, ALTERA, <http://www.altera.com/literature/ds/ds_nios_uart.pdf>
- [21] Introduction, Nios II Processor Reference Handbook, ALTERA, <http://www.altera.com.cn/literature/hb/nios2/n2cpu_nii51001.pdf>
- [22] Creating multiprocessor Nios II system tutorial, ALTERA 2011,<www.altera.com/literature/tt_tt_nios2_multiprocessor_tutorial.pdf>
- [23] SOPC Builder Applications, ALTERA 2006, <<http://www.altera.com/products/software/products/sopc/sopindex.html>>

[24] Avalon Interface Specifications, ALTERA, <www.altera.com/literature/manual/mnl_avalon_spec.pdf>

Zengkai Liu was born in Shandong, P. R. China, in 1986. He received his B.S. and M.S. degree in Mechanical Engineering from China University of Petroleum in 2009 and 2011 respectively. Currently, he is a Ph.D. candidate in Mechanical Engineering in China University of Petroleum, P. R. China. His recent research interest is the Field Programmable Gate Array based control system for subsea blowout preventer.

Yonghong Liu was born in Anhui, P. R. China, in 1965. He received his Ph.D. degree in Mechanical Manufacture from Harbin Institute of Technology, Harbin, China, in 1996. He is currently a professor and doctoral supervisor in College of Electromechanical Engineering at China University of Petroleum, China. He has published over 120 papers in some international or national journals and conferences. His current research interests include the Expansion Sand Screen for sand control, new kind of numerical control machine about combined machining and control system of the subsea drilling equipments. Dr. Liu is a member of China Nontraditional Machining Committee and Nontraditional Machining Association of Shandong. Dr. Liu is the candidate for the New Century National "Hundred, Thousand and Ten Thousand Talents Project" and prominent young and middle-aged specialist of Shandong Province.

Baoping Cai was born in Hebei, P. R. China, in 1982. He received his B.S. and M.S. degree in Mechanical Engineering from China University of Petroleum in 2006 and 2008 respectively. Currently, he is a Ph.D. candidate in Mechatronic Engineering in China University of Petroleum, P. R. China. His

recent research interest is the control system of the subsea drilling equipments.

Fei Wang was born in Shandong, P. R. China, in 1985. He received his B.S. and M.S. degree in Mechanical Engineering from China University of Petroleum in 2008 and 2010 respectively. Currently, he is a Ph.D. candidate in Mechanical Design and Theory in China University of Petroleum, China. His recent research interest is the control system of NC machine tool.

Zhili Chen was born in Shanxi, P. R. China, in 1984. He received his B.S. degree in Machinery Design and Manufacturing and Its Automation from China University of Petroleum in 2009. Currently, he is a postgraduate student in Mechanical Engineering in China University of Petroleum, P. R. China. His recent research interest is the design of NC machining simulation system

Xiaojie Tian was born in Shandong, P. R. China, in 1985. He received his B.S. and M.S. degree in Mechanical Engineering from China University of Petroleum in 2008 and 2010 respectively. Currently, he is a Ph.D. candidate in Mechanical Design and Theory in China University of Petroleum, China. His recent research interest is the control system of the subsea drilling equipments based on the field bus.

Yazhou Wang was born in Shandong, P. R. China, in 1988. He received his B.S. degree in Machinery Design and Manufacturing and Its Automation from China University of Petroleum in 2009. Currently, he is a postgraduate student in Mechanical Engineering in China University of Petroleum, P. R. China. His recent research interest is the design and analysis of the pressure vessel used in subsea equipments.

A Non-Standard Approach for the OWL Ontologies Checking and Reasoning

Yingjie Song

Dalian Maritime University, Dalian 116026, P.R. China

Email:tiantianyingjie@gmail.com

Rong Chen

Dalian Maritime University, Dalian 116026, P.R. China

Email: tsmc.dmu@gmail.com

Yaqing Liu^{1,2}

¹Dalian Maritime University, Dalian 116026, P. R. China

²Artificial Intelligence Key Laboratory of Sichuan Province, Zigong, China

Email: liuyaqing234@yeah.net

Abstract—The Semantic Web is the extension of the World Wide Web that enables people to share content beyond the boundaries of applications and websites. The understanding of Semantic Web documents is built upon ontologies that define concepts and relationships of data. Hence, the correctness of ontologies is vital. In this paper, we propose a new algorithm combined with the software engineering techniques, such as Alloy modeling language and its reasoner Alloy Analyzer to provide checking and reasoning service for OWL ontologies. First of all, we use Jena to parse OWL ontology documents. Next, the intermediate results are used as the inputs of the algorithms to generate the Alloy model. Futher, with the assistance of Alloy Analyzer, the Alloy model is checked. Experimental results show that this method can be carried out large-scale ontology reasoning and complex-property reasoning which are different from traditional ontology reasoning. Furthermore, the results provide useful information to guide the ontology modification.

Index Terms—Ontology Reasoning, OWL, Alloy, Semantic Web

I. INTRODUCTION

Tim Berners-Lee's [3] original vision for the Semantic Web was that information would be just as readable (and understandable) to a person or to a machine. The aim of the Semantic Web is to make Web resources more readily accessible to automated processes. Digital objects, whether web page, image, video, or some other file, would have embedded within them meta data that would provide context to the content and allow software to extract meaning from the file. To make sure that different agents have a common understanding of these, ontology is needed to describe. Basically, an ontology [11] is a collection of definitions of concepts and the shared understanding comes from the fact that all the agents interpret the concepts in the same way. The importance of ontologies in Semantic Web has prompted the

development of several ontology languages. Many ontology languages have been developed, especially for the semantic web, such as OIL [12], DAML [2], DAML+OIL [6] and OWL [14]. OIL (Ontology Interchange Language) is based on three elements, namely, frame-based systems, description logics, and web standards [10]. DAML (DARPA Agent Markup Language) is developed in DARPA DAML programme. DAML+OIL is a language for expressing far more sophisticated classifications and properties of resources than RDFS. OWL, which is aimed to be the standardized and broadly accepted ontology language of the Semantic Web, is compatible with early ontology languages and provides the engineer more power to express semantics.

Reasoning with ontology languages is important to ensure the quality of an ontology. Indeed reasoning can be employed in different phases during the design, maintenance and deployment of ontology [9]. Some of the popular reasoners that are available such as RACER [22], FaCT [12], FaCT++ [21], Pellet [19]. They are all based on the description logic. For OWL, a standard ontology language for the semantic web, these reasoners and algorithms can be used to solve the reasoning problems, in particular satisfiability. When attempting to use description logic to provide reasoning services for OWL, there are many differences between OWL and description logics [1]. OWL's RDF syntax allows circular syntactic structures, in addition, the description logic does not include contain-features of OWL.

Software engineering methods and tools will be introduced in our solution. In our previous work [20], this idea has been used for DAML + OIL ontology testing, and the experiment proved that this method could work on a larger scope of property checking. This paper, the lightweight modelling language for software design Alloy [15] and its fully automatic analysis tool Alloy Analyzer [16] are used for ontology reasoning. Alloy can solve the problem which dose not exist in the description logic. It's

the characteristics of circular structure and contain features that are more conducive to the expression of OWL ontologies. Alloy Analyzer is a software tool which can be used to analyze specifications written in the Alloy specification language. The Alloy Analyzer supports the analysis of partial models. As a result, it can perform incremental analysis of models as they are constructed, and provide immediate feedback to users. Alloy Analyzer is based on the new SAT-based model finder Kodkod. Kodkod applies new techniques and optimizations to the translation from relational to boolean logic (<http://alloy.mit.edu/alloy4/>). In our approach, the Alloy Analyzer is used to analyze the Alloy model based on the OWL ontology, and provide automatic reasoning and consistency checking for the semantic Web.

After a brief introduction and quick survey of OWL, Section 3 and 4 a simple ontology example is given which is described in OWL, after the analysis by jena, we give the algorithms which is used to transfer the results into Alloy model. Section 5 surveys some of the major problems that had to be resolved in the reasoning of ontologies with Alloy Analyzer, while Section 6 gives a simple example to illustrate the process of reasoning and the problems can be found easily. And Section 7 gives summary of the article.

II. THE OWL WEB ONTOLOGY LANGUAGE

A. Logical Characteristic of OWL

OWL is an ontology language that has recently been developed by the W3C Web Ontology Working Group. It is intended to provide a language that can be used to describe the classes and relations between them that are inherent in Web documents and application. OWL is defined as an extension to RDF in the form of a vocabulary entailment, i.e., the syntax of OWL is the syntax of RDF and the semantics of OWL are an extension of the semantics of RDF. In order to meet the different expressing power and computational efficiency, it comes in three flavors: OWL DL, OWL Lite and OWL FULL. OWL FULL, which contains all the language element of OWL, is the complete works of the language, and it is a syntactic and semantic extension of RDFS. OWL DL, a subset of OWL FULL, is a version of OWL with decidable inference that can be written in a frame or Description Logic manner. OWL Lite is a subset of OWL DL, and also has a lower formal complexity than OWL DL. It is the least expressive species of OWL, but it ensures the efficient reasoning.

B. Reasoners for Semantic Web

The current main reasoning tools for Semantic Web are based on description logic, such as the most widely used three well-known DL reasoners: RACER, FaCT++ and Pellet.

RACER not only can be used as a description logical system, but also can be regarded as a representation system that implements a highly optimized tableau calculus. It implements the description logic SHIQ., and facilities for algebraic reasoning including concrete

domains. But RACER can only support ABox reasoning completely.

FaCT++ is the new generation of the well-known FaCT (Fast Classification of Terminologies). A new tableaux decision procedure for SHIQ(D) is implemented. The FaCT++ system provides satisfiability testing for modal logic. In order to create a more efficient software tool and to maximize portability, it is implemented as a free open-source C++-based reasoned. However, FaCT++ can't take OWL documents directly nor any remote file. The OWL ontologies have to be translated into DIG format for the system.

Pellet is an open source, Java reasoner for OWL DL ontologies. It is based on the tableaux algorithms developed for expressive description logics. It is claimed to provide functionalities to see the species validation, check consistency of ontologies, classify the taxonomy, check entailments and answer a subset of RDOL queries.

Though these tools can reason ontologies with a high degree of automation, the complex-properties ontologies still can't be checked by them effectively. Furthermore, these reasoners can find that the ontology is error, but not be able to find out the reason.

C. Alloy and Alloy Analyzer

Alloy is a lightweight modelling language for software design, which is widely accepted as micromodels of software in the software engineering community. It is a first order relation logic and treats relations as the first element. An Alloy model consists of Signatures, Relations, Facts, Functions and Predicates. Signatures represent the entities of a system and Relations are used to describe relations between such entities. Facts and Predicates introduce constraints over such Signatures and Relations. Whereas Facts are constraints to be always valid, Predicates are named parameterized constraints for depicting operations, Functions are named expression with parameters that return results.

Alloy Analyzer is a tool for analyzing models written in Alloy. The Alloy Analyzer 4 is based on the new SAT-based model finder to support fully automated analysis of Alloy models through simulation and Assertion checking. Given a user specified scope on the model elements bounding the domain, the analyzer first translates an Alloy model into boolean formulas, and then invokes a SAT-solver to find an instance. If an instance that violates the assertion is found within the scope, the assertion is not valid and the instance is returned as a counterexample.

III. REASONING FOR OWL ONTOLOGIES

This section, we will introduce the reasoning process. In the reasoning, Jena [5], Alloy and Alloy Analyzer will be used. Jena is fundamentally an RDF platform, and it supports ontology formalisms built on top of RDF, such as DAML+OIL and OWL. As is shown in Figure 1, according to Jena, we can get Class C, Property P and Statements S in the results. Next, owl2Alloy algorithm is used to translate the results into Alloy model. And then Alloy model is analyzed by the Alloy analyzer. If there

are some errors in the ontology, they must be analyzed and corrected. The modified ontology must be reasoned again.

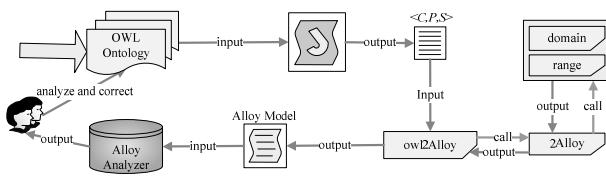


Figure 1. The specific process of ontology reasoning.

A. Parsing OWL Documents

For the requirements of reasoning, classes, properties and statements are wanted. First, Jena, which is a Java framework for building Semantic Web application, is used to analyze the OWL ontology documents. We created an ontology model of *OntModel* for handing OWL ontologies, and loaded an ontology document into the model using the *read* method. The *listClasses()* method can pick out the classes of the ontology, the *listOntProperties()* method can answer an iterator over all of the ontology properties, and the *listStatements()* method can return an iterator over all the statements in this model.

There is an ontology document in Figure 2. It contains something about animal. There are four classes defined: *Animal*, *Male*, *Man* and *Female*. *Animal* is the base class, *Male* and *Female* are disjoint subclass of the base class. *Man* is subclass of *Male*. There are also three properties in the document, they are *hasFather*, *hasParent* and *hasChild*. The properties *hasParent* and *hasChild* are inverse of each other, and the property *hasFather* is subproperty of *hasParent*. We will use Jena to parse the ontology document into three sections, which are classes, properties and statements.

```

.....
<owl:Class rdf:ID="Animal"/>
<owl:Class rdf:ID="Male">
  <rdfs:subClassOf rdf:resource="#Animal"/>
</owl:Class>

<owl:Class rdf:ID="Female">
  <rdfs:subClassOf rdf:resource="#Animal"/>
  <owl:disjointWith rdf:resource="#Male"/>
</owl:Class>
<owl:Class rdf:ID="Man">
  <rdfs:subClassOf rdf:resource="#Male"/>
</owl:Class>
.....
<owl:ObjectProperty rdf:ID="hasParent">
  <rdfs:domain rdf:resource="#Animal"/>
  <rdfs:range rdf:resource="#Animal"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:ID="hasFather">
  <rdfs:subPropertyOf rdf:resource="hasParent"/>
  <rdfs:range rdf:resource="#Male"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:ID="hasChild">
  <rdf:inverseOf rdf:resource="hasParent"/>
</owl:ObjectProperty>
  
```

Figure 2. The specific process of ontology reasoning.

The output of Jena contains a sequence of *Classes*, *Properties* and *Statements*, each having a counterpart in the original document. It can reflect the structure of animal ontology accurately. The three sections can mirror OWL vocabulary. Except *Classes*, *Properties* corresponding to the Class and Property of OWL, the rest of OWL language features can be converted into *Statements*. Jena read and parsed the animal ontology document, the result is shown in Figure 3.

```

<OWLClass Animal>
<OWLClass Male>
<OWLClass Man>
<OWLClass Female>
The number of Classes is: 4
<OWLProperty hasFather>
<OWLProperty hasParent>
<OWLProperty hasChild>
The number of Properties is: 3
[<Male>, subClassOf, <Animal>]
[<Female>, subClassOf, <Animal>]
[<Female>, disjointWith, <Male>]
[<Man>, subClassOf, <Male>]
[<hasFather>, range, <Male>]
[<hasFather>, subPropertyOf, <hasParent>]
[<hasParent>, domain, <Animal>]
[<hasParent>, range, <Animal>]
[<hasChild>, inverseOf, <hasParent>]
The number of Statements is: 9.
  
```

Figure 3. Jena outputs of “Animal” ontology

B. The Algorithms to Generate Alloy Model

The different parts of the parsing result can be converted into different Alloy element. In Alloy, *signature* represents a set of atoms. If there are two kinds of signatures which have a father-son relation, the signature that extends another signature is said to be a subsignature of the signature it extends, and its type is taken to be a subtype of the type of the signature extended, and is declared using the *in* keyword *extends*. This is similar to the relation between classes, so *Classes* are converted into signatures. The Alloy universe consists of *atoms* and *relations*, although everything that you can get your hands on and describe in the language is a relation. The rest of the outputs of Jena appears in the Alloy model in the form of *relation*. As shown below, our Algorithm *owl2Alloy*(*C*, *P*, *S*, Σ) converts the input *Classes* (denoted as *C*), *Properties* (denoted as *P*) and *Statements* (denoted as *S*) into a textual Alloy model Σ .

Algorithm 1: owl2Alloy(C, P, S, Σ) 建议用图来表示	
Input: a set C of Classes, a set P of Properties and a set S of Statements	
Output: a textual Alloy model Σ	
1. $\leftarrow \omega;$	
2. FOR each $c \in C$	
3. $\leftarrow \Sigma + 2\text{Alloy}(c, S);$	
4. FOR each $s \in S$	
5. IF $s.\text{predicate.name} = \text{"disjointWith"}$	
6. IF $s.\text{subject.subClassOf} \neq s.\text{object.subClassOf}$	
7. $\leftarrow \Sigma + \text{"pred } \{ \text{no } c1: " + s.\text{subject.name} + ", c2: " + s.\text{object.name} + " \mid c1 = c2 \}";$	
8. IF $s.\text{predicate.name} = \text{"complementOf"}$	
9. $\leftarrow \Sigma + \text{"pred } \{ " + s.\text{subject.name} + " = " + C - s.\text{object.name} + " \};$	
10. IF $s.\text{predicate}$ meets one of Conditions in Table 1 the corresponding generated Alloy predicate is appended to Σ	
11. RETURN Σ	

The owl2Alloy algorithm's input is the parsing result of Jena, and then it will generate the Alloy model Σ , which contains Alloy signatures and the relations between them. The algorithm must call the other two algorithms, 2Alloy(c, S) and domain (p, S) to complete the translation process. The top-level algorithm initializes the Alloy model Σ as an empty string ω above all. For each c , there will be a corresponding signature of the same name added to Σ , while the 2Alloy(c, S) algorithm is invoking. Next, we deal with the *Statements* through lines 4~11. *Statements* are composed of three segments like $\langle \text{subject}, \text{predicate}, \text{object} \rangle$. In the algorithm, they are represented by $s.\text{subject}$, $s.\text{predicate}$ and $s.\text{object}$ respectively. We mainly based the second one to determine the coming form. We use e to represent one of the three segments, and $e.\text{name}$ to represent the name of it. The different name of predicate will generate different predicate in Σ . Lines 5~11 identify the name of each $e.\text{predicate}$ to determine the predicates in the coming model Σ . If $e.\text{predicate.name}$ is 'disjointWith' (lines 5~7), there will be a predicate to declare that $e.\text{subject}$ and $e.\text{object}$ are disjoint from each other, and there is no individual belong to both of them. Lines 8~9 are used for the condition which $e.\text{subject}$ is a subclass of the complement of $e.\text{object}$. The others meet one of the condition in Table 1

The owl2Alloy algorithm also calls the other sub-algorithms, in which, the 2Alloy(c, S) algorithm is to generate a signature for a certain class c from the outputs of Jean. In addition to the class c , a set S of *Statements* is also used as inputs of the algorithm. At first, the first line produces a signature which is named after the input class c . Lines 2~4, to determine whether the class c is subclass of another. If $c.\text{subClassOf}$ is not empty, that is to say c is subclass of another. We use keyword "extends" in Alloy to connect the parent-signature. From 5 to 8 lines, it comes to determine the relationship between the classes. There is a property p , sub-algorithm domain (p, S) is used to determine whether c is its domain. If it comes to this, we use range (p, S) algorithm to get the range of p . Then it represents with the performance of Alloy language likes "domain {property: range}". In line 9, we get the signature σ .

TABLE I.
MORE CASES FOR CONVERTING STATEMENTS TO ALLOY

	Condition	Alloy predicates
1	$s.\text{predicate.name} = \text{"subProp}$ $e = \text{"of"}$	"pred subPropertyOf{all r:" + $s.\text{subject.range} + " r \text{ in } " +$ $s.\text{object.range} + "}"$
2	$s.\text{predicate.name} = \text{"equivalentClass"}$	"pred equivalentClass{ " + $s.\text{subject} +$ $e = " + s.\text{object} + "}"$
3	$s.\text{predicate.name} = \text{"equivalentProperty"}$ $e = " + s.\text{subject} + " = " + s.\text{object} + "$	"pred equivalentProperty { "+s.\text{subject} + " = " + s.\text{object} + "}"
4	$s.\text{predicate.name} = \text{"inverseOf"}$	"pred inverseOf { "+s.\text{subject} + " = ~ " + s.\text{object} + "}"
5	$s.\text{predicate.name} = \text{"TransitiveProperty"}$	"pred TransitivePropertyOf { a,b,c \in " + s.\text{subject} + " /a, (" + s.\text{predicate} + ") = b && b, (" + s.\text{predicate} + ") = c ⇒ a, (" + s.\text{predicate} + ") = c)"
6	$s.\text{predicate.name} = \text{"hasValue"}$	"pred hasValue {#(" + s.\text{predicate.range} + ") = 1}
7	$s.\text{predicate.name} = \text{"cardinality"}$	"pred cardinality {#(" + s.\text{predicate.range} + ") = "+ s.\text{object} + "}"
8	$s.\text{predicate.name} = \text{"maxCardinality"}$	"pred maxCardinality {#(" + s.\text{predicate.range} + ") <= "+ s.\text{object} + "}"
9	$s.\text{predicate.name} = \text{"minCardinality"}$	"pred minCardinality {#(" + s.\text{predicate.range} + ") >= "+ s.\text{object} + "}"
10	$s.\text{predicate.name} = \text{"SymmetricProperty"}$	"pred SymmetricProperty { a \in "+ s.\text{predicate.domain} + " && b \in " + s.\text{predicate.range} + " / a, (" + s.\text{predicate} + ") = b ⇒ b, (" + s.\text{predicate} + ") = a)"
11	$s.\text{predicate.name} = \text{"FunctionalProperty"}$	"pred FunctionalProperty {#(" + s.\text{predicate.range} + ") = 1}
12	$s.\text{predicate.name} = \text{"InverseFunctionalProperty"}$	"pred InverseFunctionalProperty {#(" + s.\text{predicate.domain} + ") = 1}"
13	$s.\text{predicate.name} = \text{"allValuesFrom"}$	"pred allValuesFrom { "+s.\text{subject.range} + " in "+ s.\text{object} + "}"
14	$s.\text{predicate.name} = \text{"someValuesFrom"}$	"pred someValuesFrom {some r: "+s.\text{subject.range} + " / r in "+ s.\text{object} + "}"

Algorithm 2: 2Alloy(c, S, P)
Input: a Class c , a set S of Statements, a set P of Properties

Output: a signature σ

- $\leftarrow \text{"sig" } + c.\text{name};$
- IF $c.\text{subClassOf} \neq \emptyset$
- $\leftarrow \sigma + \text{"extends" } + c.\text{subClassOf.name};$
- $\leftarrow \sigma + \{";$
- FOR each $p \in P$
- IF $\text{domain}(p, S) = c.\text{name}$
- $\leftarrow \sigma + p.\text{name} + ":" + \text{range}(p, S);$
- $\leftarrow \sigma + "\}";$
- RETURN σ ;

The sub-algorithm domain (p, S) is used to obtain the domain of the property p . For a property p , if there is a statement s meet the requirement of lines 1~2, the $s.\text{object}$ is considered to be the domain of p . Otherwise,

the domain of p 's parent is also its domain. We use parent (property) to represent the parent property of the parameter property, so the domain of a property is calculated recursively.

There is another sub-algorithm range (p, S), which is similar to the domain (p, S) algorithm. The algorithm aims to get the range of the property p .

```
Algorithm 3: domain (p, S)
Input: A property p of P, Statement S
Output: the domain of p
1. IF  $\exists s \in S, s.\text{subject} = p$  and  $s.\text{predicate} = \text{"domain"}$ 
2.  $p.\text{domain} = s.\text{object.name};$ 
3. ELSE
4.  $p.\text{domain} = \text{domain}(\text{parent}(p), S);$ 
5. RETURN  $p.\text{domain};$ 
```

As previously stated, the Animal ontology is taken as an example to illustrate how these algorithms are used.

We use the outputs of Jena as the inputs. Then we will get an Alloy model as is shown in Fig.4. In the model, there are four signatures which are *Animal*, *Female*, *Male* and *Man* corresponding to four classes in the inputs. The paternity relations between them can be described as follow: Female and Male both extend from *Animal* signature, while *Man* extends from *Male* signature. There are three properties which are *hasParent*, *hasFather* and *hasChild*. They are also the ontology's properties. The *inverseOf* and *subPropertyOf* are two predicates in the model which are converted from the statements. The former is used to illustrate the *hasParent* is the inverse of the *hasChild*. The latter show that the range of *hasFather* is subset of the range of *hasParent*.

```
sig Animal{
    hasParent:Animal,
    hasFather:Male,
    hasChild:Animal
}
sig Female extends Animal{}
sig Male extends Animal{}
sig Man extends Male{}
pred inverseOf{hasParent~hasChild}
pred subPropertyOf{all a:Animal|a.hasFather in a.hasParent}
```

Figure4. The generated Alloy model of "Animal" ontology.

C. Verifying Ontologies with the Alloy Analyzer

Ontology reasoning is an important way to ensure the correctness of ontologies. The existing tools are mainly for the conceptual-level reasoning. Alloy, in addition to meeting the conceptual-level reasoning, can also be used as a reasoned for the instance-level. In order to prove the satisfiability of the ontology model, at least one instance of the model is needed. When Alloy Analyzer is used to analyze the model, if the model is right, it will give some instances which meet the model at random. Else, a counterexample will be given to prove that the model is inadequate.

There are several tasks to do in the reasoning, which are consistency checking, subsumption reasoning and implication relation checking. If an ontology is inconsistent, then any erroneous conclusion may be deduced by software agents. Subsumption reasoning can be described as follow: if a class C_1 is more general than

another one C_2 , it subsumes that C_1 can be contained by C_2 . The implication relation can be used to get some implicated conclusion. All these tasks (是指什么任务 What does it mean "all these tasks can be converted"?) can be converted into Alloy assertion. In the reasoning, Alloy Analyzer will determine whether these assertions are met. If all of these have been met, then Alloy Analyzer will give some instances. Otherwise, the assertions which have not been satisfied will be pointed out. And Alloy Analyzer will generate a counterexample. All these can be used to search the root of the problem. In the following, we will give an example to illustrate the process of ontology reasoning.

In Fig.5, there is a segment of an Animal ontology. In the document, there are three classes: *Animal*, *Male*, *Female* and *Woman*. *Male* and *Female* both extend from *Animal* and disjoint with each other. *Woman* is a subclass of *Female*. The *Animal* class has a property named *animalHasFather* whose range is *Male*. The *femaleHasFather* is another property of the ontology. Its domain is *Female*, while its range is *Male*. The property *femaleHasFather* is the subPropertyOf *animalHasFather*.

```
.....
<owl:Class rdf:ID="Male">
  <rdfs:subClassOf rdf:resource="Animal"/>
</owl:Class>
<owl:Class rdf:ID="Female">
  <rdfs:subClassOf rdf:resource="Animal"/>
  <owl:disjointWith rdf:resource="Male"/>
</owl:Class>
<owl:Class rdf:ID="Woman">
  <rdfs:subClassOf rdf:resource="Female"/>
</owl:Class>
.....
<owl:ObjectProperty rdf:ID="animalHasFather">
  <rdfs:domain rdf:resource="Animal"/>
  <rdfs:range rdf:resource="Male"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="femaleHasFather">
  <rdfs:subPropertyOf rdf:resource="animalHasFather"/>
  <rdfs:domain rdf:resource="Female"/>
  <rdfs:range rdf:resource="Woman"/>
</owl:ObjectProperty>
.....
```

Figure 5. An example of error ontology.

The Animal ontology document can be converted into a Alloy module using our method. In the module, there are four signatures, two relations and a predicate. Then Alloy Analyzer is used to check the module.

As is shown in Fig.6, the editor panel of the user interface contains the Alloy models, and if there are some errors, they will be highlighted during model compilation. In the Animal model, the keyword 'in' is highlighted. The message panel displays the results of analysis. It shows that the left type and right type of highlighted word are always disjoint. Because the range of *animalHasFather* is *Male* and the range of *femaleHasFather* is *Woman*. And there is a *subProperty* constraint between *femaleHasFather* and *animalHasFather*, which mean that the *Woman* is *subClassOf Male*. It is a conflict.



Figure 6. Checking result of Alloy Analyzer.

According to the error message, we can find that the range of *femaleHasFather* property is not appropriate. It is modified into *Male* and the model is checked again. The Alloy Analyzer gives an instance for the modified model as shown below.

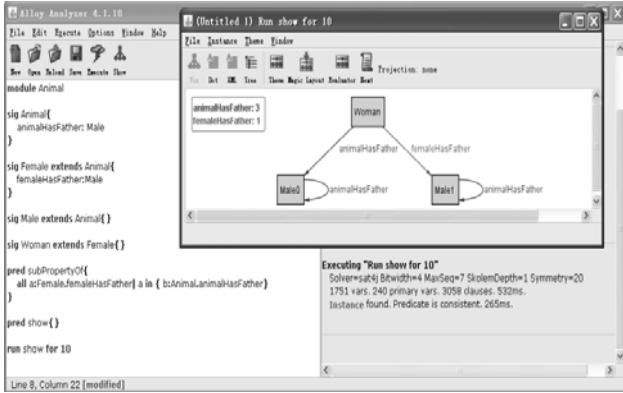


Figure 7. The modified model.

IV. RELATED WORK

Ontologies are set to play a key role in the Semantic Web by providing a source of shared and precisely defined terms that can be used in descriptions of web resources. Reasoning over such descriptions will be essential if web resources are more accessible to automated processes.

A number of ontology inference engines, such as FaCT, RACER, and FaCT++ have been developed with the advancement of ontology languages to facilitate ontology creation, management, verification, merging, etc.. They can make explicit information from knowledge and data. But complex ontology-related properties cannot be supported by them. FaCT (Fast Classification of Terminologies) is a DL classifier that can also be used for modal logic satisfiability testing. The FaCT system includes two reasoners, one for the logic SHF (ALC augmented with transitive roles, functional roles and a role hierarchy) and the other for the logic SHIQ (SHF augmented with inverse roles and qualified number restrictions), both of which use sound and complete tableau algorithms. The RACER system is a knowledge representation system that implements a highly optimized tableau calculus for a very expressive description logic. It offers reasoning services for multiple TBoxes and for

multiple ABoxes as well. The system implements the description logic ALLQHIL₊ also known as SHIQ. The DL reasoned FaCT++ implements a tableau decision procedure for the well known SHIQ description logic, with additional support for datatypes, including strings and integers. The system employs a wide range of performance enhancing optimizations, including both standard techniques and newly developed ones.

Recently, some researchers have proposed some methods to do the ontology reasoning tasks, article [17] provides a rigorous treatment of data type predicates on the concrete domain. It investigated the complexity of combined reasoning with description logics and concrete domains, and extended ALL (D), which is the basic description logic for reasoning with concrete domain, by the operators “feature agreement” and “feature disagreement”. Jeff Z. Pan [18] proposed a flexible reasoning architecture for Semantic Web ontology languages and described the prototype implementation of the reasoning architecture, based on the well-known FaCT DL reasoned. It allows users to define their own data types and data type predicates based on built-in ones and new data type reasoners can be added into the architecture without having to change the concept reasoned. Article [4] present OWL Flight, which is loosely based on OWL, but the semantics is grounded in Logic Programming rather than Description Logics, and it borrows the constraint-based modeling style common in databases. And the ontology reasoning tasks are supported by OWL DL and OWL Flight.

Similarly to our approach, article [9] proposed a new novel application domain for Alloy firstly. It believed that software engineering techniques and tools can provide automatic reasoning and consistency checking services for Semantic Web. The software modeling languages Z and its proof tool Z/EVES can also be used to verify DAML+OIL [8]. And in the article [7], a combined approach is proposed to checking Web ontologies. It used the software engineering techniques and tools, i.e., Z/EVES and Alloy Analyzer, to complement the ontology tools for checking Semantic Web documents.

We take advantage of Alloy and Alloy Analyzer to convert the OWL ontologies into Alloy model, and then using the Alloy Analyzer to automatically check and reason the generated model.

Compared with the article[23], our method can handle complex-property ontology efficiently. In the article[23], the ontology classes and properties are translated into different signatures in Alloy model, and then predicates are used to establish the relationship between them. Our approach uses the features of Alloy to convert the ontology properties to relations of Alloy model directly, for in Alloy everything is a relation. For the Alloy model, there are striking contrasts between the two methods with different scopes. In the Figure 8, there is a big difference in numbers in the analysis of the same ontology. The horizontal axis represents the Alloy model scope, while the vertical axis represents the number of variables. You can see that the greater the scope given is, the bigger the difference of the number of variables is, which is to say

that the more the number of ontology instances is, the more obvious the advantage of this method is.

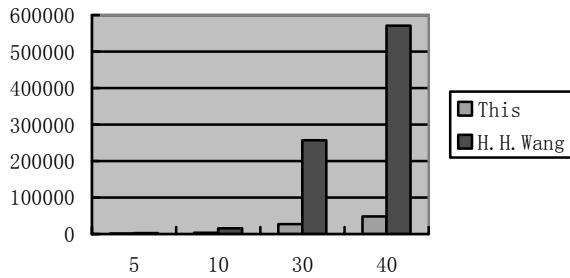


Figure 8. The variables of a same model in different methods.

The Figure 9 is similar to the previous figure, the horizontal axis represents the Alloy model scope, while the vertical axis represents the number of clauses.

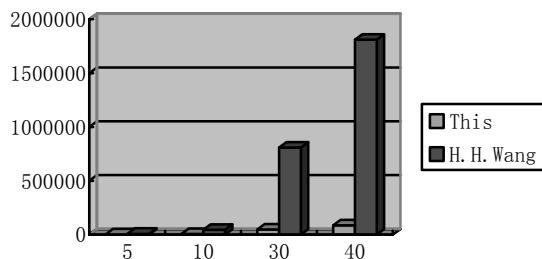


Figure 9. The clauses of a same model in different methods.

The Figure 10 is to illustrate the difference in execution time (ms).

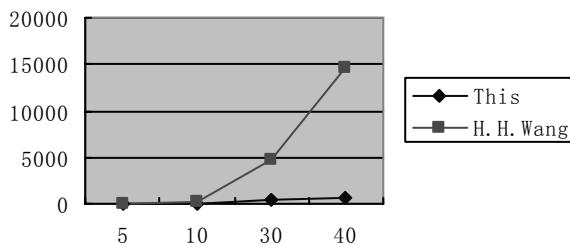


Figure 10. The time of a same model in different methods.

V. CONCLUSION

This paper presented a reasoning method for OWL ontologies. Compared to the existing reasoners for OWL, we propose a method using software engineering tools. The paper is based on the conversion mode. It can meet the complex-properties ontology reasoning. And better than the existing reasoners, it can give reasoning services in the instance level. Further, the reasoning of Alloy Analyzer can not only prove the ontology is wrong, generate a set of ontology instances, counterexample on predicates, but also can offer help to point out where the errors are.

However, this method also has its shortcomings, it can't be carried out automatically, it requires human involvement.

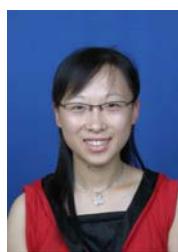
VI. ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China (61175056), Young Key Teachers Foundation Projects of Dalian Maritime University (2011QN028) and the Open Project Program of Artificial Intelligence Key Laboratory of Sichuan Province, (Sichuan University of Science and Engineering), China (2011RYJ03).

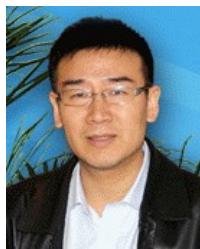
REFERENCES

- [1] B. FRANZ, C. DIEGO, L.M. DEBORAH, N. DANIELE AND F.P.-S. PETER Eds. *The description logic handbook: theory, implementation, and applications*, Cambridge University Press, 545. 2003.
- [2] A. Ankolekar, M. Burstein, J. Hobbs, O. Lassila, D. Martin, S. McIlraith, S. Narayanan, M. Paolucci, T. Payne, K. Sycara, and H. Zeng. "DAML-S: Semantic markup for Web services". In Proceedings of the International Semantic Web Working Symposium (SWWS), pages 411-430, 2001.
- [3] T. Berners-Lee, J. A. Hendler and O. Lassila, "The Semantic Web". Scientific American. v284 iMay (5). 34-43.
- [4] J. De Bruijn, R. Lara, A. Polleres, and D. Fensel, "OWL DL vs. OWL flight: conceptual modeling and reasoning for the semantic Web". In Proceedings of the Proceedings of the 14th international conference on World Wide Web, Chiba, Japan 2005 ACM, 1060836, 623-632.
- [5] J. J. Carroll, I. Dickinson, C. Dollin, D. Reynolds, A. Seavorne AND K. Wilkinson, "Jena: implementing the semantic web recommendations". In Proceedings of the Proceedings of the 13th international World Wide Web conference on Alternate track papers \& posters, New York, NY, USA2004 ACM, 1013381, 74-83.
- [6] R. S. Cost, T. Finin, A. Joshi, Y. Peng, C. Nicholas, I. Soboroff, H. Chen, L. Kagal, F. Perich, Y. Zou and S. Tolia. "ITtalks: A Case Study in the Semantic Web and DAML+OIL". IEEE Intelligent Systems 17, 40-47.
- [7] J. S. Dong, C. H. Lee, H. B. Lee, Y. F. Li and H. Wang. "A combined approach to checking web ontologies". In Proceedings of the Proceedings of the 13th international conference on World Wide Web, New York, NY, USA 2004 ACM, 988770, 714-722.
- [8] J. S. Dong, C. H. Lee, H. B. Lee, Y. F. Li and H. Wang. "Verifying DAML+OIL and Beyond in Z/EVES". In Proceedings of the Proceedings of the 26th International Conference on Software Engineering2004 IEEE Computer Society, 999425, 201-210.
- [9] J. S. Dong, J. Sun and H. Wang. "Checking and reasoning about Semantic Web through alloy". In Fme 2003: Formal Methods, Proceedings, K. Araki, S. Gnesi and D. Mandrioli Eds. Springer-Verlag Berlin, Berlin, 796-813.
- [10] D. Fensel, I. Horrocks, F. Van Harmelen, S. Decker, M. Erdmann and M. Klein. "OIL in a nutshell". In Knowledge Engineering and Knowledge Management, Proceedings - Methods, Models, and Tools, R. DIENG AND O. CORBY Eds. Springer-Verlag Berlin, Berlin, 1-16. 2000.
- [11] T. R. Fruber. "Toward principles for the design of ontologies used for knowledge sharing". Int. J. Hum. Comput. Stud. 43, 907-928. 1995.

- [12] I. Horrocks, "The FaCT System" In Proceedings of the Proceedings of the International Conference on Automated Reasoning with Analytic Tableaux and Related Methods, vol. LNAI 1397, pp. 307-312. 1998.
- [13] I. Horrocks, D. Fensel, J. Boriekstra, S. Decker, M. Erdmann, C. Goble et al. The Ontology Inference Layer OIL, Tech. Report, Vrije Universiteit, Amsterdam. Retrieved March 19, 2002.
- [14] I. Horrocks, P. F. Patel-Schneider, S. Bechhofer and D. Tsarkov."OWL rules: A proposal and prototype implementation". Journal of Web Semantics 3, 23-40. 2005.
- [15] D. Jackson. "Alloy: a lightweight object modelling notation". ACM Trans. Softw. Eng. Methodol. 11, 256-290. 2002.
- [16] D. Jackson, I. Schechter and H. Shlyacter. "Alcoa: the alloy constraint analyzer". In Proceedings of the Proceedings of the 22nd international conference on Software engineering, Limerick, Ireland2000 ACM, 337616, 730-733. 2000.
- [17] C. Lutz. "The Complexity of Reasoning with Concrete Domains Revised Version". Rheinisci-Westfälische Technische Hochschule. 1999.
- [18] J. Z. Pan. "A flexible ontology reasoning architecture for the Semantic Web". Ieee Transactions on Knowledge and Data Engineering 19, 246-260. 2007.
- [19] E. Sirin, B. Parsia, B. C. Grau, A. Kalyanpur and Y. Katz. "Pellet: A practical OWL-DL reasoner". Journal of Web Semantics 5, 51-53. 2007.
- [20] Y. Song and R. Chen. "Non-standard Reasoning Services for the Verification of DAML+OIL Ontologies". In Artificial Intelligence Applications and Innovations, H. PAPADOPOULOS, A. ANDREOU AND M. BRAMER Eds. Springer Boston, 203-210-210. 2010.
- [21] D. Tsarkov and I. Horrocks. "FaCT++ Description Logic reasoner: System description". In Automated Reasoning, Proceedings, U. FURBACH AND N. SHANKAR Eds. Springer-Verlag Berlin, Berlin, 292-297. 2006.
- [22] H. Volker, M. Ralf and A. Turhan. "RACER User's Guide and Reference Manual Universitaet Hamburg". 1999.
- [23] H. Wang, J. Dong, and J. Sun. "Reasoning support for Semantic Web ontology family languages using Alloy". Multiagent and Grid Systems 2. 2006.
- [24] A. Alnusair and T. Zhao. "Using Ontology Reasoning for Reverse Engineering Design Patterns". Models in Software Engineering, 344–358. 2010.
- [25] R. Umeton, B. Yankama, G. Nicosia and C.F. Dewey. "OREMP : Ontology Reasoning Engine for Molecular Pathways Use Case 1 : Model Alignment for Parallel Simulation with Cytosolve Use Case 2 : Aligned Ontologies". Lloydia Cincinnati, 2122-2122. 2009.
- [26] I. Conforence and C. Academy. "Decomposition Model Building and Ontology Reasoning Towards G-Layer of RGPS Requirement Meta-Model". Computer, 47-51. 2010.



Yingjie Song born in ShanDong province, China, in September 1983. She earned master degree of Computer Science and Technology from the school of information science and technology at the Dalian Maritime University, Dalian, China in 2009, and now study in the school of information science and technology at the Dalian Maritime University as a Doctoral Candidate. His research interests are in Semantic Web, Ontology Evolution and Ontology reasoning, etc.



Rong Chen received the Ph.D. degree in computer software and theory from the Jilin University, China, in 2000. He is currently a member of the school of information science and technology at the Dalian Maritime University, China. He has previously held position at Sun Yat-sen University. His research interests are in system reliability and automated software engineering joined the faculty in 2005 having previously worked at the Sun Yat-sen University and the Graz University of Technology. Currently, he research focuses on trusted software, Internet and mobile computing, semantic web, embedded software engineering, database and knowledge base, and intelligent diagnosis.



Yaqing Liu is a lecturer in the School of Information Science and Technology at Dalian Maritime University, Dalian, China. He received PhD degrees in Computer Application Technology in 2011. His research interests are in Semantic Web, Ontology Evolution and Ontology reasoning, etc. He has authored over 10+ refereed journal/conference papers and about 80% papers is indexed by SCI/EI.

Process Goose Queue Methodologies with Applications in Plant-wide Process Optimization

Jingwen Huang

College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China
Email: huangjw@mail.buct.edu.cn

Hongguang Li

College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China
Email: lihg@mail.buct.edu.cn

Abstract— Inspired by biologic nature of flying wild geese, a so-called process goose queue (PGQ) technique oriented for plant-wide optimization is established. Taking advantage of this ad-hoc structure of flying geese, a plant-wide process can be decomposed into several hierarchically connected PGQs along the direction of the objective function generation. In line with this thought, plant-wide process optimization is accordingly identical with the following and tracking issues between leading and following geese. Followed by this philosophy, related theoretical definitions and modeling principles together with enabling algorithms are explicitly introduced. With the characteristics of evolutionary optimization, PGQ approach is able to overcome the algorithmic deficiencies associated with conventional optimizations. To demonstrate the feasibility and validity of the contributions, TE process is employed as the case study.

Index Terms—plant-wide process, process optimization, process goose queue (PGQ)

I. INTRODUCTION

Traditionally, plant-wide process optimization approaches can be classified into two relatively distinct categories in terms of architecture: global or centralized architecture and decentralized architecture. The global approaches associate overall processes with economic objectives and optimize them based on rigorous models. Therein, the main impacts on optimization performance arise from model complexity and nonlinearity, as well as heavy computational burden due to enormous manipulated variables involved. Even though applications of flow-sheet simulation tools such as DMCplus, CLP and RTO of AspenTech, as well as Profit Optimizer and Profit Max of Honeywell are increasingly extensive, it is acknowledged that the corresponding optimization solutions suffer both far from analytics and hard to understand. Alternatively, the decentralized approaches decompose large-scale optimization problems into

several sub-systems mutually coordinated. Darby and White [1] exemplified that decentralized architectures could achieve the same performances with those of the global ones. Taking into account of physical structures and coupling factors among the subsystems, Sobieski[2] presented a generalized multilevel optimization approach named multidisciplinary design optimization (MDO), which is concerned with complex systems exhibiting challenges with three typical MDO architectures subsequently exploited, including concurrent subspace optimization (CSSO), collaborative optimization (CO) and Bi-Level Integrated System Synthesis (BLISS). MDO methods are widely applied to non-process industries. For example, Duddeck & Fabian[3] applied MDO to control system designs for car bodies and Silva & Valceres [4] employed MDO to those for gas turbine engines. However, most of existing researches highlight mechanical structures but rarely deal with interconnection characteristics of process variables. Another research issue of decentralized optimization focuses on hierarchical multi-objective optimization together with multi-layer optimization algorithm, in which the problems are firstly decomposed into a multi-system and subsequently dealt with using multi-objective programming, as addressed in references [5],[6]. However, owing to the demands for separable or approximately separable objectives, the above-cited method is apt to cause considerable systematic deviations in the presence of severe nonlinear relations between the global objective and sub-objectives. A popular approach at present is to recognize the weights of sub-objectives which can then be used to approximate the global objective. However, in the presence of severe nonlinear relationship between global objectives and sub-objectives, the solutions derived from this method can be far away from the actual optima.

Plant-wide industrial processes are actually connections of a variety of basic operational units which are considered as dynamic systems constituted by output process variables and input process variables. Inspired by

Corresponding author: Hongguang Li; Tel.: 86-10-64434797; Fax: 86-10-64442932.

the philosophy of flying goose queue, a certain kinds of operational units could be regarded as a goose so that the plant-wide process could be identified as a goose queue. The principle of pursuing optimal set points in plant-wide process is similar to the mechanism of flocks of geese self-organizing into V-formation. In line with this thought, we came up with a novel idea of “process goose queue (PGQ)” [7] to reconfigure plant-wide processes, thereby optimizing them in terms of economic perspectives. With the PGQ methodology, optimization model can be decomposed organically according to goose queue formation. The transformation of the subsystem optimum objective value can refer to the transfer mechanism of optimal position of through upwash models. Thus, the optimum operating points can be achieved based on the pursuing principle among geese in V-formation. As for inseparable objectives, the PGQ approaches could enjoy theoretically lossless decomposition to achieve decentralized optimization schemes.

The remainder of this paper is organized as follows. Section 2 briefly discusses the mechanism of flocks of geese self-organizing into V-formation. Section 3 proposes fundamental definitions and adjusting rules of PGQ along with an illustrative example. In Section 4 plant-wide process optimization problems are formulated based on multi-layer PGQ metrics and enabling algorithms of multi-layer PGQ for plant-wide process optimization are introduced. In Section 5, TE process is employed as a case study for exemplifying the applications. Section 6 concludes the contribution and assesses the future prospects.

II. SELF-ORGANIZATION V-FORMATION OF FLOCK OF GEESE

Goose queue refers to a flock of flying wild geese lined an instinctively V-shaped formation in mass migration, whose principal benefit lies in the increased flying efficiency, as shown in Fig.1. It is reported that geese in a V-formation may conserve 12–20 % of the energy they would need to fly alone [8], [9], [10]. A flying goose could generate an upward pressure known as upwash beneficial for a following goose maintaining its altitude and save energy. As a result, leading goose serves as the leader of the queue while following geese are responsible for following and keeping the V-formation. Cattivelli [11], [12] focused on the self-organizing V-formation of flocks of geese, thereby using a model, $f(x,y)$, to describe the upwash generated by a flying goose. Assuming that the wingspans of all gooses are constant and the upwash functions are convex, an optimum position (x_{opt}, y_{opt}) is available which could maximize the upwash as shown in Fig. 2. Every goose located at position (x_k, y_k) in the V-formation experiences the overlap upwash through $\sum_{l=1}^N f(x_k - x_l, y_k - y_l)$, by which the optimum position (x_{opt}, y_{opt}) could be achieved. The

underlying point behind this mechanism lies in that geese could measure the upwash and communicate with their neighbors.

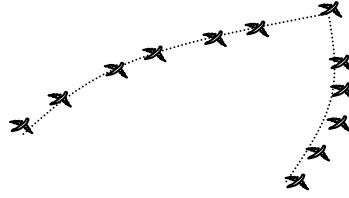


Figure 1 V-formation of flying geese.

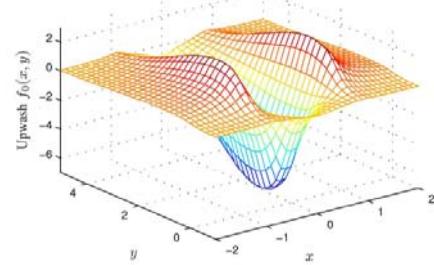


Figure 2 Upwash generated by a goose.

In simulation, Fig.3 [11] shows the resulting goose formations at different time instants, where the goose flock converges to a V-shape formation through 500 iterations or so. Goose located at position (x_k, y_k) measures the upwash with respect to the reference geese so that to pursue an optimum position (x_{opt}, y_{opt}) . After that a new estimate of the best relative position with other geese is achieved.

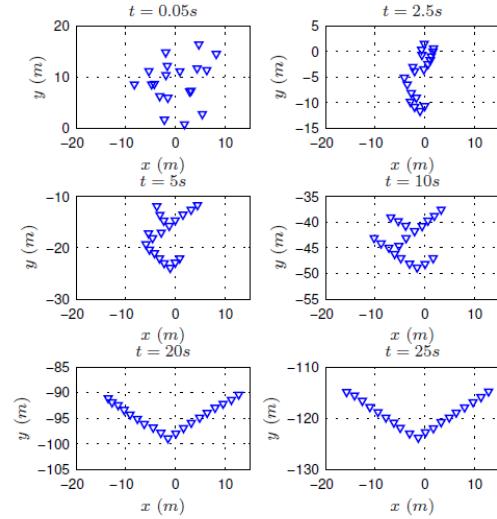


Figure 3 Goose positions at different time instants.

A steady-state V-formation is shown in Fig. 4, where the red dots indicate the positions of the geese. Notice that every goose flies in such a way that the generated upwash overlaps with the upwash from its leading goose.

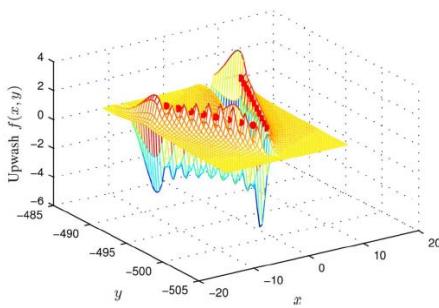


Figure 4. Upwash generated by geese in steady-state.

III. PROCESS GOOSE QUEUE

Generally, steady-state relationship among the basic operational units of a plant-wide process can be described by following equation:

$$Y = g(S, X) \quad (1)$$

where, Y, S, X and g indicate sets of the output state variables, input state variables, manipulated variables and steady-state relationship functions, respectively. Instead, we accordingly propose PGQ approaches with following descriptions.

Definition1 (PGQ)

A Process Goose Queue (PGQ) is a 4-tuple, $\text{PGQ} = (L, F_S, F_M, A)$, where,

- L is the process leading goose (PLG) such that $\{L | L \subset Y\} \neq \emptyset$, represented as (L) ;
- F_S is the supervised following goose (SFG) such that $F_S \subset S$ represented as (F_S) ;
- F_M is the manipulated following goose (MFG) such that $\{F_M | F_M \subset X\} \neq \emptyset$, represented as (F_M) ;
- A is the information arc (IA) such that $\{A | A \subseteq (L \times (F_S \cup F_M)) \cup (A \subset g)\} \neq \emptyset$, represented as inverted V;
- $F_S \cap F_M = \emptyset$, $L \cap F_M = \emptyset$, $L \cap F_S = \emptyset$, $L = A(F_S, F_M)$

The graphical description of a PGQ is illustrated in Fig.5, where, the PLG (L), SFG (F_S) and MFG (F_M) represent output process variables, Y , input process state variables, S , and input process manipulated variable, X , respectively. A , corresponds to the process steady-state models, g .

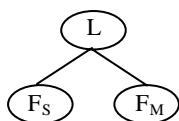


Figure 5. A PGQ (process goose queue)

Let's take a look at an example of the Williams-Otto reactor [13] which is a jacketed CSTR as shown in Fig.6.

It is operating at a temperature T_r as well as reactant flows F_a and F_b , where a six-component product, Z , is produced.

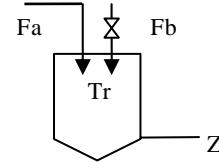


Figure 6. Williams-Otto reactor

Based on PGQ techniques, the output state variables Z , input state variables F_a , manipulated variables F_b and T_r are equivalent to PLG, SFG, MFG, respectively, as shown in Fig.7, where a steady-state model between the output and inputs corresponds to IA of the PGQ.

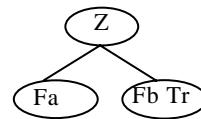


Figure 7. A PGQ for Williams-Otto reactor

Definition 2(Optimum operating states)

Optimum operating states refer to an ideal PGQ configuration in which a PLG operating at an ideal trajectory is followed by an optimum V-formation constituted by a SFG and a MFG, as described by:

$$L^* = A(F_S^*, F_M^*) \quad (2)$$

In practical industrial processes, optimum operating state could be destroyed by uncertain disturbances, which is in desperate need of adjustment to recover. Motivated by this idea, two alternative adjustment rules associated with a PGQ are specified as follows.

Rule 1 (PLG driven adjustment)

Once a PGQ operates away from its normal trajectory, PLG would try to adjust its position autonomously back to an ideal one. At the same time, SFG and MFG would operate in consistent with the activities of PLG, formulating an adapted V-formation. This kind of adjustment implies solving the following optimization problems.

$$\begin{aligned} & \min_{F_M, F_S} (L^* - L)^2 \\ \text{s.t. } & L = A(F_S, F_M) \\ & F_{SL} \leq F_S \leq F_{SU} \\ & F_{ML} \leq F_M \leq F_{MU} \end{aligned} \quad (3)$$

Rule 2 (SFG driven adjustment)

Once the V-formation of PGQ deviates from an optimum one due to SFG failing to follow it, MFG would try to adjust its position autonomously to formulate a new

optimum formation. At the same time, PLG would slightly shift its position to survive the adjustment of MFG. This kind of adjustment implies solving the following optimization problems.

$$\begin{aligned} & \min_{F_M} (L^* - L)^2 \\ \text{s.t. } & L = A(F_S^*, F_M) \\ & F_{ML} \leq F_M \leq F_{MU} \end{aligned} \quad (4)$$

Referring back to the above-mentioned example, if a new target of any component of z is demanded, the rule 1 would be launched to implement the adjustment; if $F_S = [Fa]$ deviates from the optimum operating state, the rule 2 would be triggered to implement the adjustment.

IV. PROCESS OPTIMIZATION

A. Plant-wide PGQ for Process Optimization

In order to cope with plant-wide processes, a multi-layer PGQ structure should be established additionally.

Definition 3 (Multi-layer PGQ)

A multi-layer PGQ consists of several PGQs which are organized in a hierarchical architecture. Therein, the PGQs are characterized by $PGQ_i = (L_i, F_{S_i}, F_{M_i}, A_j)$, where, $i=1, \dots, m$, indicates the depth index, the SFG of an upper PGQ may serve as the PLG of the neighbored lower PGQ in terms of increased depth index. The graphical description of a multi-layer PGQ is exemplified in Fig.8.

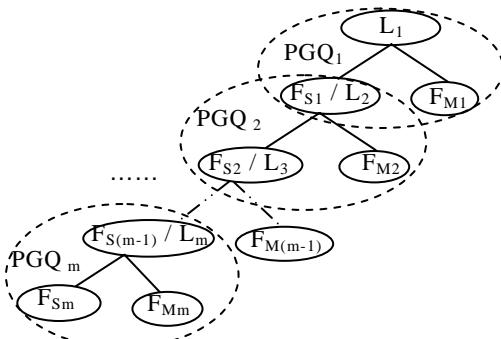


Figure 8. A multi-level PGQ

In fact, process optimization could be identified as a procedure of adjusting manipulated variables to minimize or maximize economic goals subjecting to the constraints of process models. What's more, the fact that actual formations of process models involved in a plant-wide process optimization depend on the connections of operational units accounts for particular hierarchical decompositions of process models, as shown in Fig.9.

Specifically, the economic objective function of a plant-wide optimization problem can be expressed in terms of direct related process state variables and manipulated variables. Accordingly, an additional concept about the objectives of a multi-layer PGQ is presented as

follows.

Definition 4 (PGQ-Objective) :

A PGQ-Objective is equivalent to an economic objective function of a plant-wide process, characterized by

$$P = \min \varphi(P_{S1}, P_{S2}, \dots, P_{Sj}, P_M) \quad (j = 1, 2, \dots, n) \quad (5)$$

where, P_{Sj} and P_M are process state variables and manipulated variables respectively.

Referring back to definition 1, P_{Sj} and P_M could be similarly considered as PLG and MFG of a PGQ, respectively. The graphical descriptions of a PGQ-Objective are shown in Fig.10. In this context, the procedures towards establishing a plant-wide PGQ for process optimization are summarized as follows.

(1) A plant-wide process is decomposed into several operational units / areas corresponding to the PGQs using sequential modular approaches. In the presence of a tree-structural plant-wide process, the multi-layer PGQs are consistent with the connections of the process operational areas. Otherwise, additional modeling treatments such as block segmentation, staggered breaks, and convergence calculation should be carried out before the multi-layer PGQs are obtained.

(2) Construct the economic objective functions with respect to the related process state variables P_{Sj} which serve as PLG (L_{ij}) and manipulated variables P_M .

(3) Connect each P_{Sj} (L_{ij}) with a multi-layer PGQ. Thus, a plant-wide PGQ could be realized, whose exemplary graphical description is shown in Fig.11.

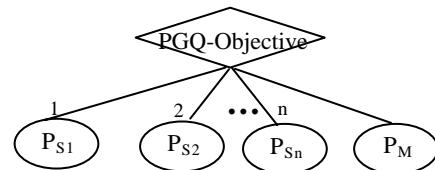


Figure 10. A PGQ (process goose queue)-Objective

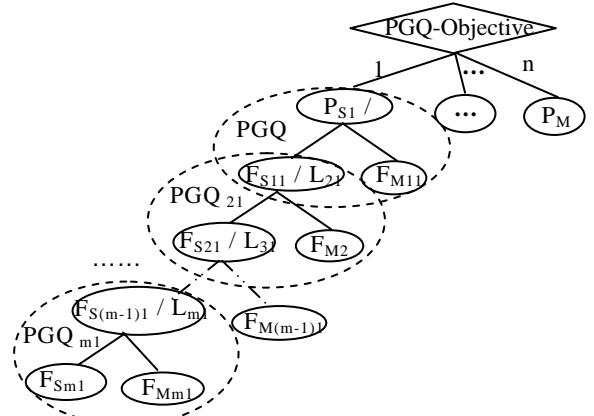


Figure 11. A plant-wide PGQ

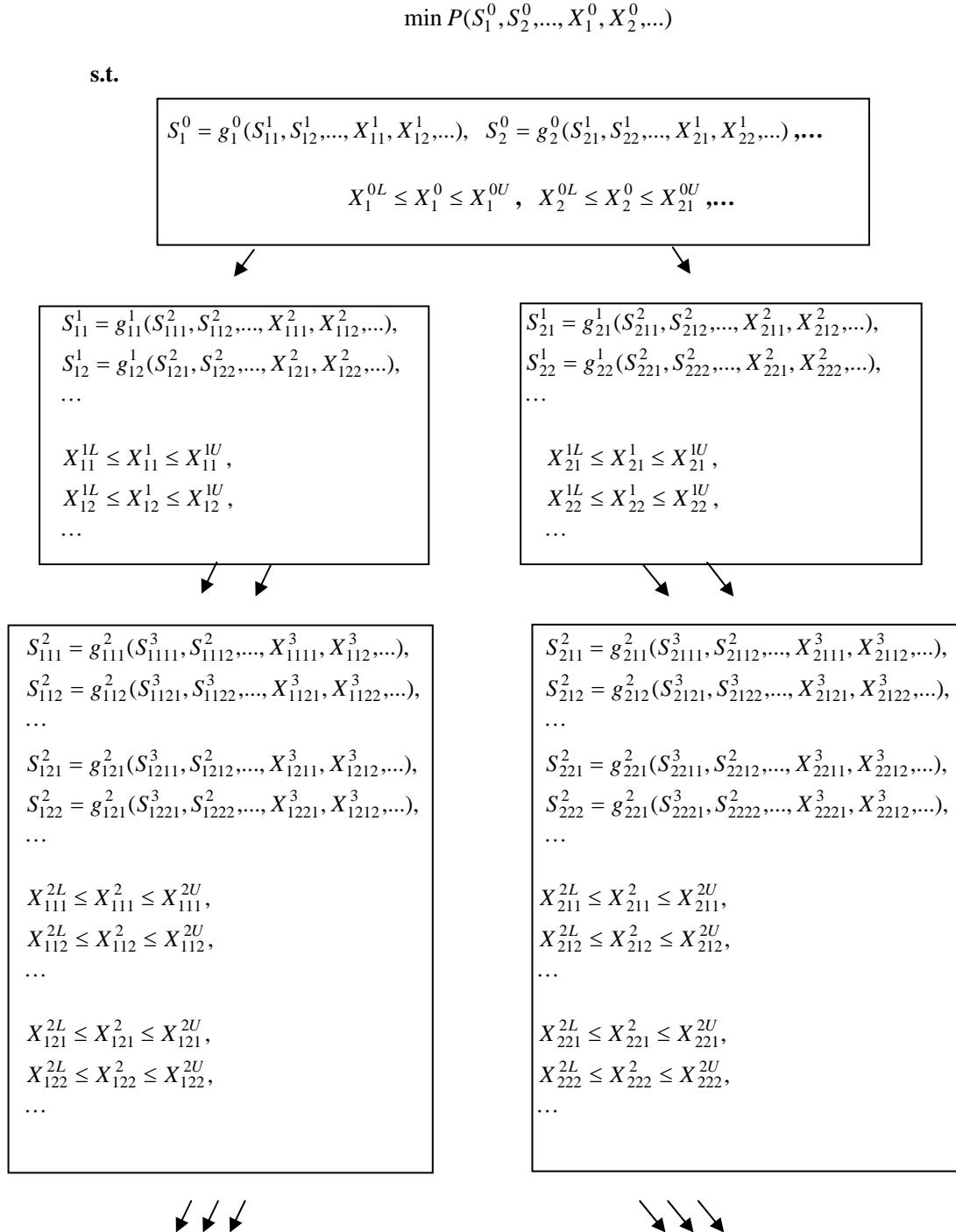


Figure 9 Decompositions of a plant-wide process optimization problem

B. Implement of Plant-wide PGQ Optimization Algorithms

According to the fundamental description of the PGQ, plant-wide process optimization is equivalent to implementing PLG autonomous adjustment rule in corresponding multi-layer PGQs. Therefore, solving a plant-wide optimization problem can be converted into implementing multi-layer PGQ optimization

algorithms which involves three kinds of interconnected tasks, such as assignment of a PGQ-Objective, configuration of the multi-layer PGQ formations and achievement of a PGQ-Objective, described as follows.

(1) Assignment of a PGQ-Objective: A PGQ-Objective could be achieved by means of assignment of optimum points of process state variables

P_{Sj} ($j=1,2,\dots,n$) and manipulated variable set P_M involved, which could be implemented by applying rule 1 (PLG driven adjustment), i.e. solving the following optimization problem.

$$\begin{aligned} \min_{P_{S_1}, \dots, P_{S_n}, P_M} & \varphi(P_{S_1}, P_{S_2}, \dots, P_{S_n}, P_M) \\ \text{s.t. } & P_{S_{jL}} \leq P_{Sj} \leq P_{S_{jU}} \quad (j = 1, 2, \dots, n) \\ & P_{M_{iL}} \leq P_M \leq P_{M_{iU}} \end{aligned} \quad (6)$$

(2) *Configuration of the multi-layer PGQ formations:* It is supposed that there is a couple of multi-layer PGQ associated with a PGQ-Objective for a practical problem. For simplicity's sake, we only consider the treatment of one multi-layer PGQ. Starting off with $i=1$, rule 1 (PLG driven adjustment) is stepwise applied to PGQ_i at an increasing index i , which implies solving the following optimization problem (Noting $F_{s0}=P_s$).

$$\begin{aligned} \min_{F_{S_i}, F_{M_i}} & (F_{S_{i-1}}^* - L_i)^2 \\ \text{s.t. } & L_i = A_i(F_{S_i}, F_{M_i}) \\ & F_{S_{iL}} \leq F_{S_i} \leq F_{S_{iU}} \\ & F_{M_{iL}} \leq F_{M_i} \leq F_{M_{iU}} \quad (i = 1, 2, \dots, m) \end{aligned} \quad (7)$$

(3) *Achievement of a PGQ-Objective:* From the PGQ with the largest depth index to the PGQ-Objective, we should update the achievement of PLG associated with each PGQ with respect to the optimum solutions (process variables) configured in step (2) until that of the PGQ-Objective.

V. CASE STUDIES

Since TE process (Tennessee Eastman Process) [14], shown in Fig.12, was proposed by Downs and Vogel (1993), it has been widely circulated in the literature as a case study due to attractive challenging properties. TE process involves five major operational units including a two-phase reactor, a partial condenser, a separator, a stripper and a compressor, in which two products are created from four reactants, an inert component B and a byproduct F, denoted by a total of eight components, A, B, C, D, E, F, G, and H instead.

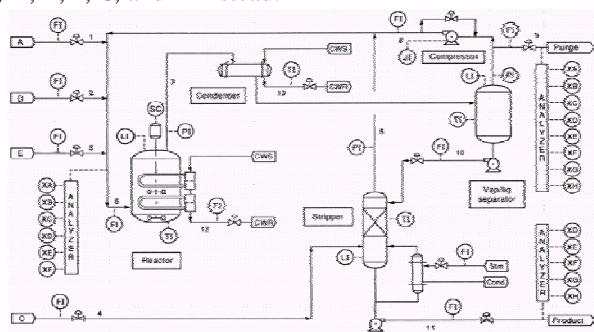


Figure 12 TE process

There are 12 manipulated variables and 41 state

variables involved in the process. Specifically, the manipulated variable vector F_M contains 10 variables, $F_M = [F_1, F_2, F_3, F_4, F_8, F_9, F_{10}, F_{11}, T_{cr}, T_{cs}]$, where, F_i is the molar flow rate of stream i [kmol/h] ($i=1,2,\dots,11$), T_{cr} and T_{cs} are temperatures of the reactor and separator. Here, the objective function corresponds to the hourly operating cost (C_{tot}) in \$/h which is aimed to be minimized. Therein, the reactant and product losses in terms of the purge and product streams, steam cost and compressor power cost and is measured by (8).

$$\begin{aligned} C_{tot} = & F 9 \sum_{\substack{i=1 \\ i=A \\ i \neq B}}^H C_{i,cst} X_{i,9} + F 11 \sum_{i=D}^F C_{i,cst} X_{i,11} \\ & + 0.0536 W_{cmp} + 0.0318 F_{steam} \end{aligned} \quad (8)$$

Based on the steady-state first-principle models, TE process can be developed into a plant-wide PGQ which includes two multi-layer PGQs as shown in Fig.13, in which the corresponding notations are listed in Table I. As a result, the following steps are implemented to solve the plant-wide PGQ optimization problem.

TABLE I.

CORRESPONDING PGQ NOTATIONS FOR TE PROCESS

PGQ	SFG/PLG	MFG
PGQ-Objective	$P_{S1}/L_{11} = [X_{i,9}]$ $P_{S2}/L_{12} = [X_{i,11}]$	$P_{M3} = [F_9 \ F_{11}]$
PGQ ₁₁	$F_{S11} / L_{21} = [P_s \ T_s]$	$F_{M11} = [F_{10} \ F_8]$
PGQ ₂₁	$F_{S21} / L_{31} = [T_R]$	$F_{M21} = [T_{CS}]$
PGQ ₃₁	$F_{S31} / L_{41} = [F_R \ P_R \ L_R]$	$F_{M31} = [F_{12} \ T_{CR}]$
PGQ ₄₁	0	$F_{M41} = [F_1 \ F_2 \ F_3]$
PGQ ₁₂	$F_{S11} / L_{21} = [T_{str} \ P_{str}]$	$F_{M12} = [F_4]$
PGQ ₂₂	$F_{S11} / L_{21} = [P_s, T_s]$	$F_{M22} = [F_{10}]$
PGQ ₃₂	$F_{S21} / L_{31} = [T_R]$	$F_{M32} = [T_{CS}]$
PGQ ₄₂	$F_{S31} / L_{41} = [F_R \ P_R \ L_R]$	$F_{M42} = [T_{CR}]$
PGQ ₅₂	0	$F_{M52} = [F_1 \ F_2 \ F_3]$

(1) *Assignment of the PGQ-Objective:* Referring to the optimization scheme carried out by Ricker [15], compressor and steamer are specified at the OFF positions. The state variables $x_{i,9}$ and $x_{i,11}$ and the manipulated variables F_9 and F_{11} involved in the objective function are selected as P_S and P_M , respectively. Table II presents assignment of the optimum points associated with the problem.

(2) *Configuration of the multi-layer PGQ formations:* There are two multi-layer PGQs involved in this case. The optimum points L_{11}^* and L_{12}^* responsible for the PGQ-Objective are followed and tracked by PGQ₁₁ and PGQ₁₂, attaining $F_{S11}^* / L_{21}^* = [P_s=2700, T_s=92]$, $F_{M11}=[F_{10}=37.2, F_8=0]$, $F_{S12}^* / L_{22}^* = [P_{str}=3330, T_{str}=66.60]$, $F_{M12}=[F_4=60.9]$, where F_{S11}^* and F_{S12}^* serves as L_{21}^* of PGQ₁₁ and L_{21}^* of PGQ₂₁ respectively. Similarly, the rest PGQs are treated along with a minimum objective value 116\$/h expected, as listed in Table II.

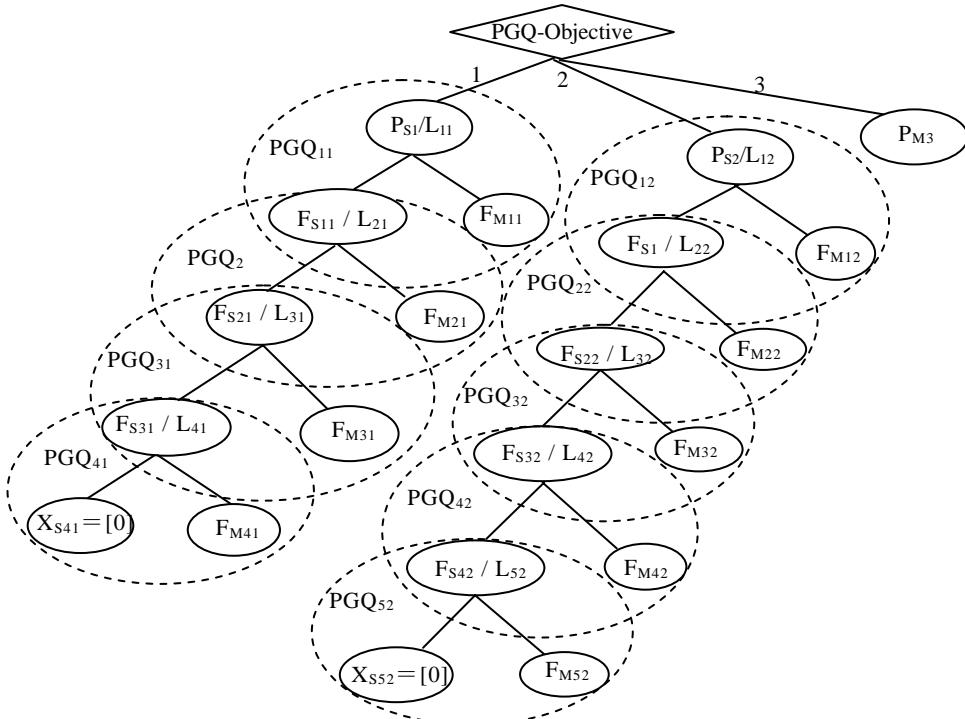


Figure 13. Plant-wide PGQ for TE process

TABLE II.

ASSIGNMENT OF THE OPTIMUM POINTS

Name of PGQ	SFG/ PLG	MFG
PGQ -Objective	$P_{S1}/L^{*}_{11} = [X_{i,9(i=A,C,D,E,F,G,H)} = 32.63 \quad 13.29 \quad 0.89 \quad 16.13 \quad 6.675 \quad 3.265]$ $P_{S2}/L^{*}_{12} = X_{i,11(i=D,E,F)} = [0.011 \quad 0.575 \quad 0.186]$	$P_{M3} = [F_9 = 24.2 \quad F_{11} = 46.4]$
PGQ ₁₁	$F^{*}_{S11}/L^{*}_{21} = [P_s = 2700, T_s = 92]$	$F_{M11} = [F_{10} = 37.2 \quad F_8 = 0]$
PGQ ₂₁	$F^{*}_{S21}/L^{*}_{31} = [T_R = 123.1]$	$F_{M21} = [T_{CS} = 13.0]$
PGQ ₃₁	$F^{*}_{S31}/L^{*}_{41} = [P_R = 2800, L_R = 65]$	$F_{M31} = [T_{CR} = 35.94]$
PGQ ₄₁	$F^{*}_{S41} = [0]$	$F_{M41} = [F_1 = 26.17 \quad F_2 = 62.89 \quad F_3 = 53.30]$
PGQ ₁₂	$F^{*}_{S12}/L^{*}_{22} = [P_{str} = 3330 \quad T_{str} = 66.60]$	$F_{M12} = [F_4 = 60.9]$
PGQ ₂₂	$F^{*}_{S22}/L^{*}_{32} = [P_s = 2700, T_s = 92]$	$F_{M22} = [F_{10} = 37.2 \quad F_8 = 0]$
PGQ ₃₂	$F^{*}_{S32}/L^{*}_{42} = [T_R = 123.1]$	$F_{M32} = [T_{CS} = 13.0]$
PGQ ₄₂	$F^{*}_{S42}/L^{*}_{52} = [P_R = 2800, L_R = 65]$	$F_{M42} = [T_{CR} = 35.94]$
PGQ ₅₂	$F^{*}_{S52} = [0]$	$F_{M52} = [F_1 = 26.17 \quad F_2 = 62.89 \quad F_3 = 53.30]$

(3) Achievement of the PGQ-Objective: Starting along the paths of $PGQ_{41} \rightarrow PGQ_{31} \rightarrow PGQ_{21} \rightarrow PGQ_{11} \rightarrow PGQ\text{-Objective}$ and $PGQ_{52} \rightarrow PGQ_{42} \rightarrow PGQ_{32} \rightarrow PGQ_{22} \rightarrow PGQ_{12} \rightarrow PGQ\text{-Objective}$, the two multi-layer PGQs are

implemented for the achievements. The resultant optimum solutions are listed in Table III, showing an actual objective value 118\$/h.

TABLE III.

FINAL OPTIMUM SOLUTIONS

Name of PGQ	SFG/ PLG	MFG
PGQ -Objective	$P_{S1}/L_{11} = [X_{i,9(i=A,C,D,E,F,G,H)} = 32.56 \ 13.61 \ 0.9162 \ 15.71 \ 5.392 \ 6.703 \ 3.308]$ $P_{S2}/L_{12} = X_{i,11(i=D,E,F)} = [0.0142 \ 0.5880 \ 0.1782]$	$P_{M3} = [F_9 = 24.57 \ F_{11} = 46.41]$
PGQ ₁₁	$F_{S11}/L_{21} = [P_s = 2702.59 \ T_s = 92.15]$	$F_{M11} = [F10 = 37.16 \ F8 = 0]$
PGQ ₂₁	$F_{S21}/L_{31} = [T_R = 123.10]$	$F_{M21} = [T_{CS} = 13.89]$
PGQ ₃₁	$F_{S31}/L_{41} = [P_R = 2798.6 \ L_R = 65.23]$	$F_{M3} = [T_{CR} = 35.97]$
PGQ ₄₁	$F_{S41} = [0]$	$F_{M4} = [F1 = 26.11 \ F2 = 62.98 \ F3 = 53.08]$
PGQ ₁₂	$F_{S12}/L_{22} = [P_{str} = 3329.88 \ T_{str} = 66.75]$	$F_{M12} = [F4 = 60.65]$
PGQ ₂₂	$F_{S22}/L_{32} = [P_s = 2702.59 \ T_s = 92.15]$	$F_{M22} = [F10 = 37.16 \ F8 = 0]$
PGQ ₃₂	$F_{S32}/L_{42} = [T_R = 123.10]$	$F_{M32} = [T_{CS} = 13.89]$
PGQ ₄₂	$F_{S42}/L_{52} = [P_R = 2798.6 \ L_R = 65.23]$	$F_{M42} = [T_{CR} = 35.97]$
PGQ ₅₂	$F_{S52} = [0]$	$F_{M52} = [F1 = 26.11 \ F2 = 62.98 \ F3 = 53.08]$

VI. CONCLUSIONS

Inspired by biologic nature of flying goose queue, this paper proposed novel PGQ strategies for plant-wide modeling and optimization, contributing to overcoming the algorithmic deficiencies associated with conventional plant-wide process optimization. To offer PGQ theoretical foundations, key definitions and enabling algorithms have been explicitly. The benefits of the proposed strategies are demonstrated through a case study of TE process. It could be expected that advantages of the PGQ approaches towards plant-wide process optimization are potentially attractive in the following aspects.

(1) Focusing on inseparable objectives, the PGQ approaches could enjoy theoretically lossless decomposition to achieve decentralized optimization schemes.

(2) The PGQ methodology is accommodated to relatively simple nominal forms of the objective functions and process models. In contrast to conventional optimization methods which need to deal with enormous manipulated variables, the PGQ approaches could take advantage of more accurate process models by solving several small-scale PGQ optimization problems, which are more beneficial for effectively utilizing as much information of process variables as possible against modeling uncertainty.

(3) It is found that the algorithms related with the PGQs and PGQ-Objective could be launched independently, helping make options for appropriate optimization algorithms more flexible.

Anyway, to promote this research issue more attractive, an in-depth investigation on PGQ real-time optimization (RTO) approaches together with the applicability potential should be strongly advisable.

REFERENCES

- [1] M.L Darbyan and D.C White. "On-line optimization of complex process units", *Chemical Engineering Progress*, 1988, vol. 67(10), pp.51-59.
- [2] J. Sobieski "Structural sizing by generalized multilevel optimization", *AIAA Journal*, 1987, vol. 1, pp.139-145
- [3] F.Duddeck, "Multidisciplinary optimization of car bodies", *Structural and Multidisciplinary Optimization*, 2008, vol. 35(4), pp.375-389.
- [4] Silva,Valceres V. R. E, Khatib,Wael and Fleming,Peter J. "Control system design for a gas turbine engine using evolutionary computing for multidisciplinary optimization", *Controle and Automação Sociedade Brasileira de Automatica*, 2007, vol. 18(4), pp. 471-478.
- [5] Ranjan Kumar, KazuhiroIzui, MasatakaYoshimura and ShinjiNishiwaki "Multi-objective hierarchical genetic algorithms for multilevel redundancy allocation optimization", *Reliability Engineering and System Safety*, 2009, vol. 94 (4), pp. 891-904.
- [6] T.Gómez, M.González, M.Luque, F.Miguel and F. Ruiz "Multiple objectives decomposition-coordination methods for hierarchical organizations", *European Journal of Operational Research*. 2001, vol. 133, pp.323.
- [7] H. Li and J. Huang, Supervision-driven Plant-wide Process Goose Queue RTO Methodologies, *Industrial and Engineering Chemistry Research*, unpublished.
- [8] C. J.Cutts and J. R. Speakman. "Energy savings in formation flight of Pink-footed Geese", *Journal of Experimental Biology*, 1994, vol. 189 (1), pp. 251-261.
- [9] I.Newton, *The Migration Ecology of Birds*. Elsevier. 2007.
- [10] D.Ekrem, U.Mitat and F.A.Ali, "Migrating birds optimization: a new meta-heuristic approach and its application to the quadratic assignment problem", *Lecture Notes in Computer Science*, 2011, vol. 6624,pp.254-263.
- [11] F. Cattivelli and A. H. Sayed, "Modeling bird flight formations using diffusion adaptation", *IEEE Transactions on Signal Processing*, 2011,vol. 59, (5).
- [12] F. Cattivelli and A. H. Sayed, "Self organization in bird flight formations using diffusion adaptation," *IEEE international Workshop on Computational Advances in*

- Multi-Sensor Adaptive Processing(CAMSAP)*,2009, pp. 49-52.
- [13] T.Williams and R.Otto, "A generalized chemical processing model for the investigation of computer control", *AIEE Transaction*, 1960, vol. 79, pp.458.
- [14] J. J.Downs and E. F.Vogel. "A plant-wide industrial process problem control", *Computers and Chemical Engineering*, 1993,vol. 17, pp. 245.
- [15] N.L.Ricker, "Optimal steady state operation of the Tennessee Eastman challenge process", *Computers and Chemical Engineering* ,1995, vol. 19, pp.949-989.



Jingwen Huang was born in Shandong province, P. R. China in 1978. She received the M.S. degree in China University of Mining and Technology in 2003. At present, she is the part time Ph. D. candidate and Lecturer in Beijing University of Chemical Technology. Her research interests are process modeling, control and optimization.



Hongguang Li was born in Jilin Province, P. R. China in 1963. He received the Ph. D. degree in East China University of Science and Technology in 2004. At present, he is professor in Beijing University of Chemical Technology. His research interests are Modeling, control and optimization of chemical process as well as computer based intelligent control for industrial plants.

A New Sub-topic Clustering Method Based on Semi-supervised Learning

Xiaodan Xu

College of Mathematics, Physics and Information Engineering, Zhejiang Normal University
Jinhua, China
Email:xuxiaodan@zjnu.cn

Abstract—Sub-topic clustering is a crucial step in multi-document summarization. The traditional k-means clustering method is not effective for topic clustering because the number of clusters k must be given in advance. This paper describes a new method for sub-topic clustering based on semi-supervised learning: the method firstly partition the set of sentences into disjoint subsets, each of which contained sentences covering exactly one topic, and labels the sentences which have high scores in the topic, then use the method of constrained-k-means to decide the number of topics, and finally get the sub-topic sets by k-Means clustering. This algorithm can dynamically generate the number of k-means clustering, and the experiment result indicates that the accuracy of clustering is improved.

Index Terms—Sub-topic clustering, semantic distance, semi-supervised learning, k-means clustering

I. INTRODUCTION

With the continuing growth of online information, it has become increasingly important to provide improved mechanisms to find and present textual information effectively. Conventional IR systems find and rank documents based on maximizing relevance to the user query. Some systems also include sub-document relevance assessments and convey this information to the user. More recently, Single document summarization systems provide and automated generic abstract or a query relevant summary. However, large scale IR and summarization have not yet been truly integrated, and the functionality challenges on a summarization system are greater in a true IR or topic-detection context.

Consider the situation where the user issues a search query, for instance on a news topic, and the retrieval system finds hundreds of closely-ranked documents in response. Many of these documents are likely to repeat much the same information, while differing in certain parts. Summaries of the individual documents would help, but are likely to be very similar to each other, unless the summarization system takes into account other summaries

that have already been generated. Multi-document summarizations are likely to be essential in such situations. Ideally, multi-document summaries should contain the key shared relevant information among all the documents only once, plus other information unique to some of the individual documents that are directly relevant to the user's query.

Though many of the same techniques used in single-document summarization can also be used in multi-document summarization, there are some significant difference such as the degree of redundancy, the compression ratio, the co-reference problem in summarization presents.

At present, the method based on text clustering is used in multi-document summarization and gets the good results. Reference [1] R.Radev present a multi-document summarizer, called MEAD, which generates summaries using cluster centroids produced by a topic detection and tracking system. MEAD uses information from the centroids of the clusters to select sentences that are most likely to be relevant to the cluster topic. Reference [2] Endre Boros get the multi-document summaries by utilizing complete sentences from the documents in the collection. In this method, classic clustering techniques were employed in an attempt to partition the set of sentences into disjoint subsets or clusters, each of which contained sentences covering exactly one topic. Clusters are ranked by their similarity with the vector of the term frequencies of all terms appearing in the documents to be summarized. In this method, the similar sentences in multi-document set are combined into one class, each class is one topic of multi-document set, and then multi-document set can be composed of sub-topic sets. Reference [3][4][5] described this method. Therefore the sub-topic detecting is important. Usually, there are two methods for detecting the sub-topics: one is based on hierarchy clustering and the other is based on k-means clustering. The hierarchy clustering method needs an end-clustering threshold, which is hard to decide. The traditional k-means clustering must be given in advance the number of clusters k, but in the actual cases, k is difficult to establish; In addition, traditional k-means algorithm has powerful local search capability, but easily

Manuscript received September 5,2011 ; revised December 19,2011 ; accepted January 1,2012

falls into local optimum. Genetic algorithm can get the global optimal solution, but the convergence is fast.

In view of this, this paper presents a new method of sub-topic clustering based on semi-supervised learning. The algorithm first partition the set of sentences into disjoint subsets or clusters, each of which contained sentences covering exactly one topic, and labels the sentences which have high scores in the topic, then use the method of constrained-means to decide the number of topics, and finally get the topic sets by k-Means clustering.

The rest of the paper is organized as follows. In section 2 relationship between the approaches here proposed and relevant literature is presented. In section 3 the semantic distance between sentences and the algorithm of sub-topic clustering are formally described. Section 4 reports experimental results and comparison with related methods. Finally, in section 5 conclusions are drawn.

II. RELATED WORK

The literature related to this work can be grouped into two main categories: semi-supervised clustering, the evaluation method of the sub-topic detecting and multi-document summarization.

In many machine learning domains (e.g. text processing, bioinformatics), there is a large supply of unlabeled data but limited labeled data, which can be expensive to generate. Consequently, semi-supervised learning, learning from a combination of both labeled and unlabeled data, has become a topic of significant recent interest.

Clustering is an unsupervised learning problem, which tries to group a set of points into clusters such that points in the same cluster are more similar to each other than point s in different clusters, under a particular similarity metric. Clustering problems can also be categorized as generative or discriminative. In the generative clustering model, parametric form of data generation is assumed, and the goal in the maximum likelihood formulation is to find the parameters that maximize the probability of generation of the data given the model. In the most general formulation, the number of clusters K is also considered to be an unknown parameter.

Semi-supervised clustering, which uses class labels or pair wise constraints on some examples to aid unsupervised clustering, has been the focus of several recent projects. Semi-supervised clustering can group data using the categories in the initial labeled data as well as extend and modify the existing set of categories as needed to reflect other regularities in the data.

Existing methods for semi-supervised clustering fall into two general approaches: search-based and similarity-based methods. In search-based approaches, the clustering algorithm itself is modified so that user-provided labels or constraints are used to bias the search for an appropriate partitioning. In similarity-based approaches, an existing clustering algorithm that uses a similarity metric is employed; however, the

Similarity metric is first trained to satisfy the labels or constraints in the supervised data. Reference[6][7][8][9]Several similarity metrics have been used for similarity-based semi-supervised clustering, including string-edit distance trained using EM, Jensen-Shannon divergence trained using gradient descent, Euclidean distance modified by a shortest-path algorithm, or Mahalanobis distances trained using Convex optimization. Several clustering algorithms using trained similarity metrics have been employed for semi-supervised clustering, including single-link and complete-link and KMeans.

Reference [10][11]Basu proposed two algorithms for semi-supervised clustering with labeled data:seeded KMeans(S-KMeans) and constrained KMeans(C-KMeans). In S-KMeans, the seed clustering is used to initialize the KMeans algorithm. Thus, rather than initializing KMeans from K random means, the centroid of the h th cluster is initialized with the centroid of the h th partition S_h of the seed set. In C-KMeans, the seed clustering is used to initialize the KMeans algorithm as described for the S-KMeans algorithm. However, in the subsequent step, the cluster memberships of the data points in the seed set are not recomputed in the assign_cluster steps of the algorithm—the cluster labels of the seed data are kept unchanged, and only the labels of the non-seed data are re-estimated. C-KMeans seeds the KMeans algorithm with the user-specified labeled data and keeps that labeling unchanged throughout the algorithm. In S-KMeans, the user-specified labeling of the seed data may be changed in the course of the algorithm. C-KMeans is appropriate when the initial seed labeling is noise-free, or if the user does not want the labels on the seed data to change, whereas S-KMeans is appropriate in the presence of noisy seeds.

Chinese researchers also get some progress in semi-supervised clustering.

Reference [12] YIN Xuesong presents a discriminative semi-supervised clustering analysis algorithm with pairwise constraints, called DSCA, which effectively utilizes supervised information to integrate dimensionality reduction and clustering. Reference [13] Wang Ling proposed a density-sensitive semi-supervised spectral clustering algorithm (DS-SSC), which incorporate the pairwise constraints knowledge and space consistency prior knowledge into original spectral clustering. Reference [14] Peng Yan proposed a semi-supervised canonical correlation analysis algorithm called Semi-CCA, which used supervision information in the form of pair wise constraints in canonical correlation analysis. Reference[16] Jin Jun described a semi-supervised robust online clustering algorithm called Semi-ROC, which introduced supervision information in the form of class labels into the previously proposed robust online clustering. Reference[17] Wang HJ proposed a semi-supervised cluster ensemble (SCE) model based on both semi-supervised learning and ensemble learning technologies.

In recent years, peoples pay more attention to the standard test sets and large scale evaluations. Two workshops on Automatic Summarization were held: the Document Understanding Conference (DUC) sponsored by the National Institute of Standards and Technology (NIST) started in 2001 in the United States. The Text Summarization Challenge (TSC) task under the NTCIR (NII-NACSIS Test Collection for IR Systems) project started in 2000 in Japan. DUC and TSC both aim to compile standard training and test collections that can be shared among researchers and to provide common and large scale evaluations in single and multiple document summarization for their participants.

III. SUB-TOPIC CLUSTERING

In this paper, D is a document collection, $D=\{d_i|i=1,2,\dots,n\}$, and d_i is a sentence collection: $d_i=\{s_{i,k}|k=1,2,\dots,m\}$, so multi-document set can be described as the set of sentences $s_{i,l} : D = s_{i,l} | s_{i,l} \in d_i$. The sentences which have the same meaning are composed to one topic T_i , the multi-document set can also describe as the set of topics: $D=\{T_i|i=1,2,\dots,k\}$. by this way, multi-document set is the sets of many sub-topics which describe the articles from different aspect. It is useful to improve the quality of multi-document summarization.

Compared with ordinary text file, the Web page includes a large number of additional information, such as html tags, script, internet link, navigating, copyright. This non-text information may influence the speed and quality of the abstraction and must be filtered before making the abstraction. Because some of the html tags (such as `<H1>`, `Title>`, etc) provide useful information for summarization, these useful tags should be kept while cleaning the web pages.

In order to clean the web pages, the following strategy is resented: Firstly, establish a regular expression of text block, and then withdraw the smallest text block which contains the main text information by expression matching. Secondly, combine pattern matching with heuristic rules to clean "the noise" and keep the useful html tags. Finally get the text information.

The so-called automatic summarization method is to detect the sub-topics and deduce the abstract from the different topics by combining some key sentences. It is favorable to understand the logic and fundamental framework of the article and so that the abstract can reflect the contents more correctly and comprehensively. We are going to discuss the sub-topic detecting as the following two steps: calculate the semantic similarity of sentences and detecting the sub-topics based on semi-supervised clustering.

A. Calculate the Semantic Distance between Sentences

Chinese language is different from English in the structure. It does not have the obvious separation symbol between the words, so it need word parsing before further

processing. The low-frequency words (only appear once) and some common words such as "the", "and", "at" are filtered because they contain little information, and the remaining words which called practicable words are used to calculate the semantic distance.

We will use the following formula to calculate the semantic distance between two sentences A and B.

$$D(A, B) = \frac{a}{a + S(A, B)} \quad (1)$$

Among that, $D(A, B)$ is the semantic distance between sentence A and sentence B, $S(A, B)$ is the semantic similarity of sentences A and B. usually, the higher the similarity between sentences, the shorter their semantic distance.

Accordingly, the following formula is used to calculate the semantic similarity of two sentences:

$$S(A, B) = \frac{\sum_{i=1}^m S(a_i, B) + \sum_{j=1}^n S(b_j, A)}{m+n} \quad (2)$$

In the formula (2), sentence A contains words a_1, a_2, \dots, a_m , and sentence B contains the words: b_1, b_2, \dots, b_n . $S(a_i, B)$ is to calculate the semantic similarity between the word a_i and the sentence B, the formula for it would be $:S(a_i, B)=\text{Max}(S(a_i, b_1), S(a_i, b_2), \dots, S(a_i, b_n))$. the same for $S(b_j, A)$.

Obviously, $S(a_i, b_j)$ is to calculate the semantic similarity of word between a_i and b_j .

$$S(a_i, b_j) = \frac{\alpha}{\alpha + D(a_i, b_j)} \quad (3)$$

In the formula, $D(a_i, b_j)$ is the semantic distance between a_i and b_j . We use a thesaurus dictionary to calculate the value of $S(a_i, b_j)$, and the dictionary is provided by Research Center for Information Retrieval (HIT-CIR) of Harbin Institute of Technology. The structure of thesaurus dictionary and its levels are as follows:

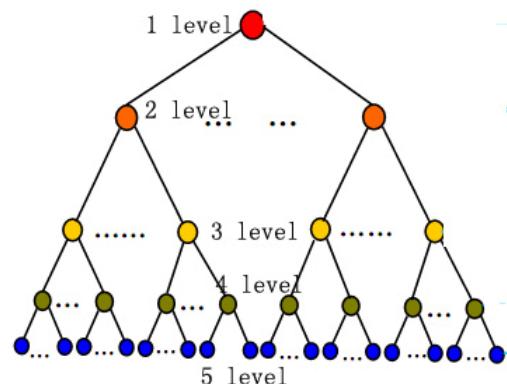


Figure 1. Tree structure of thesaurus dictionary

For the words in dictionary, there are five levels of semantics, which use different mark. The first level use capital letters (A-Z), the second level use small letters(a-z), the third level use numbers(00-99), the fourth and fifth level additional can provide more information. Each word is encoded and arranged according to their semantic relations, for example, the encode of "peach" is Bh07A28.

We use the following formula to calculate the value of $D(a_i, b_j)$:

$$D(a_i, b_j) = 2 * (6 - n) \quad (4)$$

$(2 \leq n \leq 6)$

As for each word has a semantic code, n is the start position which the semantic code is different between two words. For example, in the dictionary, the code for "peach" is Bh07A28, and that for "watermelon" is Bh07A56, so the value of n is 5 and the semantic distance between the two words can be calculated like $D(\text{peach}, \text{watermelon}) = 2 * (6 - 5) = 2$. Exceptional, for the situation $n=1$, if the two words belong to the noun class or verb class, their semantic distance would be $D(a_i, b_j) = 12$, else $D(a_i, b_j) = +\infty$.

B. Cluster the sub-topics

As the semantic distance of each sentence is calculated, the sentences which have small distance will be clustered into one class. Each class would be a sub topic.

Hierarchical clustering was used for finding the initial clusters:

1. Start with each sentence being a cluster of size 1
2. Calculate the distance between each cluster and sort a list of this
- Information so the "closest" clusters are at the top .
3. Pick the two clusters which are "closest" and merge them into a new cluster.
4. Delete the two "closest" clusters and any references to them in the distance list.
5. Go to 2.
6. Stop when have trimmed down to m clusters.

Non-hierarchical clustering, specifically k-means clustering method is given the m clusters as a starting point, with a target of trimming the number of clusters to $n(n < m)$. Since k-means may terminate with more than the target of 10 clusters, the 10 clusters with the most sentence in them are utilized.

From the above we can see that the traditional approach based on hierarchical clustering need an end-clustering threshold, which is hard to decide; the k-means clustering need the Initial Value of k, while the number of sub topics is unpredictable. The method of semi-supervised clustering we will describe in the following can effectively overcome these shortcomings and obtain good results.

1) K-means clustering based on semi-supervised

The basic idea of semi-supervised learning is to use the labeled data to predict the unlabeled data. In semi-supervised clustering; some label level or instance level

supervised information is used along with the unlabeled data in order to obtain a better clustering result.

Many of the existing semi-supervised clustering algorithms are based on the traditional clustering algorithm .On behalf of the algorithm is the semi-supervised k-means algorithm which developed from the classical k-means algorithm.

Given a set of observations $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$, where each observation is a d -dimensional real vector, k-means clustering aims to partition the n observations into k sets ($k \leq n$) $\mathbf{C} = \{C_1, C_2, \dots, C_k\}$ so as to minimize the within-cluster sum of squares.

TABLE I.
WORDS CODE LIST OF THE DICTIONARY

level	Symbols used	For example	semantic
1	Capital letters	B	Main category
2	Small letters	h	Sub category
3	Double digit	07	Detail category
4	Capital letters	A	Words cluster
5	Double digit	28	Detail words cluster

The average value of class k can be described as

$$m_k = \frac{1}{N_k} \sum_{i=1}^{N_k} x_i, \quad k=1, \dots, K,$$

and the Objective function of K-means clustering based on Euclidean distance[18] is as follows:

$$J = \sum_{k=1}^K \sum_{i=1}^{N_k} \|x_i - m_k\|^2 \quad (5)$$

K-means clustering acquires the number of clusters in advance. The random selection of the initial cluster centers will result in the instability and K-means clustering algorithm will be terminated in access to a local optimum value.

K-means clustering method usually takes k as the initial value on the condition that the number of clusters is given in advance. The following two graphs show the different conditions of data distribution with two classes.

The coordinate system in fig2, fig3 and fig4 is based on the distance between two data. The fig3 shows the data sets are disjointed and fig4 shows the data set are intersected.

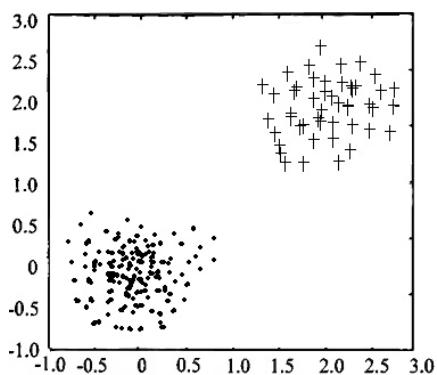


Figure 2. Distributing graph of two disjoint data sets

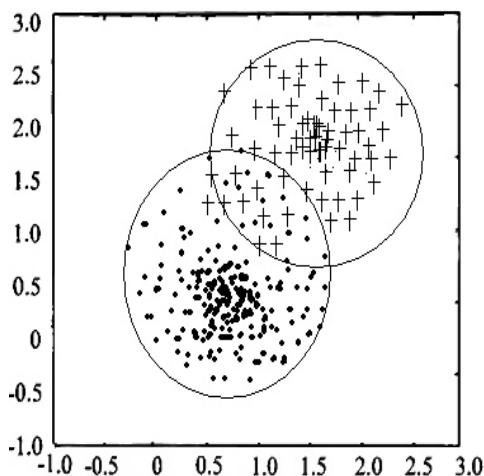


Figure 3. Distributing graph of two intersection data sets

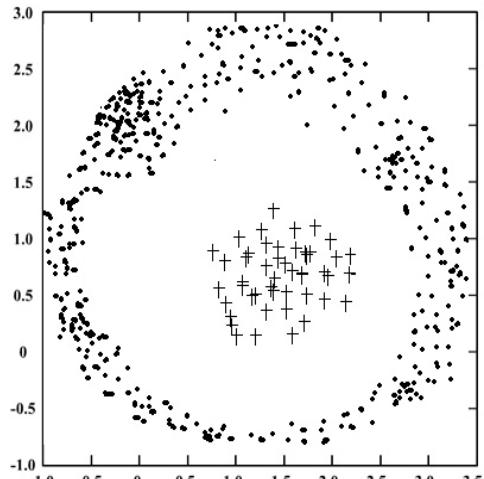


Figure 4. Distributing graph of data sets with non-convex shapes

Figure2 shows that the k-means algorithm can get good effort for disjoint data sets when k=2, while in figure3 and figure4 we can see that the effect of clustering will drop down. Data shown in fig3 contains

two categories, if k=2, data at the cross position will be assigned to error categories, and if we take the value of k greater than 2, more data can meet the rules that data inside the class have a strong similarity and between class have low similarity.

However, the increased K value method requires the following two judgments: one is what value k takes can maximize accuracy of clustering on the existing foundation; the second is how to judge the redundant marked categories when the clustering is completed.

In order to solve the problem, we try to use the semi-supervised method to determine the initial value of k:

Suppose a complete data set contains a few labeled data set L (monitoring information) and unlabeled data set U, i and j is the mark of labeled set. We first use constrained-K-means clustering method for the data set {L,U} on condition that k=2, then calculate the number(N) of incorrectly labeled data in L when different values of k, the k that make N to obtain the minimum value should be the best initial value. The formula is as follows:

$$N = \sum_{c=1}^K \min(n_{ic}, n_{jc}) \quad (6)$$

In (6), c is the class number(c=1,2...K), n_{ic} is the count of data that be marked as class i, n_{jc} is the count of data that be marked as class j. The maximum value of k is adopted as 6 according to the experience.

The constrained KMeans method can be described as follows[20]:

Input: set of data points $X=\{x_1, \dots, x_n\}$, number of clusters K, set $S=\{S_1, S_2, \dots, S_k\}$ of initial seeds

Output: disjoint k partitioning $\{x_h\}(h=1..,k)$ of X such that the KMeans objective function is optimized

1. Initialize clusters: $u_h^{(0)} \leftarrow \frac{1}{|S_h|} \sum_{x \in S_h} x$, for $h=1, \dots, K$; $t \leftarrow 0$

2. repeat until convergence

2a. assign_cluster: for $x \in S$, if $x \in S_h$ assign x to the cluster h (ie., set $X_h^{(t+1)}$). For $x \notin S$, assign x to the cluster h^* (ie., set $X_{h^*}^{(t+1)}$), for $h^* = \arg \min_h \|x - u_h^{(t)}\|^2$

2b. estimate_means: $u_h^{(t+1)} \leftarrow \frac{1}{|X_h^{(t+1)}|} \sum_{x \in X_h^{(t+1)}} x$

2c. $t \leftarrow (t + 1)$

The detail method of sub-topic clustering will be described in the following section 2).

2) the algorithm of semi-supervised clustering

While detecting the sub topics, we use the following method:

(1) Given the number of sub-topics (n=3), get the primal sets of sub-topics by hierarchy clustering, these

sets can be described as $T = (T_i)$ ($1 \leq i \leq n$), where T_i is collection of sentences, $T_i = (s_{i,k} | k = 1, 2, \dots, m)$;

(2) For each T_i , label the sentences that have high similarity and add them into the labeled data set, then get two sets: labeled data set L which contains few labeled sentences, and unlabeled data set U ;

(3) Use constrained-K-means clustering method for the data set $\{L, U\}$ on condition that $k=2$, then calculate the number(N) of incorrectly labeled data in L when different values of k , take the k that make N to obtain the minimum value.

(4) Cluster all sentences with k-means given k , and finally get the sub topics.

TABLE II.
SAMPLES IN EXPERIMENT

Theme	Collection number	The average number of sentences
military	2#,3#,12#,14#,15#	320
art	1#,4#,5#	280
amusement	6#,10#,13#	300
education	7#,8#,9#,11#	210

To determine which sentences should be selected to be included in the summary and the order in which they should appear, clusters were ranked by their similarity to the collection term frequency vector. The sentences within each cluster were then ranked by their similarity to their cluster center. And the Sentences in front of the team row of each sub-topic is detected to be included in the summary in turns .

IV. EXPERIMENTS AND EVALUATION

The presented experiment sample comes from people's network in 2001 which includes about 6000 web pages. The data consists of eight classes, namely military, lift, amusement and so on. In the experiment, we extract 15 collections of web pages; each collection contains 5-10 pages. The web pages in one collection refer to the same topic.

The collection of sample in the experiment is list in table II.

While detecting the sub-topics, we compare the semi-supervised clustering with the traditional hierarchy clustering, the result shown in table III.

In this paper, we first get the sub topics by expert, and then take the clustering accuracy P to evaluate the results.

Above that P is describes as $P = \frac{N_{right}}{N_{all}}$., N_{right} refers to

the number of sentence that to be correctly classified by our system.

The experimental results show that some documents based on hierarchical classification do not play an effect, there are two reasons: one is that the number of sub-

TABLE II.
THE ACCURACY OF CLUSTERING BY TWO METHODS

web pages set	Sub topics number	Hierarchy clustering	Semi-supervised clustering
1#	4	70%	75%
2#	2	80%	85%
3#	3	78%	78%
4#	5	55%	58%
5#	3	66%	70%
6#	5	60%	62%
7#	3	73%	77%
8#	4	70%	78%
9#	5	74%	80%
10#	4	61%	68%
11#	3	81%	88%
12#	3	80%	80%
13#	4	65%	68%
14#	4	71%	80%
15#	3	75%	82%

topics pre-given is not the best; the second is that the method based on hierarchical classification can not be backtracking. While the semi-supervised learning method proposed in this paper can be more precise in determine the number of sub-topics, which effectively improve the classification results.

After sub-topic clustering, the key sentences from different topics are extracted to combine the summary. In order to verify the validity of our approach, we extract 5 web pages sets from the military theme and get the summarization with the following 3 methods:

1. TOP-N method: to get the summary by selecting the former N sentences of each article.

2:MEAD method: it firstly select the key words as the mead of documents, and then measure the importance of one sentence by calculating the similarity between the sentence and the mead, finally get the key sentences as a summary.

3.STSB method: summary based on sub-topic clustering which is provided in this paper. The sentences within each cluster were then ranked by their similarity to their cluster center. one sentence at each iteration was selected to be included in the summary. We use the

precision of summary p to evaluate the effect, $p=(\text{the number of correct sentences})/(\text{number of all sentences in summary})$. We also use the experts' summary as the evaluation standard.

The following paragraph 5 shows the precision of summary with three methods in different compression ratio. The compression ratio are 10%,20%,30%.

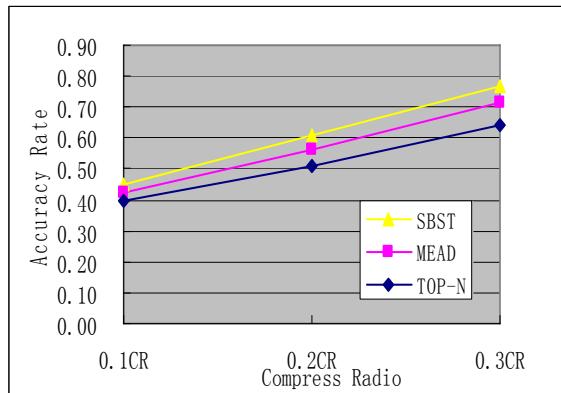


Figure 5. The prcision of summay in different compression ratio

As is shown in figure 5, the quality of most summary is satisfied. However, there is still some redundant information in the summary. In the test, we found that some summary include the sentences with the same meaning, that is because the same important message will be described many times in the sub-topic.

In order to reduce the redundancy, we take the following method: when a sentence is selected into the summary, it must abide by the rules: the key words' coverage rate of summary will be the maximum in all situations, when this sentence is added into the summary.

$$S_{Ti} = \arg \max_{Sij} \frac{|Sumword \cup \{Senword_{ij}\} \cap Muldocword|}{|Muldocword|}$$

in the formula , S_{Ti} is the candidate sentence in sub-topic T_i , $Sumword$ is the key words set in summary , and $Muldocword$ is the key words set in multi-documents. $Senwords_{ij}$ is the words set in sentence S_{ij} . The purpose is to make the summary cover the content as much as possible. Aided with this strategy, the precision of summary is increased to 80% in figure 6 (SBST').

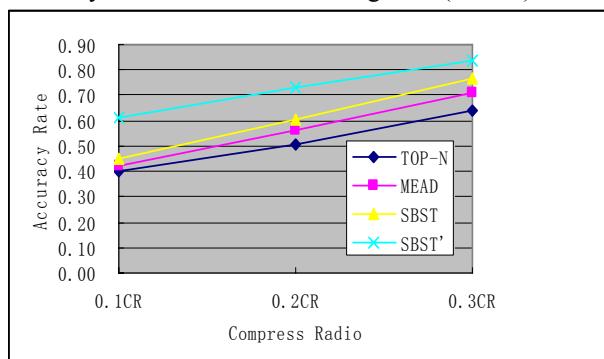


Figure 6. The new prcision of summay in different compression ratio

V.CONCLUSIONS

This paper presents a new method for sub-topic clustering based on semi-supervised clustering. In this method, the semi-supervised clustering is more effective than hierarchical clustering and k-means clustering because it can get the best value of k according to the characteristics of the data. Experiments show that the method is useful. In the next work we will study how to use the technology of natural languages understanding to improve the quality of summary, especially to improve the readability of the summary.

REFERENCES

- [1] R. Radev, Hongyan Jing and Małgorzata Budzikowska.2000.Centroid-based summarization of multiple documents:sentence extraction,utility-based evaluation, and user studies//ANLP/NAACL 2000 Workshop C,2000:21-29
- [2] Boros E et al. A clustering based approach to creating multi-document summaries// Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New Orleans, LA , 2001: 34-42
- [3] Naomi Daniel,Dragomir Radev, and Timothy Allison.Sub-event based multi-document summarization[A].In:HLTNAACL Workshop on Text Summarization.Edmonton Alberta,Canada.2003:9-16
- [4] Pascal Fung,Grace Ngai Combining Optimal Clustering and Hidden Markov Model for Extrative.In:Proceedings of the ACL 2003 workshop on multilingual summarization and question answering.2003:21-28
- [5] Qin Bing,Liu Ting,ChengShanlin,etc.Sentence Optimum Selection for Multi-Document Summarization(in Chinese).Journal of Computer Research and Development.2006.43:1129-1134
- [6] Bilenko,M.,&Mooney,R.J..Adaptive duplicate detection using learnable string similarity measures .InProceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD-2003),pp.39-48Washington,DC.
- [7] Klein,D.,Kamvar,S.D.,&Manning,C..From instance-level constraints to space-level constraints:Making the most of prior knowledge in data clustering.In Proceedings of the The Nineteenth International Conference on Machine Learning(ICML-2002) Sydney,Australia .
- [8] Hillel,A.B.,Hertz,T.,Shental,N.,&Weinshall,d. Learning distance functions using equivalence relations. In Proceedings of 20th International Conference on Machine Learning(ICML-2003)
- [9] Xing,E.P.,Ng,A.Y.,Jordan,M.I.,&Russell,S..Distance metric learning,with application to clustering with side-information.In Advances in Neural Information Processing Systems 15 .MIT Press. 2003.
- [10] Basu S,Banejee A , Mooney R J . Semi, Supervised clustering by seeding || Proc of the 19th International Conference on Machine Learning.Sydney,Australia,2002:19-26.

- [11] Basu S,Banerjee A,Mooney R J.Semi-Supervised Clustering by Seeding//Proc of the 19th International conference on Machine Learning.Sydney,Australia,2002:19-26
- [12] Yin Xuesong,Hu Enliang,Chen Songcan. Discriminative Semi-Supervised Clustering Anaylsis with Pairwise constraints. Journal of Software,2008,19(11):2791-2802(in Chinese).
- [13] Wang Ling,Bo Liefeng,Jiao Licheng. Density-Sesitive Semi-Supervised Spectral Clustering .Journal of Software,2007,18(10):2412-2422(in Chinese)
- [14] Xiao Yu,Yu Jian. Semi-Supervised Clustering Based on Affinity Propagation Algorithm. Journey of Software,2008,19(11):2803-2813(in Chinese)
- [15] Peng Yan, Zhang Daoqiang. Semi-Supervised Canonical Analysis Alogorithm. Journey of Software,2008,19(11):2822-2832(in Chinese)
- [16] Jin Jun, Zhang Daoqiang. Semi-Supervised Robust On-line Clustering Alogorithm .Journal of Computer Research and Development,2008:496-502(in Chinese)
- [17] Wang Hongjun,Li Zhishu,Qi Jianhuai,etc.Semi-Supervised Cluster Ensemble Model Based on Bayesian Network. Journey of Software,2010,21(11):2814-2824(in Chinese)
- [18] Li Kunlun,Cao Zheng,etc. Some Developments on Semi-supervised Clustering(in Chinese).PR&AI. 2009.10:735-742
- [19] ZHU X J.Semi-supervised Learning Literature Survey[R].Madison:University of Wisconsin,2008
- [20] Sugato Basu, Semi-supervised Clustering:learning with limited user feedback, Austin.2003.11



Xiaodan Xu,Jinha,China,1978.Received M.A degree of Software Engineering in National University of Defense Technology in 2005,Changsha,China. Her reaserch interests include data mining and knowledeg discover,computational intelligence ,and automatic summarization.She has worked in the Department of Mathematic Physics and Information Engineering,Zhejiang Normal University.China.
A New Method of Sub Topic Clustering in Multi-Document Summarization,Xu.

Multi-feature Fusion Face Recognition Based on Kernel Discriminate Local Preserve Projection Algorithm under Smart Environment

Di Wu^{1,2,3}

1. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou, China

2. Key Laboratory of Gansu Advanced Control for Industrial Processes, Lanzhou, China

3. Manufacturing Engineering Technology Research Center of Gansu, Lanzhou, China

Email: wudi6152007@163.com

Jie Cao^{1,2,3,4}, Jinhua Wang^{1,2,3}, Wei Li^{2,3,4}

4. College of Computer and Communication, Lanzhou University of Technology, Lanzhou, China

Email: {caoj@lut.cn, wjh0615@lut.cn, lwyz815@163.com}

Abstract—In this paper, a new face recognition method based on kernel discriminate local preserve projection(KDLPP) and Multi-feature fusion under smart environment is proposed. In order to solve the small sample size problem, combined with kernel theory and QR decomposition, a new face recognition algorithm named kernel discriminate local preserve projection is proposed based on discriminate local preserve projection algorithm. considered the external features are useful in face recognition, because hair is a highly variable feature of human face, so we combined hair features and DCT features on the feature layer. The experiments on the AMI database indicate the proposed method can enhance the accuracy of the recognition system effectively.

Index Terms—Kernel Discriminate Local Preserve Projection(KDLPP), Hair Feature, Discrete Cosine Transform, Feature Fusion

I. INTRODUCTION

With the past decades, face recognition has become a very popular area of research in pattern recognition, computer vision and machine learning. Due to the immense application potential in military, commercial, building a automated system to recognize face in still images or video clips is necessary. Face recognition can be defined as the identification of individuals from images of their face by using a stored database of face labeled with people's identities[1]. This objective is very challenging and complex because the appearances of individual's face features are always affected by the factors such as illumination conditions, aging, 3D poses, facial expression and disguise including glasses and cosmetics[2]. Some other problematic factors such as noise and occlusion also impair the performance of the face recognition algorithms.

In the last ten years, face recognition under smart meeting room environment has been raised and become an hot research area. The smart meeting environment is installed four cameras on the four corner and microphone arrays on the table to tracking and

recognizing peoples joined in the meeting[3]. However, early studies all focused on the audio features and hardly any research based on visual features. Japanese researchers tried to use the visual characteristics of video sequence to study the communication process over the conference, they extracted the eye features between the people intercourse in the meeting, and using these features to present the influence degree of speaker to other peoples[4], to our best knowledge, this is the first time of researchers using visual features to discover the multi-people communication process, the drawback of this research is lack of quantitative analyse result of the experiment. In 2007, the research of IDIAP lab tried to use motion vector and residual encoding bit rate between two frames as face features[5]. In [6], chen used Discrete Cosine Transform coefficients as face features to recognize peoples in the meeting.

Dimension reduction is a key problem in face recognition and many useful techniques for dimensionality reduction has been developed. He et al.[7,8] proposed the local preserve projections (LPP) which building a graph incorporating neighbourhood information of the data set and provides a way to the projection of the test data point. In contrast to most manifold learning algorithms, LPP possesses a remarkable advantage that it can generate an explicit map. To consider the discriminant information of recognition task, several locality preserving discriminant analysis methods have been mentioned in recent years. Hu [9] proposed an orthogonal neighbourhood preserving discriminant analysis (ONPDA) method, which effectively combines the characteristics of LDA and LPP. Yu et al.[10] presented a discriminant locality preserving projections (DLPP) method to improve the classification performance of LPP. All the mentioned locality preserving methods also suffer from the SSS problem too. So PCA approach, which discards some useful discriminatory information is often used before LPP or DLPP. Yang et al.[11] proposed a null space discriminant locality preserving projections (NDLPP) algorithms. However, NDLPP merely utilizes the

discriminant information in the null space of the locality preserving within-class scatter.

Most of the face recognition method mentioned above only use facial information, as we know, external information such as hair, facial contour and clothes also can provide the discriminant evidence[12]. Although external information are useful, but their detection, representation, analysis and application are seldom been studied in the computer vision community. Ji et al.[13] used hair features for gender classification, they used length, area and texture information and split as hair features. Liu et al.[14] also used hair features for gender recognition.

In this paper, in order to solve the small sample size problem, by incorporating the kernel trick, a new face recognition algorithm based on discriminant locality preserving projections (DLPP) method and QR decomposition is proposed, which called kernel discriminant locality preserving projections (KDLPP). The enhanced algorithm can not only handle the SSS problem, but also can adequate to describe the complex variation of face images. Considering the important role of hair features in face recognition, we study hair feature extraction and fusing with discrete cosine transform coefficient on the feature layer in order to capture the most recognize information.

The rest of this paper is organized as follows: in Section II we describe the feature extraction process of hair and face. Section III introduce the kernel discriminant locality preserving projections algorithm(KDLPP). We present our recognition method in Section IV. The experiment result are shown in Section V. Section VI offers our conclusion.

II. FEATURE EXTRACTION

A. Hair Feature Extraction

Hair is a highly variable feature of human appearance. It perhaps is the most variant aspect of human appearance. Until recently, hair features have been discarded in most of the face recognition system. To our best knowledge, there are two different algorithms in the literature about hair feature extraction. Yacoob et al.[15] developed a computational model for measuring hair appearance. They extracted several attributes of hair including color, volume, length, area, symmetry, split location and texture. These are organized as a hair feature vector. Lapedriza et al.[16] learned a model set composed by a representative set of image fragments corresponding to hair zones called building blocks set. The building blocks set is used to represent the unseen images as it is a set of puzzle pieces and the unseen image is reconstructed by covering it with the most similar fragments. We adopt the former method and modify it in this study.

The basic symbols used in the geometric hair model are depicted in Figure 1. Here G is the middle point between the left point L and the right eye point R, I is the point on the inner contour, O is the point on the outer contour, and P is the lowest point of hair region.

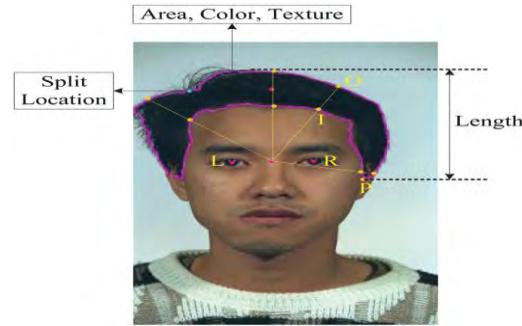


Figure 1. The geometric hair model

The hair feature extraction consist of the following three steps:

1. Extract hair length features: we define the largest distance between a point on the outer contour and P as the hair length. The normalized distance L_{hair} is defined as:

$$L_{hair} = \max(\text{dist}(O_y, P_y)) / \text{Girth}_{face} \quad (1)$$

Where Girth_{face} is the girth of the face region.

2. Extract hair area features: we define the area covered by hair as the hair surface. Based on the hair model, the normalized hair area is defined as:

$$\text{Area}_{hair} = \text{RealArea}_{hair} / \text{RealArea}_{face} \quad (2)$$

Where RealArea_{hair} is the real area of hair and RealArea_{face} is the area of face.

3. Extract hair color features: to obtain the color information in the hair region, we follow the method described in [17]. Based on this approach, the color distortion at pixel i is calculated by

$$CD_i = \|I_i - \alpha_i E_i\| \quad (3)$$

Where I_i and E_i denote the actual and the expected RGB color at pixel i respectively, the I_i is stated as follow:

$$I_i = (I_r(i), I_g(i), I_b(i)) \quad (4)$$

According to the definition above, the color distortion at pixel i is also can calculated as follows:

$$CD_i = \sqrt{\left(\frac{I_r(i) - \alpha_i \mu_r}{\delta_r}\right)^2 + \left(\frac{I_g(i) - \alpha_i \mu_g}{\delta_g}\right)^2 + \left(\frac{I_b(i) - \alpha_i \mu_b}{\delta_b}\right)^2} \quad (5)$$

Where α_i represent the current brightness with respect to the brightness of the model:

$$\alpha_i = \frac{(I_r(i) \frac{\mu_r}{\delta_r})^2 + (I_g(i) \frac{\mu_g}{\delta_g})^2 + (I_b(i) \frac{\mu_b}{\delta_b})^2}{(\frac{\mu_r}{\delta_r})^2 + (\frac{\mu_g}{\delta_g})^2 + (\frac{\mu_b}{\delta_b})^2} \quad (6)$$

Where (μ_r, μ_g, μ_b) and $(\delta_r, \delta_g, \delta_b)$ are the mean and standard deviation of color in the training set respectively. By use of equation (3),(5) and (6), we can obtain the expected RGB color values.

$$E = (E_r, E_g, E_b) \quad (7)$$

By concatenating all the hair feature mentioned above, we obtain a feature vector of hair at pixel i as follows:

$$\begin{aligned} Hair_i &= [length, area, color_i] \\ &= [L_{hair}, Area_{hair}, E_{r_i}, E_{g_i}, E_{b_i}]^T \end{aligned} \quad (8)$$

We normalized the hair region as size of $L \times N$, so the feature vector of hair region is represent as follow:

$$Hair_{vector} = \begin{bmatrix} Hair_1 & \dots & Hair_N \\ \dots & \dots & \dots \\ Hair_{L1} & \dots & Hair_{LN} \end{bmatrix} \quad (9)$$

B. Face Feature Extraction

We use discrete cosine transform(DCT) coefficients to characterize the face feature. DCT has been shown promising performance applied on the human recognition system. For an $M \times N$ image, where each image corresponds to a 2D matrix, DCT coefficients are calculated as follows[18]:

$$\begin{aligned} C(u, v) &= a(u)a(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \times \cos \frac{(2x-1)u\pi}{2M} \\ &\quad \times \cos \frac{(2y-1)v\pi}{2N} \quad u = 0, 1, \dots, M \quad v = 0, 1, \dots, N \end{aligned} \quad (10)$$

Where $a(u), a(v)$ is defined by :

$$a(u) = \begin{cases} \sqrt{1/2} & u = 0 \\ 1 & otherwise \end{cases} \quad (11)$$

$$a(v) = \begin{cases} \sqrt{1/2} & v = 0 \\ 1 & otherwise \end{cases} \quad (12)$$

$f(x, y)$ is the image intensity function and $C(u, v)$ is a 2D matrix of DCT coefficients.

III. KERNEL DISCRIMINANT LOCAL PRESERVE PROJECTION ALGORITHM

A. Discriminant Local Preserve Projection Algorithm

A set of face image sample $\{x_i\}$ can be represented as an $M \times N$ matrix $X = [x_1, x_2, \dots, x_N]$, where M is the number of pixels in the image and N is the number of samples. Each face image x_i belong one of the C face class $\{X_1, \dots, X_c\}$. DLPP tries to maximize an objective function as follows[19]:

$$\frac{\sum_{i,j=1}^C (m_i - m_j) B_{ij} (m_i - m_j)^T}{\sum_{c=1}^C \sum_{i,j=1}^{n_c} (y_i^c - y_j^c) W_{ij}^c (y_i^c - y_j^c)^T} \quad (13)$$

Where n_c is the number of samples in the c th class, y_i^c represents the i th projected vector in the c th class, m_i and m_j is separately the mean of the projected vector for the i th class and j th class, such as :

$$m_i = \frac{1}{n_i} \sum_{k=1}^{n_i} y_k^i \quad (14)$$

$$m_j = \frac{1}{n_j} \sum_{k=1}^{n_j} y_k^j \quad (15)$$

Where n_i and n_j is the number of samples in the i th class and j th class separately. W_{ij}^c represents the elements of within-class weight matrix and B_{ij} represents the elements of between-class weight matrix:

$$W_{ij}^c = \exp \left(\frac{-\|x_i^c - x_j^c\|}{\delta^2} \right) \quad (16)$$

$$B_{ij} = \exp \left(\frac{-\|f_i - f_j\|^2}{\delta^2} \right) \quad (17)$$

Where δ is an empirically determined parameter, x_i^c represents the i th vector in the c th class, and f_i is the mean of the i th class:

$$f_i = \frac{1}{n_i} \sum_{k=1}^{n_i} x_k^i \quad (18)$$

Thus ,the between-class weight matrix B and the within-class weight matrix W are defined as follows:

$$B = [B_{ij}] \quad (i, j = 1, 2, \dots, C) \quad (19)$$

$$W^i = [W_{jk}^i] \quad (j, k = 1, 2, \dots, n_i) \quad (20)$$

Suppose that the mapping from x_i to y_i is $y_i = G^T x_i$, then, the objective function (13) can be rewritten as :

$$J(G) = \frac{|G^T F H F^T G|}{|G^T X L X^T G|} \quad (21)$$

Where L and H is Laplacian matrix and defined as follows:

$$L = D - W \quad (22)$$

$$D = diag \{D_1, \dots, D_c\} \quad (23)$$

$$W = diag \{W^1, \dots, W^c\} \quad (24)$$

$$H = E - B \quad (25)$$

$$F = [f_1, f_2, \dots, f_c] \quad (26)$$

Where D_i is a diagonal matrix and its elements are column sum of W^i . E also is a diagonal matrix and its elements are column sum of B .

Now we should give the following definitions:

$$S_w^L = X L X^T \quad (27)$$

$$S_b^L = F H F^T \quad (28)$$

That the equation (21) can be rewritten as:

$$J(G) = \frac{|G^T S_b^L G|}{|G^T S_w^L G|} \quad (29)$$

The transformation matrix $G = [g_1, g_2, \dots, g_k]$ that maximize the objective function (29) can be obtained by solving the generalized eigenvalues problem:

$$S_b^L g_i = \lambda_i S_w^L g_i, \quad g_1 \geq g_2 \geq \dots \geq g_k \quad (30)$$

B. Kernel Discriminant Local Preserve Projection Algorithm

In this section, we present a new KDLPP algorithm to further improve the performance of DLPP algorithm. We use the kernel trick to handle the nonlinearity in a disciplined manner. The KDLPP algorithm involves two major steps[20]. The first step is to obtain the Gram matrix K and then to reduce the dimensionality of the original data features by applying the modified DLPP/QR algorithm.

The key idea of kernel Discriminant Local Preserve Projection Algorithm is to solve the problem of DLPP in an implicit feature space F , which is constructed by the kernel trick. Consider there is a feature mapping ϕ which maps the input data into a higher dimensional inner product space F [21]. So DLPP can be performed in F and it is equivalent to maximizing the following criterion:

$$J(G) = \frac{|G^T S_b^{L\phi} G|}{|G^T S_w^{L\phi} G|} \quad (31)$$

$$S_w^{L\phi} = X^\phi L^\phi X^{\phi T} \quad (32)$$

$$S_b^{L\phi} = F^\phi H^\phi F^{\phi T} \quad (33)$$

Referring to (31), any column of the solution G must lie in the span of all the samples in F , so there exist coefficients α_{ij} such that[22]:

$$g = \sum_{i=1}^c \sum_{j=1}^{n_i} \alpha_{ij} \phi(x_{ij}) \quad (34)$$

Where g represents any one column of the projection matrix G . In other words, we can project each vector onto an axis of F as follows:

$$g^t \phi(x) = \sum_{i=1}^c \sum_{j=1}^{n_i} \alpha_{ij} k(x_{ij}, x) = \alpha^t \varepsilon_x \quad (35)$$

Where

$$\varepsilon_x = (k(x_{11}, x), \dots, k(x_{1n_1}, x), \dots, k(x_{c1}, x), \dots, k(x_{cn_c}, x))^t \quad (36)$$

$$\alpha = (\alpha_{11}, \dots, \alpha_{1n_1}, \dots, \alpha_{ij}, \dots, \alpha_{c1}, \dots, \alpha_{cn_c})^t \quad (37)$$

$$K(x_1, x_2) = \langle \phi(x_1), \phi(x_2) \rangle \quad (38)$$

Thus, by using the definitions of $S_w^{L\phi}$, $S_b^{L\phi}$ and (35), we can obtain:

$$G^T S_b^{L\phi} G = A^T K_b^{L\phi} A \quad (39)$$

$$G^T S_w^{L\phi} G = A^T K_w^{L\phi} A \quad (40)$$

Where

$$K_w^{L\phi} = K(X) L K(X)^T \quad (41)$$

$$K_b^{L\phi} = K(F) H K(F)^T \quad (42)$$

$$K(X) = [K(x_1), K(x_2), \dots, K(x_N)] \quad (43)$$

$$K(F) = [K(f_1), K(f_2), \dots, K(f_C)] \quad (44)$$

$$K(x_i) = \varepsilon_{x_i} = (k(x_{11}, x_i), \dots, k(x_{1n_1}, x_i), \dots,$$

$$k(x_{c1}, x_i), \dots, k(x_{cn_c}, x_i))^t \quad i=1, 2, \dots, N \quad (45)$$

$$K(f_i) = \left(\frac{1}{n_i} \sum_{k=1}^{n_i} k(x_{11}, x_{ik}), \dots, \frac{1}{n_i} \sum_{k=1}^{n_i} k(x_{1n_1}, x_{ik}), \dots, \right.$$

$$\left. \frac{1}{n_i} \sum_{k=1}^{n_i} k(x_{c1}, x_{ik}), \dots, \frac{1}{n_i} \sum_{k=1}^{n_i} k(x_{cn_c}, x_{ik}) \right)^t$$

$$i=1, 2, \dots, C \quad (46)$$

So the objective of KDLPP can be written as follows:

$$J(A) = \frac{|A^T K_b^{L\phi} A|}{|A^T K_w^{L\phi} A|} = \frac{|A^T K(F) H K(F)^T A|}{|A^T K(X) L K(X)^T A|} \quad (47)$$

Therefore, similar to DLPP algorithm, the optimal solution of equation (47) can be computed by finding the leading r eigenvalues $\{\alpha_i\}_{i=1,2,\dots,r}$ of $(K_w^{L\phi})^{-1} K_b^{L\phi}$ corresponding to the nonzero eigenvalues. Once $A = [\alpha_1, \alpha_2, \dots, \alpha_r]$ is obtained, for a given pattern x ,

we can map it to a r -dimensional space spanned by the column of A . This projection is given by $y = A^T x$.

The solution of A is complexly and always suffer from the small sample size problem, so we using QR decomposition matrix analysis to handle this issue[23,24].

The first step is to decompose $K_b^{L\phi}$ as follows:

$$K_b^{L\phi} = H_b^{L\phi} (H_b^{L\phi})^T \quad (48)$$

Therefor we do QR decomposition on $H_b^{L\phi}$ by $H_b^{L\phi} = QR$, for any given matrix $G \in R^{r \times r}$, with $r = \text{rank}(H_b^{L\phi})$, the solution of A is given by $A = QG$, that

$$J_\phi(A) = \frac{\left| (QG)^t K_b^{L\phi} QG \right|}{\left| (QG)^t K_w^{L\phi} QG \right|} = \frac{\left| G^t \tilde{k}_b G \right|}{\left| G^t \tilde{k}_w G \right|} \quad (49)$$

$$\tilde{K}_b = Q^t K_b^{L\phi} Q \quad (50)$$

$$\tilde{K}_w = Q^t K_w^{L\phi} Q \quad (51)$$

The final step is to compute an optimal G by solving the largest r eigenvalues problem on

$(\tilde{K}_w)^{-1} \tilde{K}_b$. Table I resume the step of KDLPP algorithm.

TABLE I
PROCEDURE OF KDKPP ALGORITHM

Purpose: compute projection matrix A

Steps:

1. compute L and H from (22) and (25), then compute $K(X)$ and $K(F)$ from (43)and (44).
2. Compute $K_w^{L\phi}$ and $K_b^{L\phi}$ from (41) and (42).
3. Construct $H_b^{L\phi}$ from equation (48).
4. Perform QR decomposition on $H_b^{L\phi}$, $H_b^{L\phi} = QR$.
5. Compute $\tilde{K}_b = Q^t K_b^{L\phi} Q$, $\tilde{K}_w = Q^t K_w^{L\phi} Q$.
6. Compute the r eigenvalues g_i of $(\tilde{K}_w)^{-1} \tilde{K}_b$, corresponding the r largest eigenvalues.
7. The projection matrix is $A = QG$ with $G = [g_1, g_2, \dots, g_r]$.

IV. PROPOSED METHOD

In the past section we know how to extract hair features and DCT features, based on DLPP algorithm and kernel trick we give a new dimensional reduce technique called kernel discriminant local preserve projection algorithm. In this section , we give the face recognition algorithm based on multi-feature fusion and kernel discriminant local preserve projection algorithm. The main procedure of our method is depicted in Figure 2.

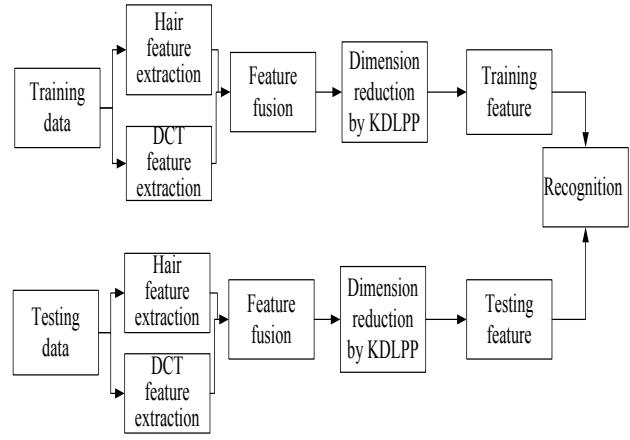


Figure 2.Block diagram of recognition process

After extract the hair features and DCT features, we combined at feature-level as follows[25]:

$$F_{fusion} = [F_{hair}^T \ F_{DCT}^T]^T \quad (52)$$

The steps of the training process is list as follows:

(1) Extract the hair features and DCT features of the images in the training set .

(2) Fusing the hair feature and DCT feature at the feature level in order to obtain fusion feature.

(3) Analyse the fusion features in the training set utilizing the KDLPP algorithm to obtain the projection matrix A .

(4) Project the training fusion feature into the lower dimensional space so that we can get the training features.

The steps of the recognizing process is list as follows:

(1) Extract the hair features and DCT features of the images in the test set.

(2) Fusing the hair feature and DCT feature at the feature level in order to obtain fusion feature.

(3) Project the test fusion feature into the lower dimensional space utilizing the projection matrix A which computed from training process.

(4) classify the test set using the Minimum Euclidean Distance classifier.

V. EXPERIMENTS

A. AMI Database

In this paper, the Augmented Multi-Party Interaction (AMI) corpus were used for our experiment.The AMI corpus consists of audio-visual data captured of our participants in a natural meeting scenario. The participants volunteered their time freely and were assigned roles such as “ project manager ” or

"marketing director" for the task of designing a new remote control device. The teams met over several sessions of varying lengths (15 – 35 minutes).

The meetings were not scripted and different activities were carried out such as presenting at a slide screen, explaining concepts on a whiteboard or discussing while sitting around a table. The participants therefore interacted naturally, including talking over each other. Data was collected in an instrumented meeting room, which contains a table, slide screen, white board and four chairs. While participants were requested to return to the same seat for the duration of a meeting session, they could move freely throughout the meeting. Different audio sources of varying distance to the speaker, and different video sources of varying views and fields-of-view represent audio-visual data of varying quality which is useful for robustness testing. Figure 3 show some samples captured from AMI database.



Figure 3. The captures of AMI database videos

In this experiment, the subset of AMI database named AMIES2016 was used. For this experiment, we captured 5 video segments from each people's video that at last a total of 20 small video segments were obtained. We denoted its as S1 to S20. For the reason of most of the image frames in the video have poor quality and no nose in the images, so we should delete it and then regular the image to guarantee nose is in the center of the image. Then select 10 frames from the video and record as 1 to 10 to construct the AMIES2016 face database. For each image, we normalized it to form the uniform size of 64*64. Figure 4 show 10 frames selected from one video.



Figure 4. Face images of AMIES2016 database videos

B.Experiment results

We randomly take k images from each class as the training data ,with $k \in \{2,3,...,9\}$, and leave the rest $10 - k$ images as the test data. The Nearest Neighbour algorithm was employed using Euclidean distance for classification. There are three small experiments taken in our experiment as follows:

Experiment A. Compare recognition accuracy based on KDLPP algorithm under different kernel functions.

The input data of the LDLPP is the kernel matrix and it is necessary to choose an adequate kernel function to construct this matrix. In this paper, we used Polynomial kernel function , Gaussian RBF kernel function and Fractional polynomial kernel function. Table II present the kernel functions used in our studies.

TABLE .II KERNEL FUNCTIONS	
1.Polynomial kernel function	
$K(x, y) = (1 + xy^d), d \in N$	
2.Gaussian RBF kernel function	
$K(x, y) = \exp(-\ x - y\ ^2 / 2\delta^2)$	
3.Fractional polynomial kernel function	
$K(x, y) = (1 + xy^d), 0 < d < 1$	

In order to illustrate the effect of kernel function choice, Figure 5 to Figure 7 show the results of KDLPP algorithm with different kernel functions. In Figure 5 we can show that for Polynomial kernel the performance decrease with the parameter d increasing. And globally gives less result than Gaussian RBF kernel function and Fractional polynomial kernel function. For Gaussian RBF kernel function, the value $\delta^2 = 10^9$ gives maximum recognition rate compares to others values of δ . The performance of Fractional polynomial kernel function with value $d = 0.4$ is good but it is lower than Gaussian RBF kernel function.

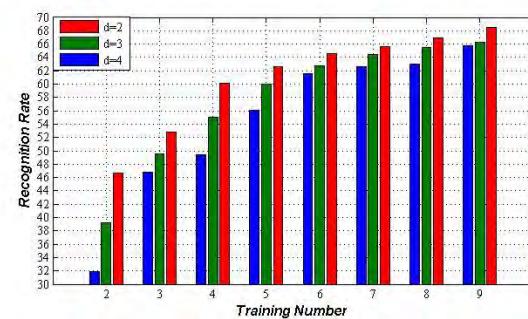


Figure 5. Recognition accuracy of KDLPP algorithm under Polynomial kernel function

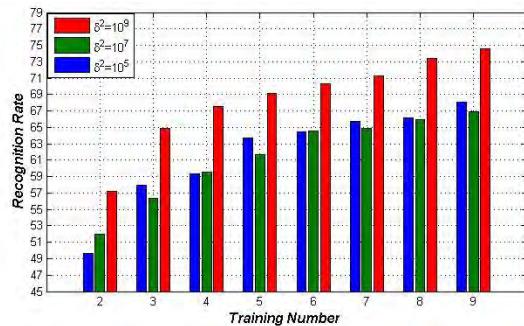


Figure 6. Recognition accuracy of KDLPP algorithm under Gaussian RBF kernel function

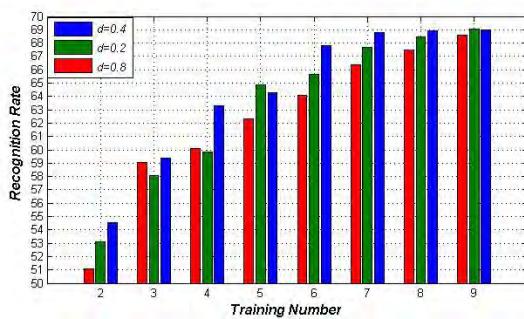


Figure 7. Recognition accuracy of KDLPP algorithm under Fractional Polynomial kernel function

Experiment B. Compare recognition accuracy based on different algorithms.

In this small experiment, we tested the FDA and DLPP methods compare to our proposed KDLPP algorithm, the kernel function used in this experiment is Gaussian RBF kernel function, the value $\delta^2 = 10^9$. Figure 8 give the recognition rate result. From the result we can show that KDLPP algorithm gives the best result under any training number situations, and FDA method give the worst result. From the figure we can also know that the face recognition rate under smart meeting environment is less than the standard face database environment because of the problem of poor quality image, lighting condition and facial expression change and so on.

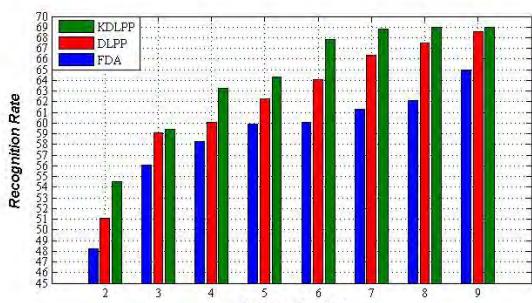


Figure 8. Recognition accuracy of different algorithms based on hair and DCT feature fusion

Experiment C. Compare recognition accuracy based on KDLPP algorithm under different features.

In this small simulation, we compare the DCT feature with the hair and DCT feature fusion under KDLPP algorithm, the kernel function used in this experiment is Gaussian RBF kernel function, the value $\delta^2 = 10^9$, the recognition rate result are shown in Figure 9. From the result we can shown that hair feature play an important role in face recognition, the recognition result can improved significantly.

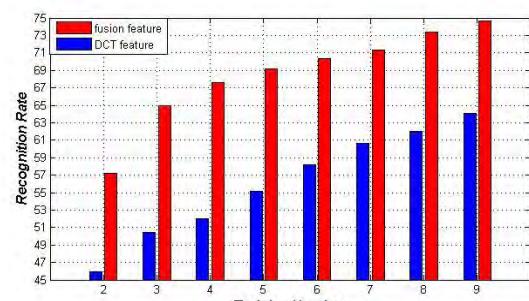


Figure 9. Recognition accuracy of different features based on KDLPP algorithm

VI. CONCLUSION

This paper investigate how to exploit effectively the hair feature information , as well as its fusion with face DCT feature for face recognition based on a new KDLPP algorithm under smart meeting room environment. The external information is crucial for face recognition, so we have presented a modified hair model for extracting hair features, by using this model, hair features are represented as length, area and color. In order to improve the accuracy we fusing it with the DCT features at the feature-level fusion for face recognition. SSS problem is always encountered by the DLPP algorithm, so we proposed a new KDLPP algorithm motivated by the idea of kernel trick and QR decomposition. By introducing a kernel function into discriminant criterion, KDLPP analyse the data in F and produces nonlinear discriminating features that then can work on more realistic situations.

From the experiment result, we can obtain the following observations:(1) hair features play an important role in face recognition, (2) implementing the fusion of hair and DCT features can achieve the best classification accuracy in all of the case in face recognition, (3) KDLPP algorithm can handle the SSS problem and can work under more realistic situations.

From this study, we believe that more external informations such as clothes should be integrated into face features to develop more relative and robust face recognition system under smart meeting room environment.

ACKNOWLEDGE

This work was supported by the Science Foundation of Gansu Province, China (1010RJZA046). The Graduate Supervisor Foundation of Education Department of Gansu Province, China (0914ZTB003). The Finance Department Foundation of Gansu Province, China (0914ZTB148).

REFERENCES

- [1] Cevikalp Hakan, Neamtu Marian, Wikes Mitch, Barkana Atalay. Discriminative common vectors for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2005;27(1):4–13.
- [2] Zhao W, Chellappa R, Phillips PJ, Rosenfeld A. Face recognition: a literature survey. *ACM Computing Surveys* 2003;35(4):399–458.
- [3] Stergiou A, Pnevmatikakis A, Polymenakos L. The AIT Multimodal person identification system for CLEAR 2007. *Multimodal Technologies for Perception of Humans, Lecture Notes in Computer Science*, Baltimore, MD: Springer, 2008: 221-232.
- [4] Carletta J, Ashby S, Bourban S, et al. The AMI meeting corpus: a preannouncement //In Proc of the Workshop on Machine Learning for Multimodal Interaction (MLMI). Edinburgh: 2005: 325-336.
- [5] G.Garau and H.Bourlard. Using audio and visual cues for speaker diarisation initialisation. In Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2010:4942 - 4945.
- [6] CHEN Yan-Xiang, LIU Ming. Audio-visual bimodal speaker identification in a smart environment[J].*JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA*.2010,40(5):486-490.
- [7] X. He, S. Yan, Y. Hu, H. Zhang, Learning a locality preserving subspace for visual recognition, in: Proceedings of the Ninth International Conference on Computer Vision, France, October 2003, pp. 385 – 392.
- [8] X. He, S. Yan, Y. Hu, P. Niyogi, H. Zhang, Face recognition using Laplacian faces, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (3) (2005) 328 – 340.
- [9] H. Hu, Orthogonal neighborhood preserving discriminant analysis for face recognition, *Pattern Recognition* 41 (2008) 2045 – 2054.
- [10] W. Yu, X. Teng, C. Liu, Face recognition using discriminant locality preserving projections, *Image Vision Computing* 24 (2006) 239 – 248.
- [11] L. Yang, W. Gong, X. Gu, W. Li, Y. Liang, Null space discriminant locality preserving projections for face recognition, *Neurocomputing* 71 (2008) 3644 – 3649.
- [12] S. Gutta, J.R.J. Huang, P.J., Wechsler, H.: Mixture of experts for classification of gender, ethnic, origin, and pose of human faces. *IEEE Trans. Neural Networks* 11(4) 948–960 (2000)
- [13] Zheng Ji, Xiao-Chen Lian and Bao-Liang Lu. Gender Classification by Information Fusion of Hair and Face. Published by In-The. November 2008.
- [14] LIU Shuang, XIE Jie-rong, LU Bao-liang. Gender Classification Using Hair Features. *Journal of Computer Simulation*. 2009;26(2):212-216.
- [15] Yacoob, Y.: Detection and analysis of hair. *IEEE Trans. Pattern Anal. Mach. Intell.* 28(7)1164–1169 (2006)
- [16] Member-Larry S. Davis, Lapedriza, A., Masip, D., Vitria, J.: Are external face features useful for automatic face classification? In: *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*
- [17] Yacoob, Y., Davis, L.: Detection, analysis and matching of hair. *The tenth IEEE International Conference on Computer Vision* 1 741–748 (2005)
- [18] Z M Hafed ,M D Levine. Face recognition using the discrete cosine transform. *International Journal of Computer Vision* ,2001 ,43 (3) :167 - 188.
- [19] W. Yu, X. Teng, C. Liu, Face recognition using discriminant locality preserving projections, *Image Vision Computing* 24 (2006) 239–248.
- [20] Baochang Zhang,Yu Qiao. Face recognition based on gradient gabor feature and Efficient Kernel Fisher analysis. *Neural Computing & Applications*,2010,19(4):617-623.
- [21] Wen-Chung Kao, Ming-Chai Hsu, Yueh-Yiing Yang. Local contrast enhancement and adaptive feature extraction for illumination invariant face recognition. *Pattern Recognition*,2010,43(5): 1736-1747.
- [22] PARK H,PARK C H. A comparison of generalized linear discriminant analysis algorithm. *Pattern Recognition*, 2008,41(3):1083-1097.
- [23] Kazuhiro Hotta. Local normalized linear summation kernel for fast and robust recognition. *Pattern Recognition*, 2010, 43(3):906-913.
- [24] MA Xiaohong,ZHAO Linlin.A robust audio watermarking method based on QR decomposition and lifting wavelet transform. *Journal of Dalian University of Technology*,2010,50(2):278-282
- [25] Shufu Xie, Shiguang Shan, Xilin Chen, Jie Chen. Fusing Local Patterns of Gabor Magnitude and Phase for Face Recognition. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 2010,19(5):1349-1361.
- [26] <http://www.amiproject.org/>.



DI WU, born in Xiangtan Hunan province of China in 1985. Received bachelor degree in communication system from Jiujiang university ,China,in 2007, and received master degree in signal and information processing from Lan Zhou university of technology,China, in 2010, now he pursue his doctor degree in control theory and control Engineering from Lan Zhou university of technology,China. His main research area is information fusion theory and application, multi-person speech

recognition.

He joined in The Science Foundation of Gansu Province and The Graduate Supervisor Foundation of Education Department of Gansu Province. also he present some papers:"Face Recognition Based On Pulse Coupled Neural Network";"Combination of SVM and Score Normalization for Person Identification based on audio-visual feature fusion" and so on . all indexed by Engineering Index.



JIE CAO, born in October 1966, Suzhou, Anhui, China. JIE CAO received B.Tech. degree from Ganshu Institute of Technology in 1987, Lanzhou, China, and the M. Tech. degree from Xi'an Jiaotong University in 1994. Hers research interests lie in the areas of information fusion and Intelligent Transportation (ITS).

She is a professor of Lanzhou University of Technology, doctoral tutor, and the "second level" candidates of Gansu leading talent. She presided over the completion of the "Gelatin production process of integrated automation control systems and process parameters optimization," during 2007 to 2010, and got the second Award of Gansu Provincial Science and Technology Progress. She participate Canada and China inter-governmental cooperation project "Regional planning and transport system" of the Canadian International Development Agency (acceptance); organized implementation of the National Technology Support Program "for the key industries of manufacturing information integration platform and application" by the Ministry of Science and acceptance; chaired or participated in 20 projects , currently hosted mainly in the research project are: Natural Science Foundation of Gansu Province, " visual detection, identification and tracking Research based on the sports car "; Gansu Higher operating costs of basic scientific research,"multi-speaker recognition based on audio and video feature fusion "; Natural Science Foundation of Gansu Province, " multi-Speaker Tracking based on the audio and video feature fusion ".



JINHUA WANG , born in Tianshui, Gansu province of China in 1978. Received B.Tech. degree from Xi'an university of science and technology in 2001,Xi'an, China, and the M.Tech. degree from Lanzhou university of technology in 2010. Her research interests lie in the areas of Information Fusion and Intelligent Transportation (ITS).

She is a Lecturer of Lanzhou University of Technology, and participated in 10 projects. She won first prize in science and technology progress of Gansu Province in 2010.Currently joined mainly in the research project are: Natural Science Foundation of Gansu Province, " visual detection, identification and tracking Research based on the sports car "; Gansu Higher operating costs of basic scientific research,"multi-speaker recognition based on audio and video feature fusion "; Natural Science Foundation of Gansu Province, " multi-Speaker Tracking based on the audio and video feature fusion ", ADB loan Lanzhou urban transport project advanced traffic control system (ATCS) project (design). Also she present some papers: " Research on moving vehicles tracking algorithm based on feature points and AKF " , " A Object Tracking Algorithm Based on the Points of Feature Fusion " , " The maneuvering target tracking based on the improved "current" statistical model and AKF " , and so on.



WEI LI , born in Xinyang, He Nan province of China in 1982. Received bachelor degree in Electronic and Information Engineering from Harbin Engineering University, China in 2007, and now he is a master candidate in Signal and Information Processing in Lan Zhou university of technology, China. His research interests lie in the areas of information fusion theory and application, multi-person tracking.

He is an engineer, and joined in The Science Foundation of Gansu Province and The Graduate Supervisor Foundation of Education Department of Gansu Province. also he present some papers: "Investigation of a high-precision algorithm for adaptive particle filtering"; "Object Tracking Method Based on Multi-feature Fusion"; "Speaker Tracking Based on Regularized Particle Filter", and so on.

Safety Separation Assessment in Free Flight Based on Conflict Area

Zhang Zhaoning

(Civil Aviation University of China / College of Air Traffic Management, Tianjin, China)

Email:zzhaoning@263.net

Sun Chang, Zhou Peng

(Civil Aviation University of China / College of Air Traffic Management, Tianjin, China)

Email:sc45668@sina.com, wind-diao@qq.com

Abstract—Free flight is an effective way to solve the congestions of air traffic flow. In order to guarantee the flight security, there is great significance to study on collision risk assessment in free flight. This paper applies the idea of collision risk of fixed route based on conflict area for reference, firstly designs a conflict area and establishes the collision risk model in free flight on the basis of it, and then gives the calculation of the parameters in the model. The model considers the influence of communication, navigation and surveillance performance to the probability of overlap and introduces the probability of the failure of controller's monitoring. The numerical example shows that the model can evaluate the collision risk in free flight effectively. The inverse problem of collision risk model is explored, then the minimum safety separation in free flight is got by simulating, and the advice of reduce the separation is given.

Index Terms—free flight, conflict area, collision risk

I. INTRODUCTION

Under the current air traffic control mode, the flight route of the civil aircraft is set up according to the radio beacon limited by ground-based navigation system. Since these facilities can not be established in any place, the aircraft usually can not choose the most direct route to the destination, so that the utilization of the airspace is not enough and the world's air routes are increasingly congested. To solve this problem, an American named William . Hatton proposed the idea of free flight in 1965, which transferred the control of aviation from the ground to the sky, so that the pilots can choose their own route to solve the traffic congestion. On October 1995, the U.S. Radio Technical Commission for Aeronautics (RTCA) defined the free flight formally as^[1,2]: "...a safe and efficient flight operating capability under instrument flight rules (IFR) in which the operators have the freedom to select their path and speed in real time. Air traffic

restrictions are only imposed to ensure separation, to preclude exceeding airport capacity, to prevent unauthorized flight through Special Use Airspace, and to ensure safety of flight. Restrictions are limited in extent and duration to correct the identified problem. Any activity which removes restrictions represents a move toward free flight." In order to guarantee the flight security, it is necessary to study the collision risk in free flight environment.

There are many results in the study of collision risk both at home and abroad, divided into non-free flight and free flight. In non-free flight, the best known is the Reich model which was established in the 1960s on the analysis of long-range air traffic in the longitudinal, lateral and vertical direction respectively, and the model is mainly applied to the calculation of the relationship between collision risk and the interval^[3]. However, the Reich model is not suitable for the collision risk calculation of cross route, so some researchers have proposed the collision risk assessment methods of cross route, and the collision model based on conflict area is established in [4]. In free flight, [5] presented the estimation methods of collision probability in free flight, and used the Monte Carlo method and examples to analyze. [6] used the fault tree analysis method to establish the reduced aircraft separation risk assessment model (RASRAM), the model has a quantified analysis between the relationship of reducing the security interval and collision risk. [7] studied the collision risk assessment model under the route in free flight, and used the Monte Carlo method to simulate. The study of collision risk assessment in free flight in the domestic currently is still in its infancy stage, [8] analyzed the common models in free flight and the application scope, merits and drawbacks of each model, which put forward the tendencies of the research of collision risk assessment in free flight in the future. [9] proposed the effect factors of aircraft positioning error in free flight, and the effect value of each important factor on collision risk was computed respectively.

At present, the collision risk model in free flight is mainly based on ideas of parallel routes, but the flight route is multi-directional in free flight, it is more likely to have cross conflict between two aircraft, so it is more

Based on National Science Foundation "Safety Assessment of Airborne Separation for Free Flight based on Stochastic Differential Equation" (No.71171190)

Based on the National High Technology Research and Development Program of China (863 Program-2006AA12A113)

objective to use the idea of collision risk in cross route. The influence of communication, navigation and surveillance (CNS) performance to the probability of overlap is considered, the probability of the failure of controller's monitoring is introduced and finally the collision risk model in free flight is established. It has important study significance to guarantee the flight security.

II. ESTABLISHMENT OF COLLISION RISK MODEL BASED ON CONFLICT AREA IN FREE FLIGHT

The model mainly uses the collision risk method of fixed route based on conflict area, which firstly designs a conflict area, secondly establishes the collision risk model in free flight, and then gives the calculation of each parameter in the model.

A. Design of Conflict Area

Free flight can not only reduce the flight time of the airline and reduce fuel consumption, but also obtain more amount of flight because of the full use of airspace. However, the increase in the number of flight and the multidimensionality of the path of free flight also increase the likelihood of flight conflict, the aircraft in the given airspace may have the cross conflict with the other aircraft, as shown in Figure 2.

In the fixed cross route, a conflict area around the route intersection is set, the controller guarantees the flight security by means of control that makes sure two planes not appear simultaneously in the same region. In order to establish the collision risk model based on conflict area, we must first understand three basic concepts: critical volume of collision, circular protected area (CPA) and conflict area^[4].

Critical volume of collision is a cylinder that its radius is the sum of the radius of the two planes and its height is the half of the sum of the height of the two planes; circular protected area uses the lateral separation standard to determine the lateral separation point and draws the circle with the radius of the distance between the lateral separation point and the route intersection; the rectangular area determined by the circle and the intersection of two intersecting routes is the conflict area, as the shadow shown in Figure 1.

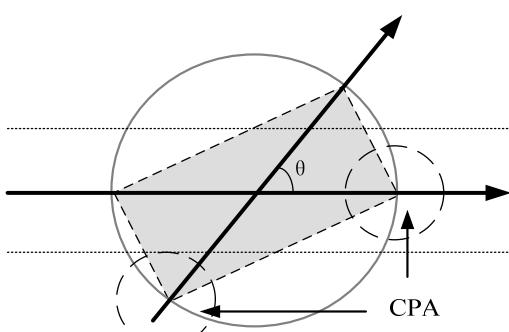


Figure 1. Conflict Area of Cross Route

As the free flight environment is not related to the specific route, so design a similar conflict area refer to cross route, as shown in Figure 3. Where, d_1 denotes the

distance between the intersection and the lateral separation point, S_y denotes the minimum of lateral separation, θ denotes the angle of the cross conflict.

B. Establishment of Collision Risk Model

In order to establish the collision risk model in free flight, the following principal assumptions are made: 1) the location of the aircraft is independent mutually; 2) the effect of the weather or other factors is not considered; 3) the horizontal and vertical position of the aircraft is independent mutually; 4) the two conflict aircraft are of the same kind, the case of adjacent planes is considered; 5) the controllers only monitor the aircraft in free flight, not implement control before the short term conflict alert alarms.

The two aircraft in the conflict area at the same time have the collision risk, the model studies two cases by the method of weight: the controller operates normally and fails to monitor. The calculation of collision risk is as follow:

$$CR = 2 \times VOP \times NP \times \{(1-\alpha) \times HCP + \alpha \times NHCP\} \quad (1)$$

Here, CR denotes the collision risk, with the number of fatal accidents per flight hour to represent; VOP denotes the probability of the vertical overlap in the same flight level; NP denotes the average number of aircraft passed the intersection per flight hour; α denotes the proportion of the planes through intersection to all the planes through the intersection in case of the failure of controller's monitoring; HCP denotes the probability of horizontal overlap in case of controller's monitoring; $NHCP$ denotes the probability of horizontal overlap in case of the failure of controller's monitoring.

According to some reference data^[10], the probability of the failure of controller's monitoring is very small, the probability of horizontal overlap in case of failure is also small, so $\alpha \times NHCP$ is relatively much smaller than others and can be ignored. Thus the assessment model can be further approximately simplified as follow:

$$CR = 2 \times VOP \times NP \times (1-\alpha) \times HCP \quad (2)$$

Here, NP can be obtained from the flight data, the calculation of VOP , HCP and α will be analyzed in the following.

1) Calculation of the probability of vertical overlap

Most of the collision risk models mainly consider the influence of navigation performance to collision risk, a few consider the influence of communication and surveillance performance. In free flight environment, we gradually get out of specific equipment requirements of CNS, but give requirements from its performance which can be achieved. Here, from the positioning error the plane caused by CNS performance when flying, the probability of vertical overlap under the specified intervals is calculated.

General the yaw error of the aircraft caused by CNS performance meets the normal distribution^[11]. Assume

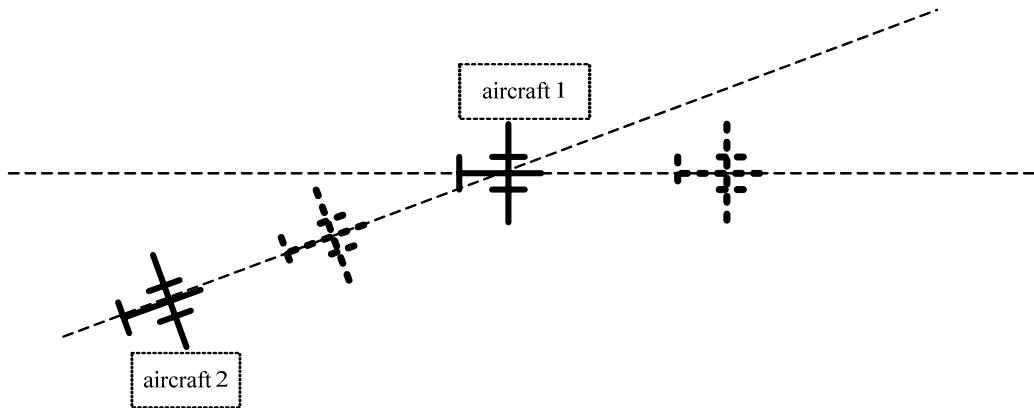


Figure 2. Cross Conflict in Free Flight

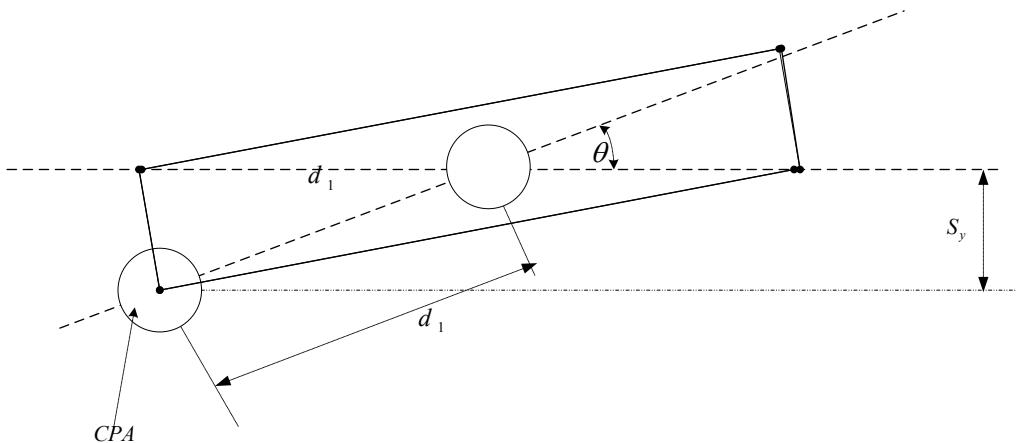


Figure 3. Circular Protected Area and Conflict Area

that in free flight the position error of vertical direction caused by CNS performance meets the normal distribution of $N(0, \sigma^2)$. Further assume that aircraft position errors respectively caused by CNS are independent mutually and all meet the normal distribution: $N_C(0, \sigma_C^2), N_N(0, \sigma_N^2), N_S(0, \sigma_S^2)$. So it is easy to deduce:

$$\sigma^2 = \sigma_C^2 + \sigma_N^2 + \sigma_S^2 \quad (3)$$

The probability density function of flight vertical collision risk can be expressed as follow:

$$f(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z^2}{2\sigma^2}\right) \quad (4)$$

Now only consider the yaw caused by CNS performance environment, assumed that the deviation to the above of expected route is positive, the below to be negative. Suppose that the flight time of the aircraft flying in free flight airspace is T , and T is evenly divided into n , $n=T/t$. The vertical deviation distance of the aircraft caused by CNS performance z in every t period meets $z \sim (0, \sigma^2)$. z_{i1}, z_{i2} respectively denotes the vertical deviation distance of the first and the second aircraft in the i th ($i=1, 2, \dots, n$) t period, the total vertical deviation distance in flight time T can be expressed as

$$\begin{aligned} Z_1 &= \sum_{i=1}^n z_{i1} \\ Z_2 &= \sum_{i=1}^n z_{i2} \end{aligned} \quad (5)$$

Thus the vertical distance Z between the two aircraft in flight time T is $Z=d+Z_1-Z_2$, among which d denotes the initial vertical distance between the two aircraft.

Because $Z_1 \sim N_1(0, n\sigma^2)$, $Z_2 \sim N_2(0, n\sigma^2)$, the following formulas can be deduced:

$$\begin{aligned} Z &= d + Z_1 - Z_2 \sim d + N(0, 2n\sigma^2) \\ &= N(d, 2n\sigma^2) \end{aligned} \quad (6)$$

The formula (4) shows that the probability density function of vertical distance between aircraft is as follow:

$$f(z) = \frac{1}{2\sqrt{n\pi}\sigma} \exp\left[-\frac{(z-d)^2}{4n\sigma^2}\right] \quad (7)$$

Where the vertical collision risk caused by CNS performance in free flight can be expressed as follow:

$$P_z = \int_{-S_z}^{S_z} f(z) dz \\ = \int_{-S_z}^{S_z} \frac{1}{2\sqrt{n\pi}\sigma} \exp\left[-\frac{(z-d)^2}{4n\sigma^2}\right] dz \quad (8)$$

where S_z denotes the given vertical separation between the two aircraft in free flight.

In 1994, the required navigation performance (RNP) in ICAO RNP Manual^[12] is defined as: when the aircraft operates in a certain route, airspace or area, RNP is determined by the value to achieve the expected navigation performance accuracy at least 95% of flight time. The definition of required communication performance (RCP) and required surveillance performance (RSP) is similar. Since under the required security interval, the vertical yaw error caused respectively by CNS performance, once the CNS performance environment is determined, it is necessary to ensure that 95% of flight time is in the specified accuracy.

Assume that CNS performance in free flight is RNP n_1 , RCP n_2 , RSP n_3 , and then the following relationship is obtained^[13]:

$$\int_{-n_1}^{n_1} \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{z^2}{2\sigma_n^2}\right) dz = 0.95 \\ \int_{-n_2V}^{n_2V} \frac{1}{\sqrt{2\pi}\sigma_c} \exp\left(-\frac{z^2}{2\sigma_c^2}\right) dz = 0.95 \quad (9) \\ \int_{-n_3V}^{n_3V} \frac{1}{\sqrt{2\pi}\sigma_s} \exp\left(-\frac{z^2}{2\sigma_s^2}\right) dz = 0.95$$

σ_n , σ_c , σ_s respectively denotes the variance component of deviation distance caused by navigation, communication and surveillance performance in the vertical direction, V denotes the vertical speed of the two aircraft in free flight.

We can deduce that:

$$\sigma_n = 0.5102n_1 \\ \sigma_c = 0.5102n_2V \\ \sigma_s = 0.5102n_3V \quad (10)$$

So the standard deviation σ of the vertical deviation distance in free flight can be expressed as follow:

$$\sigma^2 = \sigma_n^2 + \sigma_c^2 + \sigma_s^2 \\ = 0.2063(n_1^2 + n_2^2V^2 + n_3^2V^2) \quad (11)$$

That is

$$\sigma = \sqrt{0.2063(n_1^2 + n_2^2V^2 + n_3^2V^2)} \quad (12)$$

Put (12) into (8), we can deduce that

$$P_z = \int_{-S_z}^{S_z} f(z) dz \\ = \int_{-S_z}^{S_z} \frac{1}{2\sqrt{n\pi}\sqrt{0.2063(n_1^2 + n_2^2V^2 + n_3^2V^2)}} \\ \times \exp\left[-\frac{(z-d)^2}{4n\times 0.2063(n_1^2 + n_2^2V^2 + n_3^2V^2)}\right] dz \quad (13)$$

The result of P_z is VOP .

2) Calculation of the probability of horizontal overlap in case of controller's monitoring

There are already some literatures to calculate the probability of horizontal overlap in the domestic^[14], here we use the idea to deduce the probability of horizontal overlap in case of controller's monitoring.

R_{min} denotes the actual distance of the two aircraft in the nearest point, R_{col} denotes the size of the aircraft fuselage, then

$$HCP = P(R_{min} < R_{col}) \quad (14)$$

$$\text{Where } R_{min} = \left| \frac{t_0 V_1 V_2 \sin \theta}{\Delta V} \right|.$$

Because $\hat{x}_0 = x_0 + \varepsilon$, $x_0 = V_2 t_0$, so

$$R_{min} = \left| \frac{(\hat{x}_0 - \varepsilon) V_1 \sin \theta}{\Delta V} \right| \quad (15)$$

Then we can deduce that

$$HCP = P(R_{min} < R_{col}) = P\left(\left| \frac{(\hat{x}_0 - \varepsilon) V_1 \sin \theta}{\Delta V} \right| < R_{col}\right) \\ = P\left(\left| (\hat{x}_0 - \varepsilon) \right| < R_{col} \left| \frac{\Delta V}{V_1 \sin \theta} \right| \right) \\ = P\left(\left| (\hat{x}_0 - \varepsilon) \right| < A \times R_{col} \right) \\ = P\left(-A \times R_{col} < (\hat{x}_0 - \varepsilon) < A \times R_{col}\right) \quad (16)$$

Where, $A = |\Delta V/V_1 \sin \theta|$, V_1 denotes the speed of the first aircraft, ΔV denotes the vector of $V_1 - V_2$, θ denotes the angle between the two cross aircraft.

So the probability of horizontal overlap of the aircraft in case of controller's monitoring on the cross route is

$$HCP = \int_{-x}^x \int_{\hat{x}_0 - A \cdot R_{col}}^{\hat{x}_0 + A \cdot R_{col}} f(\varepsilon) g(\hat{x}_0) d\varepsilon d\hat{x}_0 \quad (17)$$

Where, \hat{x}_0 denotes the estimated distance between aircraft 2 and the intersection when aircraft 1 is located in the intersection, ε denotes the error of \hat{x}_0 .

① Calculation of probability density function $g(\hat{x}_0)$

Because the two aircraft can not appear in the conflict area at the same time, so \hat{x}_0 should meet $\hat{x}_0 \geq 2d_1$.

Where d_1 denotes the distance between the intersection and the lateral separation point, it is calculated as follow:

$$d_1 = 60 \times \frac{180}{\pi} \times \arcsin \left[\frac{\sin \left(\frac{\pi S_y}{60 \times 180} \right)}{\sin \left(\frac{\pi \theta}{180} \right)} \right] \quad (18)$$

S_y denotes the minimum of given lateral separation in free flight.

To simplify the calculations, assumed that \hat{x}_0 meets the uniform distribution between $(2d_1, 2d_1 + D)$, D is a constant, then

$$g(\hat{x}_0) = \begin{cases} \frac{1}{D}, & 2d_1 < x < 2d_1 + D \\ 0, & \text{others} \end{cases} \quad (19)$$

According to the relative data in non-free flight, $D=250$, then

$$g(\hat{x}_0) = \begin{cases} \frac{1}{250}, & 2d_1 < x < 2d_1 + 250 \\ 0, & \text{others} \end{cases} \quad (20)$$

② Calculation of probability density function $f(\varepsilon)$

Because the probability of collision is usually small and heavily depends on the tail of the distribution ε , so assumed that ε meets the double exponential distribution:

$$f(\varepsilon) = \frac{1}{2\lambda} e^{-\frac{|\varepsilon|}{\lambda}} \quad (21)$$

We can deduce that

$$HCP = \frac{1}{500} \int_{2d_1}^{2d_1+250} \int_{\hat{x}_0 - A \cdot R_{col}}^{\hat{x}_0 + A \cdot R_{col}} \frac{1}{\lambda} e^{-\frac{|\varepsilon|}{\lambda}} d\varepsilon d\hat{x}_0 \quad (22)$$

Calculate it by integral, the result is as follow:

$$\begin{aligned} HCP &= \frac{\lambda}{500} \left[\exp \left(\frac{AR_{col}}{\lambda} \right) - \exp \left(-\frac{AR_{col}}{\lambda} \right) \right] \\ &\quad \times \left[\exp \left(-\frac{2d_1}{\lambda} \right) - \exp \left(-\frac{2d_1 + 250}{\lambda} \right) \right] \end{aligned} \quad (23)$$

Where, $\lambda=RNP/2.996$, calculated by the required navigation performance in free flight.

3) Calculation of the probability in case of the failure of controller's monitoring

The large domestic transport aircraft is fitted with airborne collision avoidance system, which effectively prevents the air collision accidents. In free flight environment, the presence of collision avoidance system is essential, and much more important and dependent than non-free flight. The collision avoidance system includes two kinds: ground collision avoidance system ——Short Term Conflict Alert (STCA) and airborne collision

avoidance system——Traffic alert and Collision Avoidance System / Airborne Collision Avoidance System (TCAS / ACAS). The former is for controllers, which sends an alarm message to the controller before flight interval is less than the standard interval, the controller takes appropriate measures to make the aircraft return to the normal safe interval; if the control intervention fails, the interval continues to reduce, until the TCAS triggers an alarm message to the pilot, and provides the appropriate Traffic Advisory (TA) and Resolution Advisory (RA). Since in free flight, the controllers only implement monitoring to the aircraft, not control before the ground collision avoidance system alerts, so we assume that the probability of the failure of controller's monitoring is equivalent to the failure probability of ground collision avoidance system. The flow of the aircraft within a certain time meets uniform distribution, so in case of the failure of monitor, the proportion of the planes through intersection to all the planes through the intersection can be calculated as the failure probability of ground collision avoidance system.

[10] calculated the failure probability of TCAS and STCA by Fault Tree Analysis (FTA) and analyzed the system reliability successfully. The fault tree also takes human factors and CNS performance into account, responses the free flight environment comprehensively.

FTA is one of the primary analysis methods in safety engineering system. In the design of the system, FTA analyzes a variety of factors (including hardware, software, environment, human factors) that may result in system failure, draws logic diagram (ie, fault tree), and then determines the possible combinations or probability of occurrence that make the system failure. People can take appropriate measures to improve the reliability of the system. The basic principle of STCA is: if the system determines the current interval of the two aircraft is less than the separation standard, the system considers that the two aircraft have a conflict, then it issues a warning signal to the controller; if the system determines the two aircraft has a probability of dangerous approach according to the current position, but not yet have a conflict, the system considers that the two aircraft have a potential conflict, then it issues a early warning signal to the controller.

According to the working principle of FTA and STCA, assume "S" as the top event, which means the failure of the STCA system, use $A_i (i=1,2,3)$ as the middle events, $X_j (j=1,2,3)$ as the basic events, then we can get the fault tree of the failure of the STCA system, as shown in Figure 4.

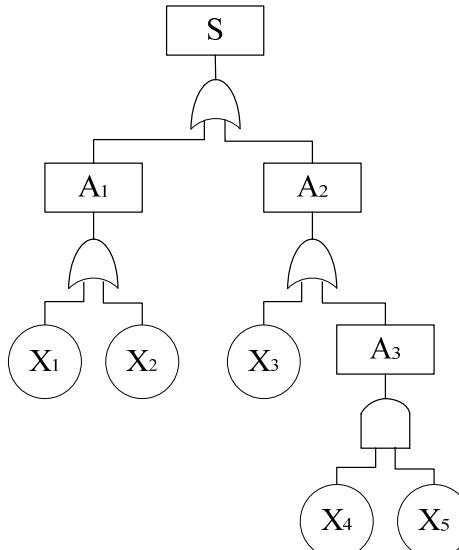


Figure 4 The Fault Tree of the Failure of the STCA System

The specific event and the probability of the occurrence of each event is shown in Table I.

TABLE I
NAME OF THE EVENT AND THE PROBABILITY OF EACH EVENT

SYMBOL	EVENT	PROBABILITY
A_1	the pilots don't take the conflict disentangle strategies	$P(X_1) + P(X_2)$
A_2	the pilots take the conflict disentangle strategies falsely	$P(X_3) + P(A_3)$
A_3	the pilots misunderstand the conflict disentangle strategies	$P(X_4) \cdot P(X_5)$
X_1	the response of the ground workers delay	3.0×10^{-3}
X_2	failure of the communication	1.0×10^{-7}
X_3	the ground workers take the strategies falsely	4.0×10^{-4}
X_4	the pilots understand the conflict disentangle strategies falsely	4.0×10^{-4}
X_5	the ground workers don't correct timely	2.0×10^{-1}

According to Figure 3, use the basic knowledge of FTA, we can get the probability of occurrence of the top event is:

$$\begin{aligned}
 P(S) &= P(A_1) + P(A_2) \\
 &= P(X_1) + P(X_2) + P(X_3) + P(A_3) \\
 &= P(X_1) + P(X_2) + P(X_3) + P(X_4) \cdot P(X_5)
 \end{aligned} \tag{24}$$

According to the probability of each event shown in Table 1, by the formula (24) we can calculate that the failure probability of the ground collision avoidance system is $p=p(S)=3.48 \times 10^{-3}$, it can be served as α , that is $\alpha=3.48 \times 10^{-3}$.

III. NUMERICAL EXAMPLE

However, free flight is still in the design and experiment stage now, this paper uses a part of Shanghai control area as a collision risk assessment area in free flight. Select the free flight area including routes G204 and A470, and the aircraft type is B747-400. Here, we fix the CNS performance environment of the plane (RNP1, RCP400, RSP2), the average number of aircraft passed the intersection per flight hour is 15. The reference data is shown in Table II.

TABLEII
RELATIVE PARAMETER VALUE IN FREE FLIGHT (UNIT: NMILE, KT)

PARAMETER	n	n_1	n_2	n_3
VALUE	30	1	400	2
PARAMETER	V	d	S_z	S_y
VALUE	2.5	0.165	0.28	1.62
PARAMETER	R_{col}	V_1	θ	α
VALUE	0.038	495	88	3.48×10^{-3}

Calculate the probability of vertical overlap VOP according to (13), calculate the probability of horizontal overlap HCP according to (23), then put $\alpha=3.48 \times 10^{-3}$, $NP=15$ into (2), we can get that $CR=6.37 \times 10^{-9}$ times / per flight hour.

The evaluate result shows that the collision risk under the given conditions is 6.37×10^{-9} times / per flight hour, meets the safety level of 1.5×10^{-8} times / per flight hour set by ICAO. It shows that the collision risk assessment model in free flight can be successfully used to estimate collision risk, and demonstrates that the assessment model is feasible.

IV. CALCULATION OF THE MINIMUM SAFETY SEPARATION

A. the Improved Model of Safety Separation in Free Flight

The collision risk model based on conflict area in free flight mentioned before is

$$CR = 2 \times VOP \times NP \times (1 - \alpha) \times HCP$$

Here,

$$\begin{aligned}
 VOP &= \int_{-S_z}^{S_z} \frac{1}{2\sqrt{n\pi} \sqrt{0.2063(n_1^2 + n_2^2 V^2 + n_3^2 V^2)}} \\
 &\quad \times \exp\left[-\frac{(z-d)^2}{4n \times 0.2063(n_1^2 + n_2^2 V^2 + n_3^2 V^2)}\right] dz \\
 HCP &= \frac{\lambda}{500} \left[\exp\left(\frac{AR_{col}}{\lambda}\right) - \exp\left(-\frac{AR_{col}}{\lambda}\right) \right] \\
 &\quad \times \left[\exp\left(-\frac{2d_1}{\lambda}\right) - \exp\left(-\frac{2d_1 + 250}{\lambda}\right) \right]
 \end{aligned}$$

Define

$$T = 2 \times NP \times (1 - \alpha) \times \frac{\lambda}{500} \left[\exp\left(\frac{AR_{col}}{\lambda}\right) - \exp\left(-\frac{AR_{col}}{\lambda}\right) \right]$$

$$\times \int_{-S_z}^{S_z} \frac{1}{2\sqrt{n\pi} \sqrt{0.2063(n_1^2 + n_2^2 V^2 + n_3^2 V^2)}} dz$$

$$\times \exp\left[-\frac{(z-d)^2}{4n \times 0.2063(n_1^2 + n_2^2 V^2 + n_3^2 V^2)}\right] dz$$

Then

$$CR = T \times \left[\exp\left(-\frac{2d_1}{\lambda}\right) - \exp\left(-\frac{2d_1 + 250}{\lambda}\right) \right] \quad (25)$$

Define the given safety level as P_{Given} ,

$$P_{Given} = CR \quad (26)$$

We can get the improved model of safety separation in free flight is

$$\left[\exp\left(-\frac{2d_1}{\lambda}\right) - \exp\left(-\frac{2d_1 + 250}{\lambda}\right) \right] = \frac{P_{Given}}{T} \quad (27)$$

When the other conditions are known, for a given safety level P_{Given} , the safety separation d_1 in free flight can be calculated by relevant algorithm and formula (27).

B. Algorithm

Since formula (27) is a complex exponential equation which is non-linear, we can't get the answer directly but by the iterative algorithm.

Iterative algorithm is a basic method to solve the problem by computer, which can use the characteristics of computers that includes the fast computing speed and suitable for repetitive operation. It allows the computer to repeat a set of instructions, in each execution of this instruction, a new value can be induced from its original value of the variable.

Here, in order to calculate the safety separation d_1 , we use the equal step iterative algorithm. Take the lateral separation of parallel route as the lateral separation in free flight and implement the algorithm based on the initial value. The flow chart of algorithm is shown in Figure 5. Where, ε denotes the accuracy of control, p_t denotes the safety level, d_0 denotes the initial value of the exponential equation, t' denotes the step length.

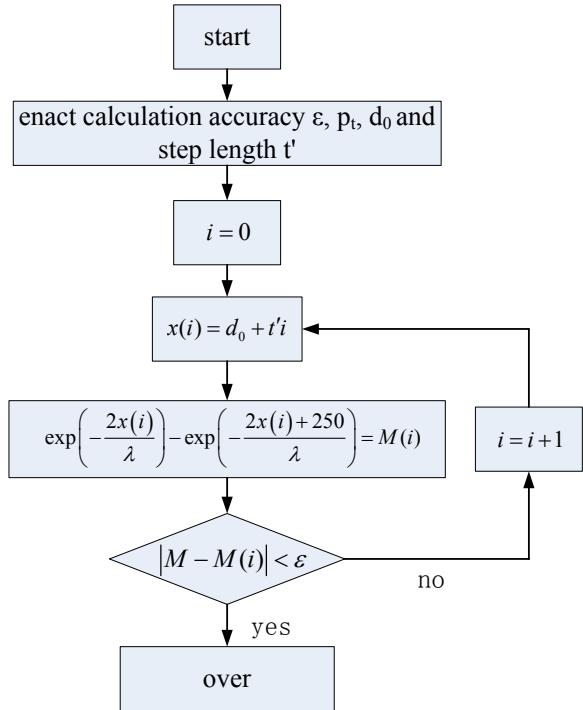


Figure 5 The Flow Chart of Algorithm

Consider the calculation of $x(i)$ as d_1 that meets the exponential equation, then the minimum safety separation in free flight can be obtained from this.

C. Result and Analysis

Use MATLAB to simulate, the specific values are the same as the numerical example below, the simulation result is shown in Table III.

TABLE III
SIMULATION RESULT

parameters			simulation result
ε	p_t	t'	d_1/n mile
5.0×10^{-21}	1.5×10^{-8}	0.01	14.6

The simulation result shows that: in a given safety level, the safety separation in free flight calculated by equal step iterative algorithm is 14.6n mile. From the relevant provisions of flight separation, the minimum separation standard of aircraft on the cross route is the distance measured by the distance measure equipment with the longitudinal separation of 40km, which requires the two aircraft fly on the same flight level and the crossing angle is less than 90°. The simulation result is smaller than the required standard of flight separation on the fixed route, it is because that the required accuracy in free flight is larger than traditional fixed route flight, the model and algorithm presented in this paper is based on CNS performance of Next Generation Civil Aviation Transportation System (NGCATS). It shows that in the case of ensuring the safety level set by ICAO and

meeting the required CNS performance, the safety separation in free flight can be appropriately reduced.

V. CONCLUSION

This paper presents a preliminary study on collision risk model in free flight. We have applied the idea of collision risk of fixed route based on conflict area for reference, by considering the influence of CNS performance to the probability of overlap and introduce the probability of the failure of controller's monitoring and at last established the collision risk model in free flight. The numerical example shows that the model can evaluate the collision risk in free flight effectively. The inverse problem of collision risk model is explored, then the minimum safety separation in free flight is got by simulating, and the advice of reduce the separation is given. But the probability of horizontal overlap in case of the failure of controller's monitoring is not considered in the model, which is a future research direction.

REFERENCES

- [1] Final Report of RTCA Task Force 3-Free Flight Implementation[R]. Washington DC: RTCA Inc,1995.
- [2] Special Report on Free Flight[R].Aviation Week and Space Technology, 1995.
- [3] Reich.P.G. Analysis of long-range air traffic systems: separation standards I II III [J]. Journal of the Institute of Navigation. 1966, 19(1):88-98, 1966, 19(2)169-186, 1966, 19(3):331-347.
- [4] Anderson D, Lin X G.Collision risk model for a crossing track separation methodology [J].Journal of Navigation, 1996, 49(3):337-349.
- [5] Paielli, R. A, Erzberger, H. Conflict probability estimation for free flight[R].NASA Ames Research Center, Moffett Field, CA: NAS 1.15:110411; A-962310 ,1996.
- [6] Cassell Rick, Smith Alex, Shepherd Roger. Risk assessment model for free flight - terminal area reduced separation[C]//Proceedings of the 1996 15th AIAA/IEEE Digital Avionics Systems Conference, Atlanta: Piscataway, 1996: 73-80.
- [7] Everdij, Mariken H.C., Blom, Henk A.P., Bakker, Bert G.J. Modelling lateral spacing and separation for airborne separation assurance using Petri nets [J]. Journal of Simulation, 2007, 83(5):401-414.
- [8] Lu Tingting, Zhang Zhaoning, Liu jimin. Study on Safety Assessment of Free Flight Collision Risk[J]. Aeronautical Computing Technique, 2010,40(6):25-29.
- [9] Cai Ming, Zhang Zhaoning, Wang Lili. Research on Collision Risk in Free Flight[J]. Aeronautical Computing Technique,2011, 41(1): 51-56.
- [10] Sui Dong. New method for safety assessment of parallel routes[J].Transactions of Nanjing University of Aeronautics & Astronautics, 2009,26(1):36-43.
- [11] Geert Moek, Edward Lutz, William Mosberg.Risk Assessment of RNP 10 and RVSM in the South Atlantic Flight Identification Regions [R]. Annapolis: ARINC, 2001.
- [12] Doc 9613-AN/937, Manual on Required Navigation Performance[S].ICAO, Second Edition, 1999.
- [13] Zhang Zhaoning, Wang Lili, Li Dongbin. Introduction to Security Assessment on Airborne Separation[M].Science Press, Beijing, 2009.
- [14] Zhang Zhaoning, Zhang Xiaoyan. Research and Application on Collision Risk of Crossing Tracks. Aeronautical Computing Technique, 2007,37(2):1-4.

Zhang Zhaoning was born in Hebei Province in China on July 1964. He received the BS, MS and Ph.D degrees from Hebei Normal University, Nankai University and Tianjin University respectively. He is now a professor of College of Air Traffic Management in Civil Aviation University of China . His current research interests include air traffic plan and management, traffic information engineering and control, signal and information processing.

Sun Chang was born in Liaoning Province in China on December 1986. She received the BS degree from Civil Aviation University of China. She is now a postgraduate of College of Air Traffic Management in Civil Aviation University of China. Her current research interests include air traffic plan and management, emphasis on security assessment on airborne separation and air traffic flow management.

Zhou Peng was born in Henan Province in China on January 1987. He received the BS degree from Guilin University. He is now a postgraduate of College of Air Traffic Management in Civil Aviation University of China. His current research is mainly on security assessment on airborne separation.

Design of Three-axis ED Milling Machine Based on the PMAC Motion Card

Fei Wang

College of Electromechanical Engineering, China University of Petroleum, Dongying, China
Email: wangfei208@sina.com

Yonghong Liu*, Zhili Chen, Renjie Ji, Xiaojie Tian, Zengkai Liu

College of Electromechanical Engineering, China University of Petroleum, Dongying, China
Email: liuyh@upc.edu.cn, chenzhili_218@163.com, jirenjie202@yahoo.cn,
tianxj20050101@163.com, liuzengk@163.com

Abstract—This paper presents a new three-axis ED milling machine, which is based on the PMAC motion card. The machine is composed of the motion system, the control system, the working fluid supplying system and the EDM pulse generator. First, the motion system, the working fluid supplying system, the hardware of the control system, and the EDM pulse generator have been built, then the software system of the industrial computer has been developed based on the Visual Basic 6.0 in order to control the machine and monitor the status of the machine. Using the PMAC motion card, the servo control program, the decoding program and the PLC program have been programmed. The software system of the EDM pulse generator is implemented based on a good real time operating system called μ C/OS-II, and the software system is composed of three main tasks, namely the data processing task, the communication task and the monitoring task. The milling experiments have been carried out on the ED milling machine based on the PMAC motion card, and the results show that the ED milling machine is steady and reliable when milling with three-axis.

Index Terms— ED milling; PMAC; EDM pulse generator; real time operating system

I. INTRODUCTION

Electrical discharge machining (EDM) is a processing method of transforming electric energy into heat energy with the pulse voltage added between the workpiece and the electrode, so it can remove the material from the workpiece [1]. As a non-traditional material removal process, EDM has long been employed in manufacturing of molds, as well as in automotive, aerospace and surgical components due to its unique merit of machining by thermal-mechanical effect regardless of the hardness of the materials [2]. EDM has been developed in the late 1940s, and it does not make direct contact between the workpiece and the electrode, as a result, it can eliminate the effect of the mechanical stresses, chatter and vibration problems.[3-5].

Electrical discharge (ED) milling is an evolution of CNC contouring EDM [6-7]. A rotating cylindrical electrode follows a programmed path in order to obtain the desired shape of a part, like a cutter used in

conventional computerized numerical controlled (CNC) milling. Compared to traditional sinking EDM, the use of simple electrodes in ED milling eliminates the need for customized shaped electrodes [8-11]. In the ED milling, the simple shape electrode does layer-by-layer milling to get a three-dimensional complex parts, at the same time, electrical discharges occur repeatedly to remove materials along the programmed path. According to the discharge status between the electrode and the workpiece, the control system determines the forward and withdrawal feedrate of the electrode, as a result, a dedicated ED milling control system need to be developed to control the ED milling and edit motion program. Ref. [12] used one kind of milling EDM control system, a discharge power supply provided the power for the system, the computerized numerical controller interpolated algorithm in the interrupt service routine, which was triggered in the fixed time. The feedrate was determined by a gap controller, which maintained the discharge in the gap, the interrupt service routine modified the current position of the electrode on the main axis. Ref. [13] designed a new control system for a micro-ED milling machine. The motion system was made up of three-axis rotary motor driven linear stages and a spindle motor. They were controlled by a motion controller board installed on a control PC, one data acquisition board and one analog image acquisition board were installed in the control PC, a vision system monitored the electrical machining condition, which was represented by the behavior of the gap voltage and current during the EDM process. The data acquisition board countered the number of discharge pulses in real time.

In this study, a new kind of three-axis ED milling machine based on the PMAC motion card has been designed. After building the motion system and the working fluid supplying system, the Microsoft Visual Basic 6.0 is used to build the software system of the industrial computer. The software system of the PMAC motion card is developed. Then, based on the μ C/OS-II, the EDM pulse generator software system is developed, which contains three main tasks, due to the well real time performance, the EDM pulse generator can bear more

work than the traditional pulse generator, as a result, the hardware of the system is simplified. Finally, some milling experiments have been carried out, and the new ED milling machine is proved to be stable and reliable. The article is structured as follows: Section I introduces the principle of ED milling and the research status. Section II describes the design of the motion system and the working fluid supplying. Section III describes the software system of the control system, including the industrial computer, the PMAC motion card and the EDM pulse generator. Section IV introduces some ED milling experiments carried out by the ED milling machine with different axes. Section V summarizes the article.

II. DESIGN OF THE MACHINE TOOL

The structure of the ED milling machine is shown in Fig. 1. The machine is composed of four main elements, namely the motion system, the working fluid supplying system, the control system, and the EDM pulse generator. The photograph of the ED milling machine is shown in Fig. 2.

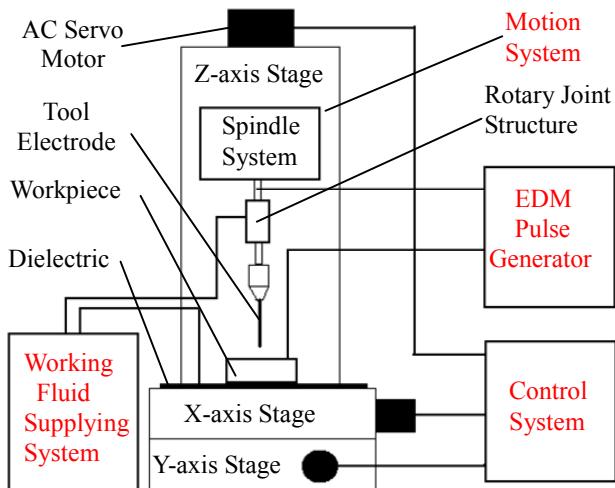


Figure 1. Structure of the ED milling machine.



Fig 2. Photograph of the ED milling machine.

A. Design of the Motion System

The photograph of the motion system is shown in Fig. 3. The motion system plays a key role in a machine tool, which will carry out the command of the control system, it is absolutely necessary. The machine tool is C-type structure in this design. The motion system is composed of the X-axis stage, the Y-axis stage, the Z-axis stage, and the spindle system. The movement is controlled by the AC servo motors, and the servo motor of the Z-axis has more power than other servo motors, as it must drive the Z-axis smoothly to overcome the gravity of the Z-axis stage. When the machine tool is shut down, a normally closed electric clutch will hold the servo motor of the Z-axis. The position loop of the machine tool is semi-closed loop, which is made of the digital incremental encoder fixed at the end of the servo motor. The spindle system provides the rotation for the electrode, which is composed of the AC motor, the AC inverter, the collecting ring, and the brush. The spindle motor fixed at the Z-axis stage is surrounded by the dielectric to prevent the discharge current from damaging the machine tool, the collecting ring and the brush will import the discharge current to the AC motor shaft. The AC motor is added to provide rotation of the electrode, with the help of the AC inverter, the speed of the AC motor is 0-3000rpm. The stage which loads the workpiece is covered with dielectric in order to prevent the discharge current from damaging the machine tool.



Fig 3. Photograph of the motion system.

B. Design of the Working Fluid Supplying System

There are three basic types of flushing in the ED

milling, and they are side flushing, electrode injection flushing and electrode suction flushing, Ref. [14] found that an optimized flushing pressure and electrode rotation can achieve an approximately constant cutting feed rate, and the electrode injection flushing and wider working range of flushing pressure can get higher material removal rate(MRR) value than the side flushing and the suction flushing. The MRR, the electrode wear rate (EWR), surface roughness (SR) and cutting feed rate all increase with increased flushing pressure, indicating that superior performance may be obtained from a higher flushing pressure during machining.

In this study, a new working fluid supplying system is designed in order to adjust the pressure and achieve the injection flushing. The system is made of the circulatory pump, the filter, the AC converter, the pipeline, and the rotary joint structure. The flushing pressure is adjusted by the circulatory pump controlled by the AC converter. The pipeline connects the pump and the rotary joint structure to transport the working fluid cleaned by the filter. The cutaway view of the rotary joint structure is shown in Fig. 4, the holder connects the rotary joint structure to the Z-axis stage to stop it from rotating. A threaded hole is designed on the shell of the rotary joint structure in order to connect the pipeline and seal the working fluid. There is a ring groove on each side of the threaded hole to hold the O-ring. In this design the maximum pressure of the working fluid is 1MPa, the maximum speed of the shaft is 3000rpm, the inner diameter of the O-ring is 20mm, the diameter of the shaft is 19.4mm, and the diameter of the ring groove is 20mm. A ring groove on the shaft is designed to store the working fluid to prevent the pressure fluctuation. The photograph of the rotary joint structure is shown in Fig. 5.

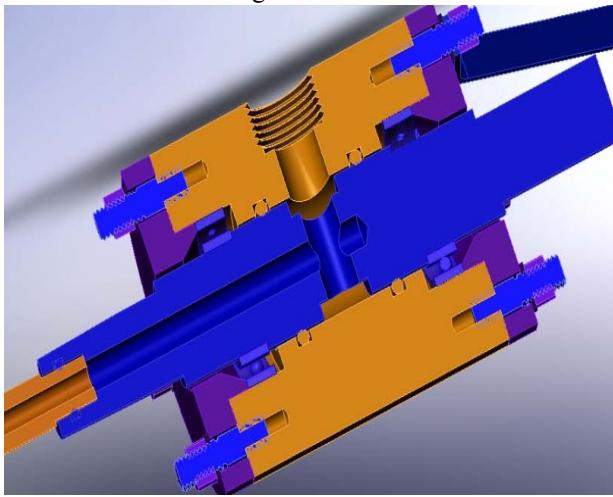


Fig 4. Cutaway view of the rotary joint structure.

III. DESIGN OF THE SOFTWARE SYSTEM

The control system structure of the ED milling machine is shown in Fig. 6. As shown in Fig. 6, the control system structure is master slave module, the industrial computer is the master, in this control system, the slave must have good real-time performance, so it is formed by the PMAC motion card and the micro control

unit (MCU) of the EDM pulse generator. The industrial computer transfers commands and files to the PMAC motion card via the Ethernet. The PMAC motion card sends the analog signal to the AC servo motor drive of the X-axis stage, the Y-axis stage, and the Z-axis stage, also it control the AC converter of the spindle and the filter system with analog signals. Some digital input-output signal is pressed by the PLC connected to the PMAC motion card. Commands such as the pulse on-time, the pulse interval time are downloaded to the MCU of the EDM pulse generator by the industrial computer via the RS232, the discharge status between the workpiece and the tool electrode is collected by the analog to digital converter, and it is send to the industrial computer through the USB. The software structure of the control system is shown in Fig. 7.



Fig 5. Photograph of the rotary joint structure.

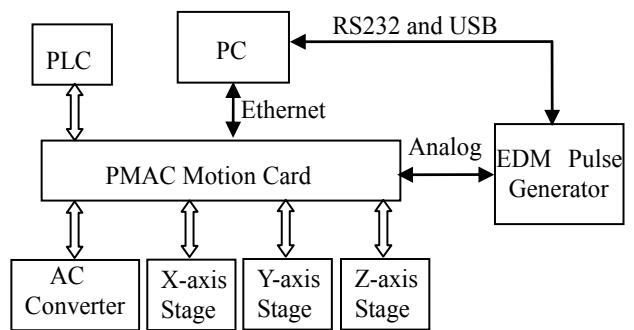


Fig 6. Structure of the control system.

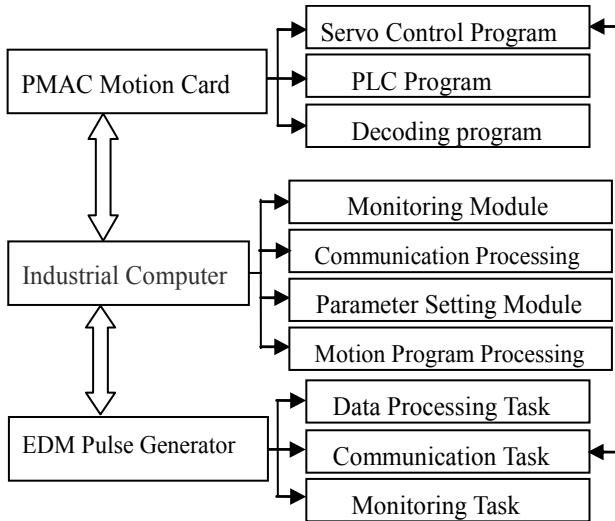


Fig. 7. Structure of the software system.

A. Software System of the Industrial Computer

The industrial computer is the master computer, which plays a key role in the control system, it must complete some important function such as the human machine interface and the management of the background task. All the programs of the master computer are developed in the Windows environment, and they are programmed based on the Visual Basic 6.0. All of the programs have less demand of real-time, as a result, the industrial computer is not able to collect the discharge status and process the motion programs. The software system is made of four main modules, the monitoring module, the communication processing module, the parameter setting module, and the motion program processing module. The monitoring module plays an important role in the human machine interface, it processes the information coming from the EDM pulse generator and the PMAC motion card, and it displays the voltage waveform, the current waveform, the status of the motion card, and the running status of the PLC, it makes the user to better understand the current status of the machine. The communication processing module establishes the communication between the master and the slave, and receives the signal from the slave. The parameter setting module sets the parameter of the PMAC motion card to finish the configuration, and it sets the parameter of the EDM pulse generator, such as the open voltage, the peak current, the pulse on-time, and the pulse interval time, the parameter setting module also determines some machining parameters according to the material of the workpiece. The motion program processing module finishes the editing of the motion program, it transforms the motion program edited by the user into the file, which can be recognized by the PMAC motion card, at the same time it examines the motion program to avoid the syntax error. Different user interfaces of the industrial computer are shown in Fig. 8.



(a) Editing mode



(b) Manual operation mode



(c) Manual operation mode

Fig. 8. Different user interfaces of the industrial computer.

B. Software System of the PMAC Motion Card

As the core element of the control system, the PMAC motion card should be competent for the control of the motion system, the AC converter, and some other function of the PLC, so it should be able to finish the real-time task. The structure of the software system of the master computer is shown in Fig. 7. The software system of the PMAC motion card is made of three programs,

namely the PLC program, the servo control program, and the decoding program. The PLC program will process the digital input-output signal of the machine. It collects the signal of the keyboard on the control panel, and controls the auxiliary equipment, such as the working fluid, and the electric clutch of the Z-axis stage. The PLC program can also achieve to control the ED milling machine with the handwheel. The servo control program determines the feed rate of the motion system according to the analog signal from the EDM pulse generator, the flow diagram of the servo control program is shown in Fig. 9. The decoding program decodes the G-code, the M-code, and the S-code in the motion program edited by the user, and then it converts those codes into the code that the PMAC motion card can recognize.

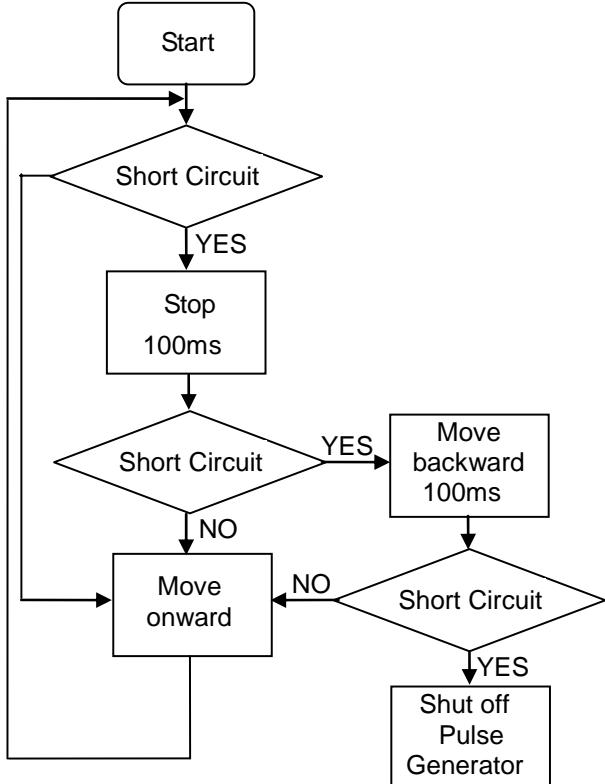


Fig 9. Flow diagram of the servo control program.

C. Software System of the EDM Pulse Generator

In this control system, in addition to the traditional functions, such as controlling the peak current, the open voltage, the pulse on-time, and the pulse interval time, the EDM pulse generator is able to realize the data processing of the discharge status and the communication function. In the MCU of the EDM pulse generator, the real time operating system μ C/OS-II has been transplanted to ensure the stability of the system. Three tasks, such as the monitoring task, the data processing task, and the communication task, are established. The monitoring task collects the discharge status through the analog to digital converter, and it completes self-diagnosis to protect the EDM pulse generator. The data processing task processes the data from the monitoring task, the feed rate of the motion system is calculated by an algorithm programmed in this task to maintain the

discharge. The communication task transports the feedrate of the motion system calculated by the data processing task to the PMAC motion card through the digital to analog converter, and it also sends those discharge status data to the industrial computer through the USB. The relationship between the three tasks is shown in Fig. 10. It can be seen that the monitoring task and the data processing task are connected by the timeout message queues to monitor the system in real-time, the data processing task and the communication task are connected by the mailbox, and the interrupt service routine is programmed in order to receive the command from the RS232.

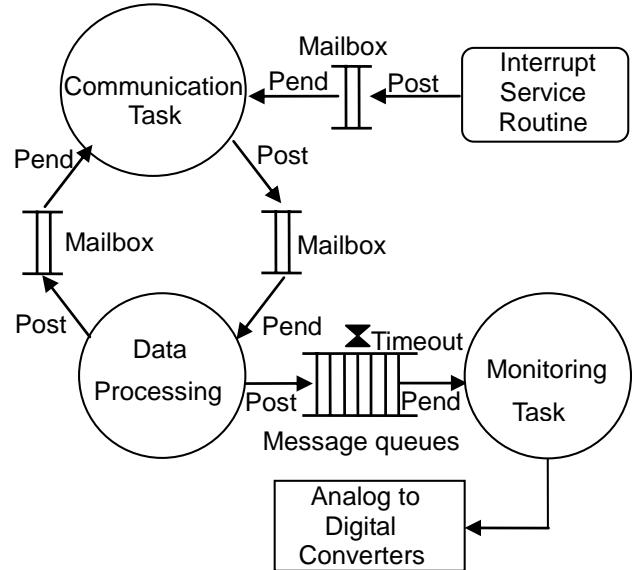


Fig 10. Relationship between the three tasks.

IV. EXPERIMENTS

Some ED milling experiments have been carried out to estimate the stability and the reliability of the designed ED milling machine. In the experiments, the tool electrode is the hollow-tube copper with the outer diameter of 10mm and the inner diameter of 5mm, the working fluid is emulsion. The photograph of the workpiece machined by the ED milling machine with different axes is shown in Fig. 11.

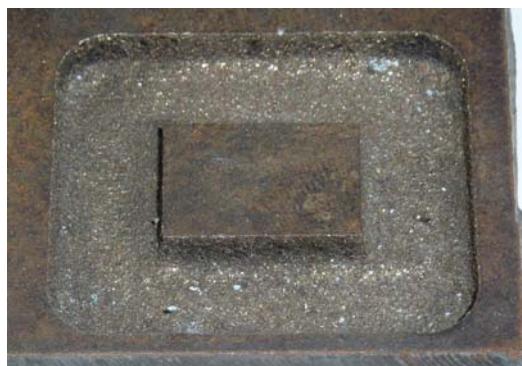
V. CONCLUSION

Based on the PMAC motion card, the three-axis ED milling machine has been designed. The motion system, the control system, the working fluid supplying system, and the EDM pulse generator are designed. The software system of the industrial computer is developed based on the Visual Basic 6.0. Using the PMAC control card, the PLC program, the servo control program, and the decoding program are developed. The real time operating system μ C/OS-II has been transplanted to ensure the stability of the system. According to this paper, the following conclusion can be drawn:

(1) The designed three-axis ED milling machine is able to machine the workpiece steadily and reliably during three-axis ED milling.

(2) In the working fluid supplying system, the AC converter is able to control the pressure of the working fluid through changing the rotation speed of the circulatory pump. The rotary joint structure can achieve electrode injection flushing, and the rotation of the electrode tool.

(3) The control system of the EDM pulse generator is able to finish multitask after the transplanting of the real time operating system due to its good real time performance. The calculation of the feed rate in the data processing task reduces the work load of the PMAC motion card, and more complex algorithm can be programmed in the data processing task.



(a) Single-axis



(b) Two-axis



(c) Three-axis

Figure 11. Photograph of the workpiece machined by the ED milling machine with different axes.

ACKNOWLEDGMENT

The authors wish to acknowledge the financial support of the Ministry of Science and Technology of the People's Republic of China (Grant No. 2009GJC60047) and National Natural Science Foundation of China (Grant No.50675225).

REFERENCES

- [1] F. Z. Han, L. Chen, D. W. Yu and X. G. Zhou, "Basic study on pulse generator for micro-EDM," *Int J Adv Manuf Technol*, vol. 33, no. 5-6, June. 2007, pp. 474-479, doi: 10.1007/s00170-006-0483-9.
- [2] F. Z. Han, Y. X. Wang and M. Zhou, "High-speed EDM milling with moving electric arcs," *Int J Mach Tools Manuf*, vol. 49, no. 1, January. 2009, pp. 20-24, doi: 10.1016/j.ijmachtools.2008.08.005.
- [3] M. A. Norliana1, S. Darius G and F. B. Md., "A review on current research trends in electrical discharge machining (EDM)," *Int J Mach Tools Manuf*, vol. 47, no. 7-8, June. 2007, pp. 1214-1228, doi: 10.1016/j.ijmachtools.2006.08.026
- [4] K.H. Ho and S.T. Newman, "State of the art electrical discharge machining (EDM)," *Int J Mach Tools Manuf*, vol. 42, no. 13, pp. 1287-1300, doi: 10.1016/S0890-6955(03)00162-7.
- [5] S. Singh, S. Maheshwari and P.C. Pandey, "Some investigations into the electric discharge machining of hardened tool steel using different," *J Mater Process Technol*, vol. 149, no. 1-3, pp. 272-277, doi: 10.1016/j.jmatprotec.2003.11.046
- [6] J. P. Kruth, B. Lauwers, and W. Clappaert. "A study of EDM pocketing," Proceedings of the ISEM, Magdeburg, Germany, May, 1992, pp.271-294.
- [7] J. P. Kruth, and P. Bleys. "Machining curvilinear surfaces by NC electro-discharge machining," *Proceedings of the Second International Conference on MMSS*, Krakow, Poland, September, 2000, vol. 4, pp. 31-40.
- [8] P. Bleys, J.-P. Kruth and B. Lauwers, "Sensing and compensation of tool wear in milling EDM," *J Mater Process Technol*, vol. 149, no. 1-3, June. 2004, pp. 139-146, doi: 10.1016/j.jmatprotec.2003.11.042.
- [9] Y. F. Chang and R. C. Hong, "Parametric curve machining of a CNC milling EDM," *Int J Mach Tools Manuf*, vol. 45, no. 7-8, June. 2005, pp. 941-948, doi: 10.1016/j.ijmachtools.2004.10.011.
- [10] Z. Y. Yu, T. Masuzawa and M. Fujino, "Micro-EDM for three-dimensional cavities - development of uniform wear method," *CIRP Ann Manuf Technol*, vol. 47, no. 1, 1998, pp. 169-172, doi:10.1016/S0007-8506(07)62810-8.
- [11] K. P. Rajurkar and Z. Y. Yu, "3D micro-EDM using CAD/CAM," vol. 49, no. 1, 2000, pp. 127-130, doi:10.1016/S0007-8506(07)62911-4.
- [12] Y. F. Chang and R. C. Hong, "Parametric curve machining of a CNC milling EDM," *Int J Mach Tools Manuf*, vol. 45, no. 7-8, June. 2005, pp. 941-948, doi: 10.1016/j.ijmachtools.2004.10.011.
- [13] J. W. Jung, Y. H. Jeong, B. K. Min and S. J. Lee, "Model-based pulse frequency control for micro-EDM milling using real-time discharge pulse monitoring," *J Manuf Sci Eng Trans ASME*, vol. 130, no. 3, June. 2008, pp. 0311061-03110611, doi: 10.1115/1.2917305.
- [14] B, H, Yan and C, C, Wang, "The machining characteristics of Al₂O₃/6061Al composite using rotary electro-discharge machining with a tube electrode," *J Mater Process Technol*, vol. 95, no. 1-3, October. 1999, pp. 222-231, doi: 10.1016/S0924-0136(99)00322-2.

Fei Wang received his B. S. and M. S. degree in Electromechanics Engineering from China University of Petroleum in 2008 and 2010, respectively. Currently, he is a Ph.D. candidate in College of Electromechanical Engineering in China University of Petroleum, China. His recent research interest is control system of electrical discharge machining.

Yonghong Liu was born in Anhui, P. R. China, in 1965. He received his Ph.D. degree in Mechanical Manufacture from Harbin Institute of Technology, Harbin, China, in 1996.

He is currently a professor and doctoral supervisor in College of Electromechanical Engineering at China University of Petroleum, China. He has published over 120 papers in some international or national journals and conferences. His current research interests include the Expansion Sand Screen for sand control, new kind of numerical control machine about combined machining and control system of the subsea drilling equipments. Dr. Liu is a member of China Nontraditional Machining Committee and Nontraditional Machining Association of Shandong. Dr. Liu is the candidate for the New Century National "Hundred, Thousand and Ten Thousand Talents Project" and prominent young and middle-aged specialist of Shandong Province.

Zhili Chen was born in Shanxi, P. R. China, in 1984. He received his B.S. degree in Machinery Design and Manufacturing and Its Automation from China University of Petroleum in 2009. Currently, he is a postgraduate student in Mechanical Engineering in China University of Petroleum, P. R.

China. His recent research interest is the design of NC machining simulation system

Renjie Ji received his B. S. and M. S. degree in Electromechanics Engineering from China University of Petroleum in 2005 and 2007, respectively. Currently, he is a Ph.D. candidate in College of Electromechanical Engineering in China University of Petroleum, China. His recent research interest is electrical discharge machining of engineering ceramics.

Xiaojie Tian was born in Shandong, P. R. China, in 1985. He received his B.S. and M.S. degree in Mechanical Engineering from China University of Petroleum in 2008 and 2010 respectively. Currently, he is a Ph.D. candidate in Mechanical Design and Theory in China University of Petroleum, China. His recent research interest is the control system of the subsea drilling equipments based on the field bus.

Zengkai Liu was born in Shandong, P. R. China, in 1986. He received his B.S. and M.S. degree in Mechanical Engineering from China University of Petroleum in 2009 and 2011 respectively. Currently, he is a Ph.D. candidate in Mechanical Engineering in China University of Petroleum, P. R. China. His recent research interest is the Field Programmable Gate Array based control system for subsea blowout preventer.

A Case Study of Model Checking Retail Banking System with SPIN

Huiling Shi

Shandong Provincial Key Laboratory of Computer Network
Shandong Computer Science Center, Jinan 250014, P. R. China
Email:shihl@keylab.net

Wenke Ma, Meihong Yang and Xinchang Zhang

Shandong Provincial Key Laboratory of Computer Network
Shandong Computer Science Center, Jinan 250014, P. R. China

Abstract—Model checking is an important technique for ensuring the correctness of investigated system. However, the model checking tools subject to the state-space explosion problem, which is an ignored hurdle to the practical application of the technique. This paper presents a case study of model checking the business flow of retail banking System, through an example of verifying automatic teller machine (ATM) with SPIN. We present the specific approach to effectively abstract the related part of ATM system, and give our experiment results. The verification results show that model checking is feasible technique for verifying the ATM system.

Index Terms—model checking, spin, verification, automatic teller machine, retail banking

I. INTRODUCTION

Retail bank is an important bank business, which offers a range of services to individual customers and small businesses, rather than to large companies and other banks. The services usually include current accounts, savings accounts, investment advice and broking, and loans and mortgages. Retail banks must enable customers to securely and reliably conduct transactions. Traditionally, retail banks have provided these services directly to the customer via branches. At present, retail banks also offer their services by telephone, the internet and automatic teller machine (ATM) as well. In particular, some operate solely via the internet and do not have facilities to serve customers at physical outlets. Additionally, some other organizations, such as supermarkets, have now entered the banking sector and also offer a wide range of banking services. Clearly, it is very necessary to ensure the correctness of services of retail banking system.

Testing is an indispensable step to try to ensure the correctness of a system. However, testing can never completely identify all the defects within an investigated

system. Model checking is a method for formally verifying finite-state concurrent systems. In model checking, properties about the system under verification are usually expressed as temporal logic formulas, and efficient algorithms are used to traverse the system model

To check whether the properties hold or not. Model checking is attractive for the system in which problems of concurrency and distribution make traditional testing challenging. In recent years, there have been many papers [1-10] which report the successful instances of using model checking to provide the validate system verifications.

SPIN(Simple Promela Interpreter) [5] [6] [7] is a generic verification system that supports the design and verification of asynchronous process systems. This model checker accepts design specifications written in the verification language PROMELA (a Process Meta Language) [8] [9] [10], and it accepts correctness claims specified in the syntax of standard Linear Temporal Logic (LTL) [11]. The input language of the model checker SPIN allows us to build high-level models of distributed systems from three basic components: asynchronous processes, message channels, and data objects.

Model checking seems to be a promising approach for ensuring the correctness of retail banking system. However, the model checking tools subject to the state-space explosion problem, which is an ignored hurdle to the practical application of the technique. This paper presents a case study of model checking the business flow of retail banking system, through an example of verifying automatic teller machine (ATM) with SPIN. In the case study, we will present the specific approach to effectively abstract the related part of ATM system, and give our experiment results.

The rest of the paper is organized as follows. We introduce the related work in Section II. Section III introduces the model checking tool SPIN and extended finite state machine (EFSM), which is used to modeling the ATM system. In Section IV, we present the detail approach for modeling the ATM system. The experiment

Manuscript received July 30, 2011; revised September 5, 2011; accepted December 19, 2011.

Project number: 61070039, and BS2011DX033
Contact author: huiling shi.

results are given in Section V. Finally, we summarize the paper in Section VI.

II. RELATED WORK

Model checking is an important method for formally verifying state transition systems because it allows the fully automatic analysis of investigated system. The application of model checking consists of several tasks. The first step is modeling investigated system, i.e., the system to be verified should be converted in a formalism accepted by a model checking tool. In the second step it is necessary to state the properties that the system must satisfy. The specifications are expressed in a logical formalism. It is common to use a temporal logic that can describe how the system evolves over time. The third step is the verification. Often it involves some human assistance for example to perform the analysis of verification results. In case of a negative result, the user is provided with an error trace. This can be assumed as a counterexample for the checked property and means that the system does not verify the property. Hence the system design can be modified and checked again.

So far the model checking has been largely implied, a complete state of the art can be found in [1-13].

In [3], The Dynamic Host Configuration Protocol (DHCP) is studied according to the concept of modeling and verification. In [11], the paper proposed a formal method for the verification of ebXML based e-commerce system. And the approach allowed to highlight some weakness of the protocol mainly due to the lack of a clear and complete set of specifications. In [12], the paper shows how finite model-checking based approach can be applied to analyze properties of ad-hoc sensor networks. It proves that SPIN and finite model checking are appropriate for studying properties of ad-hoc sensor networks specifications. In [13], the paper gives an approach to verify the object model of rCOS using model checker Spin. And a case study presents to show how the approach works.

The troublesome problem of using model checking technique is the state-space explosion. There are several approaches to combat this problem, which can be classified into two categories, i.e. simplifying the system model by higher abstraction (e.g. [9]), and reducing the resource consumption in the process of model checking (e.g. [14]).

III. PRELIMINARIES

A. SPIN

Model checking is used, for example, for the verification of protocols and hardware circuits [1] [2]. Many tools, called model checkers, have been developed to this aim. The most famous one is SPIN [7]. Trying all possible interleaving to see which ones can lead to failure would be astoundingly complex. To avoid this SPIN uses a theory of partial order reduction [14] to group process executions into equivalence classes.

To check the compliance of a system with a logic system property specified in linear temporal logic, SPIN

first converts the formula into a test automaton that works much like an observer or monitor of the system executions. While building the system executions, the monitor is consulted at every step to see if violations occurred. If a violation is detected, SPIN displays the exact interleaving sequence leading from the initial system state to the state where the violation was detected. This serves as a counter-example to the correctness claims and facilitates diagnosis of the detected violation.

SPIN accepts design specifications written in the verification language PROMELA, and it accepts correctness claims specified in Linear Temporal Logic (LTL).

PROMELA is a language for building verification models that represent an abstract of a system, which contains only those aspects that are relevant to the properties one wants to verify [9]. A PROMELA program consists of processes, message channels, and variables. Processes are defined globally; while message channels and variables can be declared either globally or locally within a process. Processes are used to specify system behaviors, and channels and global variables are used to define the environment in which the processes run. Examples and further details about the PROMELA language can be found in references [8][9].

LTL is a modal logic aimed at encoding how states evolve over time. It has been proven to have good expressivity and more natural language like statements for verification. LTL has three unary modal operators (X, F , and G) and three binary modal operators (U, R, W) [11]. A formula $X\varphi$ is true in particular state if and only if the formula φ is true in the next state; $G\varphi$ is true if and only φ is true from now on; $F\varphi$ is true if φ is or will be true at some time in the future; $\varphi U \psi$ is true if ψ will eventually become true and φ stays true until then.

B. EFSM Model

In a finite state machine (FSM), the transition is associated with a set of input boolean conditions and a set of output boolean functions. Different from FSM, the transition of EFSM can be expressed by an “if statement” consisting of a set of trigger conditions. If trigger conditions are all satisfied, the transition is fired, bringing the machine from the current state to the next state and performing the specified data operations.

Formally, EFSM M is defined as the tuple $(S, s_0, V, M_V, P, M_P, I, O, T)$, in which:

- S is a finite set of states, s_0 is the initial state, $s_0 \in S$;
- V is the finite set of the internal variable (environment variable), and the range of the internal variable is D_V ;
- M_V is the set of the initial (or default) value of variables in V , in which any element can be expressed as a tuple (s, v) , $s \in S$, $v \in D_V$;
- P is the input and output parameters;
- M_P is the set of the initial (or default) value of variables in P , in which any element can be expressed as a tuple (p, u) , $p \in I \cup O$, $u \in D_p$, D_p is the range of the input and output parameters;

- I is a set of the input symbols;
- O is a set of the output symbols;
- T is a finite set of state transition.

A state transition $t(t \in T)$ is defined as the tuple $(s, x, y, g_P, g_E, op, e)$, where:

- s and e are the start (head) state and the end (tail) state;
- g_P is the input and output conditions to determine;
- g_E is the conditions to determine of the variable required for migration;
- x and y are the input and output symbols;
- op is output operations.

IV. MODELING THE BUSINESS FLOW OF ATM

ATMs are the most immediately visible type of retail banking technology. The main operations of ATMs include balance and transaction enquiries, withdrawals, deposits and accounts transfer. In this section, we will present the model (including EFSM and Promela models) of the main business flow of the ATMs. The model has been simplified in order to obtain a minor number of states to manage in the formal verification. Particularly, encryption has been omitted.

A. EFSM Model

ATM can be used to login with a card and a pin, perform transactions against the account (deposit, withdraw, inquire balance), and logoff after desired transactions. A user gets 3 chances to login with a valid pin, after which the card is locked until reset by an official. There are usual restrictions on amounts that can be withdrawn and the number of withdrawals (6) that can be made.

ATM in the actual course of three parties are involved: the cardholder, the terminal, the bank sever. Cardholder interacts with the banking system through a terminal, for withdrawals, transfers, inquiries. Terminal receives a request from the cardholder to handle all the business logic from the terminal, and forwards the request to bank server, while waiting for the bank server's response, and forwards the response to the terminal. Bank server receives a request from the terminal to make the approval or rejection of the response and make the appropriate accounting treatment.

In order to use model checker SPIN, we have to describe the specifications of ATM in PROMELA language. But before describing the specifications in PROMELA, we can model the specification in Extended Finite State Machine (EFSM) that can be readily expressed in PROMELA. Fig. 1 shows specifications for ATM expressed in EFSM. Variable CA is said that the number of unsuccessful logon attempts, if CA is greater than or equal 3, then the card is locked. Variable CW is said that the number of successful withdrawals, if CW is greater than 6, then emerges an error.

Fig. 1 shows the EFSM model of Automated Teller Machine, containing 4 states and 13 state transitions. The label of transition "PIN/rOk[CA<=2,CW<6],Logon" means that, when in the implementation of "Logon" operation and the input symbol of the state PIN satisfying " CA<=2" and "CW<6", the state will be converted from the state "Init" to the state "Withdraw/deposit/balinq", and with the outputting of "rOk". Similarly, The label of transition "PIN/Invalid[CA=2],Logon" means that, when in the implementation of "Logon" operation and the input symbol of the state PIN satisfying " CA=2", the state will be converted from the state "Re_Logon" to the state "CardLocked" with outputting of "Invalid"

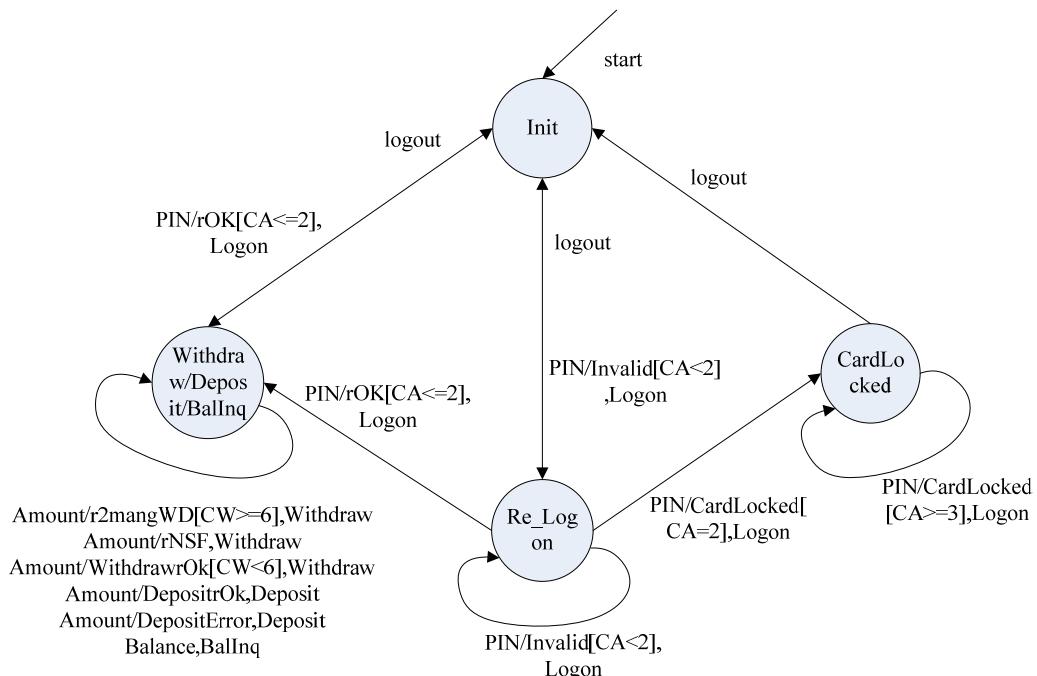


Figure 1. EFSM of automated teller machine

B. PROMELA Model

PROMELA program consists of asynchronous processes, message channels, and data objects. Process Characterizes the behavior of systems, channels and global variables used to define the process execution environment. PROMELA is similar to the C programming language that allows dynamic creation of parallel processes, and processes can be synchronized via a message channel (using the meeting point (rendezvousPort) and asynchronous (buffered) communication).

Though EFSM model is expressed, we need to decide how to model the exchange of messages among three parties. The message flow is shown in Fig. 2. Message is defined as “mtype={PIN,InvalidPIN,CardLocked,OkPIN,Operate_Logon,Operate_Withdraw,Operate_Deposit,Operate_BalInq,Operate_Logout,WithdrawrOk,DepositOk,rNSF,r2ManyWD,DepositError,WithdrawError,SmbolRequest,SmbolResponse }”.

In order to avoid the state explosion, the type of amount is “Byte”, and “Byte” type can also simulate the amount of transfer between the three parties and changes.

Message channels are used to model the exchange of data between processes. Channels are declared as shown in Fig. 3. “MAX” represents the number of Terminals. The number of Cardholders is equal to terminals. The same to say, one bank servers MAX terminals. “CusToATM[Max]” is an array of type channel. “CusToATM” sends messages from cardholder to terminal, and each message is said to consist of four fields: the first is declared to be of type byte that represents the cardholder No, the second is of type mtype that represents operation from cardholder to terminal, the third is of type byte that represents amount relates and the last is of type mtype that represents messages related to operations. Similarly, “ATMToCus” sends messages from terminal to cardholder. “ATMToBank” is declared to be capable of storing up to MAX messages from terminal to bank, and each message is said to consist of four fields: the first is declared to be of type byte, the second is of type mtype, the third is of type byte and the last is of type mtype. “BankToATM” sends messages from bank to terminal similarly to “ATMToBank”.

According to the analysis above, we create models in PROMELA as shown in Fig. 4, Fig.5 and Fig.6. We instantiate processes as shown in Fig. 7 by using a predefined operator called “run”. In addition to, some global variables are defined, such as “AccountBalance” that indicates the current account balance and changes with successful deposit or withdrawal operations.

Because the model checker runs in a closed condition, the user can not participate in the process of running. Therefore, in order to traverse all cases, the initial account balance, and the amount of each deposit and withdrawal will be careful to design. For example, if the initial account balance is “20” and amount of withdrawal is 40, then the model checker may only traverse a path with outputting “rNSF”, while the other states are not reachable. In this case, we can not decide that the model does not meet properties. In addition to, For example,

given the initial value of “ $CW=0$ ”, to covers the transition “ $Amount/r2manyWD [CW>=6], withdraw$ ” in the EFSM of Fig. 1, we need to invoke the transition labeled “ $Amount/rok[CW<6], withdraw$ ” 5 times. As an example of an infeasible transition, given the initial balance of 100 and the amount withdrawal of 5, the transition labeled “ $Amount/rNSF, Withdraw$ ” is not feasible. In order to model more realistic simulation of the situation, the amount of deposit and withdrawal required some changes. In Fig. 4, the amount of deposit is determined concurrently between the fixed values by using “..”.

In Fig. 7, we instantiate processes with the order from Bank Server, ATM to Cardholder

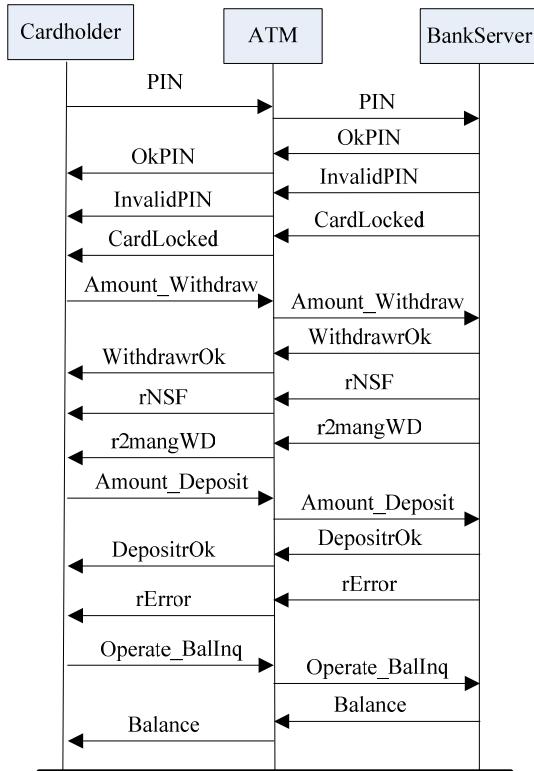


Figure 2. Message flow of ATM

```

Chan CusToATM[MAX]=[0]of{byte, mtype, byte, mtype}; /*message from customer to ATM*/
Chan ATMToCus[MAX]=[0]of{byte, mtype, byte, mtype}; /* message from ATM to customer*/
Chan ATMToBank=[MAX]of{byte, mtype, byte, mtype}; /*message from ATM to bank */
Chan BankToATM=[MAX]of{byte, mtype, byte, mtype}; /*message from bank to ATM */
  
```

Figure 3. .Channels' definition in PROMELA

```

proctype Cardholder(chan cusToATM;chan atmToCus;byte j)
{ ...
Cus_Logon_Request:
if
::cusToATM!j,Operate_Logon,0,PIN->
if
::atmToCus?eval(j),Operate_Logon,0,OkPIN->
printf(" You are welcome!");
goto SelectOperator
::atmToCus?eval(j),Operate_Logon,0,InvalidPIN->
printf(" PIN is Invalid!");
goto Cus_Logon_Request
::atmToCus?eval(j),Operate_Logon,0,CardLocked ->
printf(" Card is Locked! ");
goto Cus_CardIsLocked
fi;
Cus_CardIsLocked:
if
::CA=0 ->
goto Cus_Logon_Request/*simulate next day*/
::cusToATM!j,Operate_Logout,0,SmbolRequest->
/* logout*/
...
fi;
SelectOperator:
if
:: printf("CUS_SelectOperator:Deposit \n")->
if
::Amount_Deposit=20
::Amount_Deposit=30
fi;
if
::cusToATM!j,Operate_Deposit,Amount_Deposit
,SmbolRequest->
if
::atmToCus?eval(j),Operate_Deposit,0,DepositError->
printf(" Error, deposit is not ok \n");
goto SelectOperator
...
fi;
:: printf("CUS_SelectOperator:WithDraw \n")->
if
::Amount_Withdraw=20
::Amount_Withdraw=30
fi;
if
::cusToATM!j,Operate_Withdraw,Amount Withdraw
,SmbolRequest ->
if
::atmToCus?eval(j),Operate_Withdraw,0
,WithdrawOk->
printf("withDraw is ok \n");
goto SelectOperator
...
fi;
fi;
...
}

```

Figure 4. Fragment of cardholder process in PROMELA

```

proctype ATM(chan cusToATM;chan atmToCus;
chan atmToBank;chan bankToATM;byte i)
{ ...
ATM_Begin:
if
::cusToATM?eval(i),Operate_Logon,0,PIN->
atmToBank!i,Operate_Logon,0,PIN;
if
:: bankToATM?eval(i),Operate_Logon,0,OkPIN ->
atmToCus!i,Operate_Logon,0,OkPIN;
goto ATM_Begin
:: bankToATM?eval(i),Operate_Logon,0,InvalidPIN->
atmToCus!i,Operate_Logon,0,InvalidPIN;
goto ATM_Begin
:: bankToATM?eval(i),Operate_Logon,0,CardLocked->
atmToCus!i,Operate_Logon,0,CardLocked;
goto ATM_Begin
fi
:: cusToATM?eval(i),Operate_Deposit
,ATM_Amount_Deposit, SmbolRequest->
::atmToBank!i,Operate_Deposit,ATM_Amount
_Deposit,SmbolRequest->
if
::bankToATM?eval(i),Operate_Deposit,0,DepositOk->
atmToCus!i,Operate_Deposit,0,DepositOk;
goto ATM_Begin
::bankToATM?eval(i),Operate_Deposit,0,DepositError->
atmToCus!i,Operate_Deposit,0,DepositError;
goto ATM_Begin
fi
:: cusToATM?eval(i),Operate_Withdraw,
ATM_Amount_Withdraw,SmbolRequest->
::atmToBank!i,Operate_Withdraw,
ATM_Amount_Withdraw,SmbolRequest->
if
::bankToATM?eval(i),Operate_Withdraw,0
,WithdrawOk->
atmToCus!i,Operate_Withdraw,0,WithdrawOk;
goto ATM_Begin
::bankToATM?eval(i),Operate_Withdraw,0,rNsf->
atmToCus!i,Operate_Withdraw,0,rNsf;
goto ATM_Begin
...
}

```

Figure 5. Fragment of ATM process in PROMELA

```

proctype BankServer(chan atmToBank;chan bankToATM )
{ ...
Bank_Begin:
if
::atmToBank?k,Operate_Logon,0,PIN->
if
::(CA==0)->
if
::bankToATM!k,Operate_Logon,0,OkPIN->
goto Bank_Begin
::bankToATM!k,Operate_Logon,0,InvalidPIN->
CA=CA+1; goto Bank_Begin
fi
::(CA==1)->
if
::bankToATM!k,Operate_Logon,0,OkPIN->
goto Bank_Begin
::bankToATM!k,Operate_Logon,0,InvalidPIN->
CA=CA+1; goto Bank_Begin
fi
::(CA==2)->
if
::bankToATM!k,Operate_Logon,0,OkPIN->
goto Bank_Begin
::bankToATM!k,Operate_Logon,0,CardLocked->
CA=CA+1; goto Bank_Begin
fi
fi
:: atmToBank?p,Operate_Withdraw,Bank_Amount_
Withdraw, SmbolRequest->
if
::(Bank_Amount_Withdraw>Balance)->
bankToATM!p,Operate_Withdraw,0,rNsf;
goto Bank_Begin
::(Bank_Amount_Withdraw<=Balance)->
if
::(CW<6)->
if
:: bankToATM!p,Operate_Withdraw,0,WithdrawrOk->
Balance=Balance-Bank_Amount_Withdraw;
CW=CW+1; goto Bank_Begin
::bankToATM!p,Operate_Withdraw,0,WithdrawError->
goto Bank_Begin
fi
::(CW==6)->
bankToATM!p,Operate_Withdraw,0,r2ManyWD;
goto Bank_Begin
fi
fi
...
}

```

Figure 6. Fragment of bankserver process in PROMELA

```

init {
atomic{
run BankServer(ATMToBank,BankToATM);
run ATM(CusToATM[0],ATMToCus[0],ATMToBank,BankToATM,0);
run ATM(CusToATM[1],ATMToCus[1],ATMToBank,BankToATM,1);
run ATM(CusToATM[2],ATMToCus[2],ATMToBank,BankToATM,2);
run Customer(CusToATM[0],ATMToCus[0],0);
run Customer(CusToATM[1],ATMToCus[1],1);
run Customer(CusToATM[2],ATMToCus[2],2);
...
}
}

```

Figure 7. Fragment of init process in PROMELA

V. PROPERTIES DEFINITION AND VERIFICATION

A. Properties Definition

SPIN assists users in finding unreachable codes or deadlocks. In addition, SPIN also verifies LTL properties that we are interested in against PROMELA models. LTL allows expressing temporal properties we expect the system behavior will conform to during the system lifetime.

About ATMs, the first property ensures that deposits and withdrawals are mutually exclusive operations forever. It can be expressed in LTL formula as “!($\langle\rangle$ (withdrawning &&depositing))”. And PROMELA Model also needs insert few control statements, as in shown in Fig. 8.

```
proctype BankServer(chan atmToBank;chan bankToATM )
{
...
if
  :bankToATM!p,Operate_Withdraw,0,rNsf->
    goto Bank_Begin
  ::(withdrawning==false && depositing==false)->
    atomic{
      withdrawal=true;
      bankToATM!p,Operate_Withdraw,0,WithdrawOk;
      withdrawal=false;
      goto Bank_Begin
    }
...
fi
...
}
```

Figure 8. Fragment of process with control statements

The second property to be verified should ensure that card holders should be allowed access to deposit, withdraw and query after entering the correct password. It can be expressed in LTL formula as: $[\![pinIsOk->\langle\rangle(depositPermit /\!\!/ withdrawPermit /\!\!/ balingPermit /\!\!/ logoutPermit)]\!]$

The third property to be verified should ensure that As long as the withdrawal amount is not greater than the account balance, finally withdrawal will be successful. It can be expressed in LTL formula as:

$[\![amountNoMoreThanBalance->\langle\rangle withdrawIsOk]\!]$

B. Verification

We stated some properties that an ATM should verify. Let us now observe whether they are or not verified in the analyzed PROMELA model. We performed the experiments with SPIN version 6.1.0 on a cloud computing platform. The allocated resources from the platform are as follows: Intel xeon E7540 2GHz; 8GB of RAM; Linux version 2.6.18.

We verified successfully that PROMELA models meet all the properties with 3 or less cardholders, and No deadlock or non-progress cycle were found. With 3 cardholders, PROMELA models are sufficient to simulate the business of ATM, and do not need to verify with more than 3 cardholders. The Figs below summarize some of the performance measures of the verification. Fig. 9 shows number of states with deferent cardholders. Fig.

10 shows number of transitions with deferent cardholders. Fig.11 shows total memory usages with deferent cardholders. Fig.12 shows elapsed times with deferent cardholders. Fig.13 shows DFS-search with deferent cardholders. All Figs match index movement tendency.

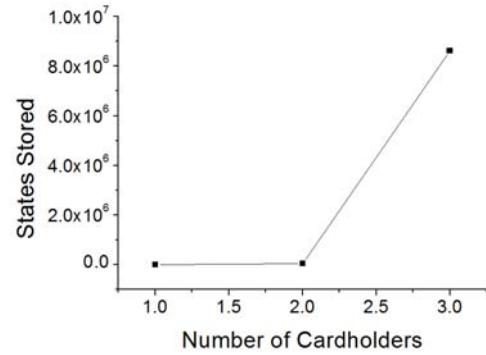


Figure 9. States stored of 1-3 cardholders

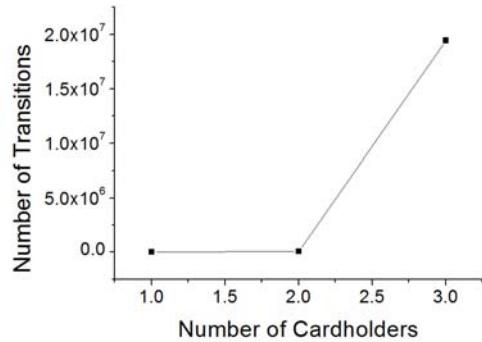


Figure 10. Transitions of 1-3 cardholders

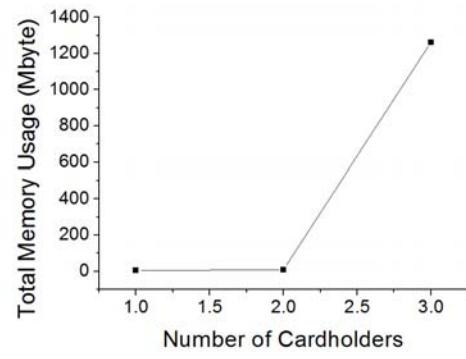


Figure 11. Total memory usage of 1-3 cardholders

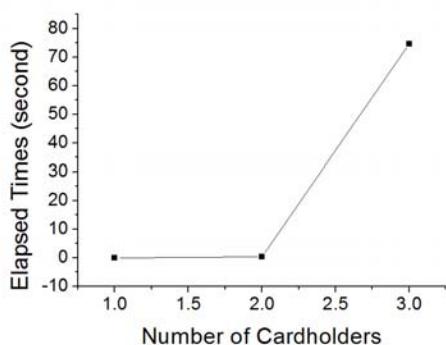


Figure 12. Elapsed times of 1-3 cardholders

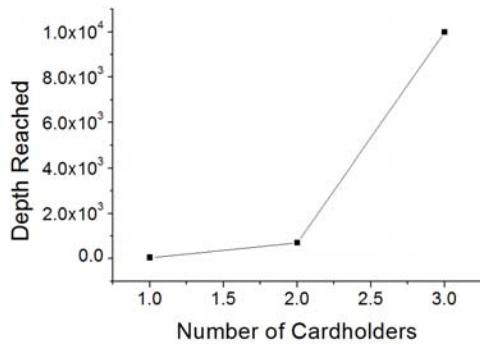


Figure 13. Depth reached of 1-3 cardholders

VI. CONCLUSIONS AND FUTURE WORK

In this paper we give an approach to verify the business flow of retail banking system, through an example of verifying automatic teller machine (ATM) with SPIN. We consider ATM as extended finite state machines that can be presented in PROMELA. We show how properties of ATM can be expressed in the form of linear temporal logic statements and then verified by applying model checker SPIN. It proves that SPIN and model checking are appropriate for studying the business flow of ATMs. And in our future research, we will expand the field used by model checking with SPIN, for example e-banking system, mobile banking system. Also we will study approaches to solve the state explosion.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant No. 61070039, the Fund for the Doctoral Research of Shandong Academy of Sciences under Grant No. 2010-12, the Outstanding Young and Middle-aged Scholars Foundation of Shandong Province of China under Contract No. BS2011DX033.

REFERENCES.

- [1] Lu Simei, Zhang Jianlin, Luo Liming. "The Automatic Verification and Improvement of SET Protocol Model with SMV". In: *Proceedings of Information Engineering and Electronic Commerce*, 2009.
- [2] Stylianos Basagiannis, Panagiotis Katsaros, Andrew Pombortsis. "An intruder model with message inspection for model checking security protocols". *Computers & Security*, 29, pp:16-34, 2010.
- [3] Syed M.S. Islam, Mohammed H. Sqalli, Sohel Khan. "Modeling and Formal Verification of DHCP Using SPIN". *International Journal of Computer Science & Applications*. 6(3), pp:145-159, 2006.
- [4] Andreas Both, Wolf Zimmermann and Rene Franke. "Model Checking of Component Protocol Conformance - Optimizations by Reducing False Negatives". *Electronic Notes in Theoretical Computer Science*, 263, pp:67-94, 2010.
- [5] Barry Long, Juergen Dingel, T.C. Nicholas Graham. "Experience Applying the SPIN Model Checker to an Industrial Telecommunications System". In: *proc. of 30th International Conference on Software Engineering*, ACM 2008:693-702.
- [6] Li Jing, Li Jinhua. "Model Checking the SET Purchasing Process Protocol with SPIN". In: *proc. Of 5th International Conference on Wireless Communications, Networking and Mobile Computing*, 2009.
- [7] On-the-fly LTL model checking with SPIN. <http://spinroot.com/spin/whatispin.html>
- [8] G.J. Holzmann, "The model checker SPIN", *IEEE Transactions on SE*, vol.23 (5), pp. 279-295, 1997
- [9] G. J. Holzmann, "The SPIN Model Checker: Primer and Reference Manual", Addison Wesley, 2004.
- [10] Michael Huth, Mark Ryan, "Logic in Computer Science: Modelling and Reasoning about System", Cambridge University Press, 1999.
- [11] Marina Mongiello, "Finite-state verification of the ebXML protocol", *Electronic Commerce Research and Applications* pp: 147-169, 2006(5).
- [12] Vladimir A. Oleshchuk, "Modeling, specification and verification of ad-hoc sensor networks using SPIN", *Computer Standards & Interfaces*, pp: 159-165, 2005(28).
- [13] Xiao Yu, Zheng Wang, Geguang Pu, etc. "The Verification of rCOS Using Spin". *Electronic Notes in Theoretical Computer Science*, pp: 49-67, 2008(207).
- [14] Flanagan C and Godefroid P. "Dynamic partial-order reduction for model checking software". *ACM SIGPLAN Notices*, 40(1), pp: 110-121, 2005.

Huiling Shi(1978-),was born in Heze, China. She received her B.E degree in 2001 from Computer and Applications, Shandong University of Technology in Zibo, her M.E. degree in 2004 from Computer and Applications, Shandong University in Jinan. She works in Shandong Computer Science Center. Her main research field is in software testing, formal verification software quality assurance, and so on.

The Investigation of Fault Diagnosis Based on GA-HPSO-NN

Bin Peng

Key Laboratory of Digital Manufacturing Technology and Application, The Ministry of Education,
Lanzhou University of Technology, Lanzhou, 730050, China
School of Mechanical and Electronical Engineering, Lanzhou University of Technology
Email: pengb2000@163.com

Chang-an Hu, Rong-zhen Zhao and Yu-qiao Zheng

Key Laboratory of Digital Manufacturing Technology and Application, The Ministry of Education,
Lanzhou University of Technology, Lanzhou, 730050, China
School of Mechanical and Electronical Engineering, Lanzhou University of Technology
Email: ajitajitai@163.com, haorongzhen@lut.cn, yuqiaozheng@lut.cn

Abstract—At present, although most fault diagnosis methods of rotating machinery is qualitatively used, it is gravely lacking in quantitative accuracy. So a novel algorithm GA-HPSO combining with the advantages of genetic algorithm (GA), simulated annealing (SA) and particle swarm optimization (PSO) was provided to train neural network (NN). The proportion and blend methods were applied to the novel algorithm. Information entropy was used to take fault signals. Four kinds of spectral entropies and six kinds of typical rotor faults were used as input and output data. NN classifier based on GA-HPSO was set up. The simulation results indicate that GA-HPSO has a better ability to escape from a local minimum and is more effective than the conventional single algorithm. It can rapidly and accurately realize fault data classification. It provides a new method for fault diagnosis.

Index Terms—Information Entropy; GA-HPSO-NN; Fault Diagnosis

I. INTRODUCTION

With the development of modern industry, the large rotating machines are developing toward high speed and high efficiency increasingly. The relationship of fault and sign is unclear which hampers operator's ability to diagnose and eliminate equipment failures before faults happen [1]. With the prevalence of computer technology, the intelligence monitoring system is more important. Fault diagnosis is the main embodiment of intellectual monitoring system, so the level of diagnose technology influences the function of intellectual monitoring system. It also has important meaning to keep normal running, reduce product cost, raise product efficiency and ensure product safety for whole system.

The coupling weights that distributed in NN are used

to express the diagnosis knowledge. It can realize the complex non-linear mapping relation of fault and sign through associative memory, pattern matching and similar induction. It is widely used in fault diagnosis of rotor system, especially for pattern recognition of multiple faults and many signs.

Back propagation algorithm (BP) based on gradient learning is the most common training method for NN, but it is easy to fall in the local best solution. GA, PSO and SA are some of the well-known meta-heuristic algorithms. GA shows unique advantages in establishing system structure and global optimization. PSO is an evolutionary computational method, which may be conveniently employed to execute random and global search. SA is a generic probabilistic metaheuristic for the global optimization problem. It accepts the current optimal solution at a probability after searching, so it can overcome local minimum point. There are some scholars and many experts who used different algorithms to train NN and solve different problems. Ya-xiang Xu et al. adopted adaptive PSO to optimize NN [2]. K. Premalatha et al. proved that GA-PSO had advantages over standard PSO [3]. Chang-cai Cui et al. put forward a novel heuristic method GA-PSO to optimize engineering problem [4]. Wen-yi Wang et al. proposed a effective optimization method GA-PSO [5]. Jun Liu et al. adopted improved PSO to train NN [6].

In order to avoid the shortcoming of the standard single algorithm, a novel algorithm GA-HPSO-NN combining with the advantages of GA, PSO, SA and NN was put forward based on previous research. Fault vibration signals were evaluated from the time, frequency and time-frequency domains in extraction sample. Multi-angle characteristics of rotor fault signals are the input variables of NN. The output values are six kinds of typical fault. GA was used for coarse search of weights and threshold under optimized NN structure, then HPSO with cross factor was used for exact search. In order to improve the ability of the new algorithm to escape from a

Corresponding author: Bin Peng, School of Mechanical and Electronical Engineering, Lanzhou University of Technology, 730050, China, Email: pengb2000@163.com

local optimum, SA was used to modify GA. The final weights and bias of NN were gotten through GA-HPSO. Computer simulation results are provided to compare the performance of GA, PSO and GA-HPSO. The results show that the fault diagnosis accuracy is effectively improved using GA-HPSO [7].

II. GA-HPSO-NN

A. Fault Feature Extraction

There is much relevant information with actual condition in the original signal detection. The extracted characteristic should be able to reflect the regularity and sensitivity of actual condition, contain regularity fault characteristic and have better fault separability. So the fault feature extraction and processing is the key to realize NN classification. The information entropy expresses statistical characterization of the whole signal which is used to measure the overall uncertainty of information source. The value is smaller, then the information is more certain and unordered degree is smaller. Singular value spectral entropy of time domain, power spectral entropy of frequency domain, wavelet energy and space spectrum entropy of time-frequency domain are used to measure rotor fault signal index [8].

The information entropy can be defined as follows:

$X = \{x_1, x_2, \dots, x_n\}$ expresses a whole set of rotor vibration signal. Probability x_i of every component is $P = (x_i)$ and $\sum_{i=1}^n P(x_i) = 1$. The information entropy $H(X)$ of X is:

$$H(X) = -\sum_{i=1}^n P(x_i) \log P(x_i) = -\sum_{i=1}^n P_i \log P_i \quad (1)$$

Four kinds of spectral entropy can be gotten through formula (1).

B. GA

GA is stochastic search technique. It first proposed by Holland, is inspired by the mechanism of natural selection and natural genetics. GA represents a highly parallel adaptive search process. GA has received considerable attention regarding their potential as a class of stochastic searching algorithms for complex problems and has been successfully applied in the area of industrial engineering. GA can avoid the problems inherent in more traditional approaches. Restrictions on the range of the parameter-space are imposed only by observations and by the physics of the model. Although the parameter-space so-defined is often quite large, the GA provides a relatively efficient means of searching globally for the best-fit model. While it is difficult for GA to find precise values for the set of best-fit parameters, they are well suited to search for the region of parameter-space that contains the global minimum. In this sense, the GA is an objective means of obtaining a good first guess for a

more traditional method which can narrow in on the precise values and uncertainties of the best-fit.

The following list shows the general procedure of GA as described by Mitsuo Gen and RunWei Cheng.

Procedure of GA

```

start
    initialize X(t);
    t=0;
    while (not termination condition)
        do
            Evaluate fitness of X(t) of each individual;
            Selection operation to X(t);
            Crossover operation to X(t);
            Mutation operation X(t);
            X(t+1)=X(t);
        end while
    end

```

GA starts from a population of randomly generated individuals and happens in generations called population. Each individual in the population is called a chromosome, representing a solution to the problem at hand. A chromosome is a string of symbols; it is usually, but not necessarily, a binary bit string. The chromosomes evolve through successive iterations, called generations. In each generation, the fitness of every individual in the population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), and modified (recombined and possibly randomly mutated) to form a new population. Fitter chromosomes have higher probabilities of being selected. The new population is then used in the next iteration of the algorithm. Some of the parents are rejected and an equal number of offsprings are accepted as the replacement of these parents so as to keep the population size constant. After several generations the algorithms converge to the best chromosome, which hopefully represents the optimum, or at least suboptimal, solution to the problem [9-10].

C. SA

SA is a well-known algorithm used to solve discrete optimization problems. SA is based on the annealing of metals. If a metal is cooled slowly, it forms into a smooth piece because its molecules have entered a crystal structure. This crystal structure represents the minimum energy state, or the optimal solution, for an optimization problem. If a metal is cooled too fast, the metal will form a jagged piece that is rough and covered with bumps. These bumps and jagged edges represent the local minimums and maximums. Kirkpatrick originally thought of using SA on computer related problems. He did this in 1983 and applied SA to various optimization problems. From there, many other people have worked on it and have applied it to many optimization problems. The algorithm borrows the annealing analogy from Statistical Mechanics. In the search process, the SA accepts not only better but also worse neighboring solutions with a certain probability. Such mechanism can be regarded as a trial to

explore new space for new solutions, either better or worse. The probability of accepting a worse solution is larger at higher initial temperature. As the temperature decreases, the probability of accepting worse solutions gradually approaches zero. This feature means that the SA technique makes it possible to jump out of a local optimum to search for the global optimum. So SA is a good algorithm because it is relatively general and tends to not get stuck in local minimum [11-12].

D. PSO

PSO is an evolutionary computation technique and simulates the behavior of birds flocking developed by Dr. Eberhart and Dr. Kennedy in 1995. Individuals in the community have the ability to control their own behavior based on certain internal and external information. It means that each individual has certain sensory ability and can perceive the local best and global best position of the individual. The particle takes its next action according to the current condition and obtained information. So the whole community displays a certain intelligence. When solving an optimization question, each individual position is correspondingly regarded as a latent solution. According to the above rules, the global optimal solution can be obtained through repeatedly adjusting these latent solutions. For the nth iteration, the particles of PSO change according to the following two formulas:

$$x_{id}^{n+1} = x_{id}^n + v_{id}^{n+1} \quad (2)$$

$$v_{id}^{n+1} = \begin{cases} w \cdot v_{id}^n + c_1 \cdot \text{rand}() \cdot (p_{id} - x_{id}^n) + \\ c_2 \cdot \text{rand}() \cdot (p_{gd} - x_{id}^n) \end{cases} \quad (3)$$

$$i=1,2,\dots,M$$

Where M is the particle sum; v_{id}^n is the d th weight of flight velocity vector for the n th iteration for particle i ; x_{id}^n is the d th weight of position vector for the n th iteration for particle i ; p_{id} is the d th weight of $Pbest$ for particle i ; p_{gd} is the d th weight of $Gbest$ for particle i . $Pbest$ is the best of a particle; $Gbest$ is the best of all particles; c_1 and c_2 are learning factors; $\text{rand}()$ is a random number between 0-1; w is an inertia weight function. The new velocity of particle i is computed using the formula (2) through three parts:

(1) The first is the previous time velocity of particle i . It shows the present condition and can balance global and local search ability.

(2) The second is the recognition part. It indicates the particle thought, enables the particle have the strong global search ability and avoid a local minimum,

(3) The third is the social part. It realizes information sharing between particles.

Under the influence of the three parts, the particles adjust their positions based on experience and information sharing. Finally the globe best solution can be obtained [13-14].

E. NN

NN is an information processing paradigm that is inspired by the way biological nervous systems. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (nodes) working in unison to solve specific problems. NN is configured for a specific application, such as data classification or pattern recognition, through a learning process. Learning in biological systems involves adjustments to the connection strengths that exist between the nodes.

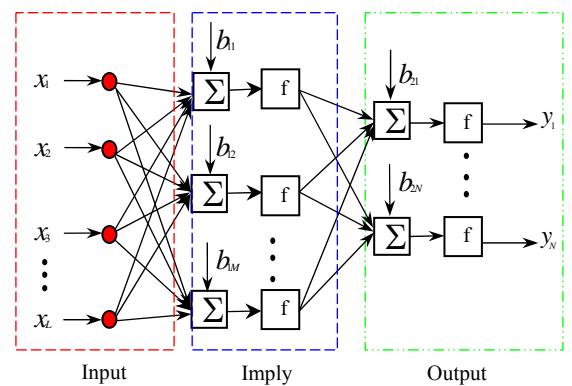


Fig.1 Standard BPNN

Fig.1 is the standard BPNN. It is made up of input layer, imply layer and output layer. If there are L inputs and N outputs, there is a nonlinearity mapping from input to output. Much research displayed that enough imply layer nodes can approach any continuous function. Relational coefficient was applied to simplify the number of the hidden layer nodes. It can be realized as follows:

O_{pi} is the output of imply layer node i under studying sample p . O_{pj} is the output of imply layer node j under studying sample p . N is learning samples total.

$$\text{So } \overline{O}_i = \frac{1}{N} \sum_{p=1}^N O_{pi} \quad (4)$$

$$\overline{O}_j = \frac{1}{N} \sum_{p=1}^N O_{pj} \quad (5)$$

$$X_p = \overline{O}_{pi} - \frac{1}{N} \sum_{p=1}^N O_{pi} = O_{pi} - \overline{O}_i \quad (6)$$

$$Y_p = \overline{O}_{pj} - \frac{1}{N} \sum_{p=1}^N O_{pj} = O_{pj} - \overline{O}_j \quad (7)$$

So relational coefficient of O_{pi} and O_{pj} is:

$$\rho_{ij} = \frac{\sum_{p=1}^N X_p \times Y_p}{\sqrt{\sum_{p=1}^N X_p^2} \times \sqrt{\sum_{p=1}^N Y_p^2}} \quad (8)$$

Where $\rho_{ij} \leq 1$, it reflects function repetition rate of node i and j. If $\rho_{ij} \geq 0.9$, node i and j are incorporated.

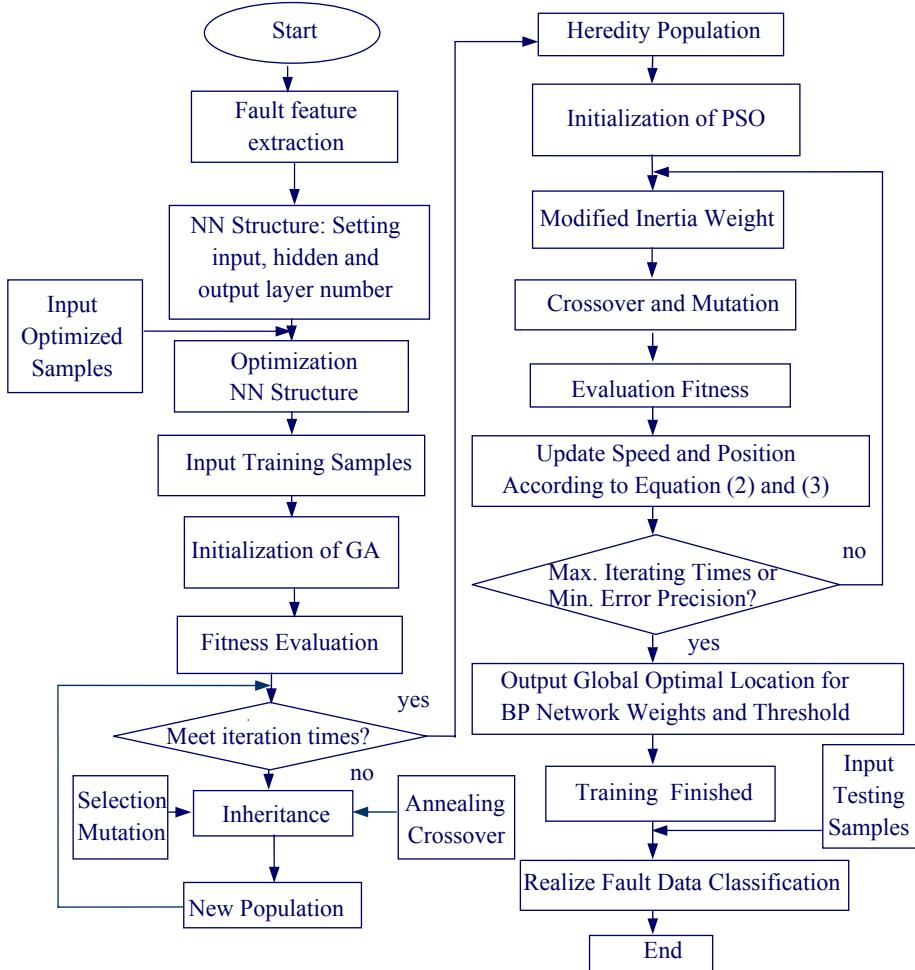


Fig.2 Rotor fault diagnosis flow chart based on GA-HPSO-NN

F. GA-HPSO-NN

There are many approaches to train NN using GA and PSO combination. The most common way is order and embedded combination. For order combination way, GA is used only for position vector of the particle. Particle velocity vector is difficult to track particle, so the effect is not good. While embedded method was too simple, the final effect is also not very ideal. So a novel GA-HPSO-NN combining GA, SA, PSO and NN, was proposed. After running M generations GA, then N generations HPSO was run. M and N value was selected according to the certain proportion. At the same time SA, selection, mutation and crossover were used to GA, and mutation and crossover were used to PSO. Fig.2 is the proposed rotor fault diagnosis flow chart based on GA-HPSO-NN.

The weights and biases optimization of NN was divided into two stages: GA-SA-NN and HPSO-NN.

The article is a small sample problem. The sample size was selected 60 according to convergence time and precision. Every individual of population is chromosome including weight and threshold of NN.

Dimension of population can be defined as follows:

$$D = r \times S_1 + S_1 \times S_2 + S_1 + S_2 \quad (9)$$

Where r is input layer number of NN. S_1 is hidden node number of NN. S_2 is output layer number of NN.

In order to simplify coding and avoid inhomogeneity of initialized weights and threshold, real number coding and grid distribution method was applied [15-16].

GA-SA-NN**Step 1 Start.****Step 2 Fault Feature Extraction.****Step 3 NN Structure.**

Four kinds of spectral entropy of fault diagnosis are the input variable of NN. The output values are six kinds of typical faults of rotor: rotor misalignment, mass unbalanced, contact rubbing, loose, oil film whirl and oil whip. In order to speed up the training speed, improve anti-noise ability, and avoid NN into local optimal, noise samples were added to the original data. In order to avoid excessive input value produce saturation state, the input data was normalized from 0.4 to 0.9, and interval is 0.1.

Step 4 Optimization NN Structure and Input Training Samples.

Rotor test-bed is shown in Fig.3. 13 eddy current sensors are arranged in different position in test-bed. Number 1 to 12 are rotor vibration signal. Number 13 is speed signal. Testing data of speed=3200r/s are selected as calculated data. 5000 data samples of one of the channels are selected to calculate four types of spectral entropy. 15 sets of data are selected to optimize hidden layer nodes of NN. 60 sets of data are selected to train NN. 10 sets of data are used to test NN. Learning rate of NN is 0.05.

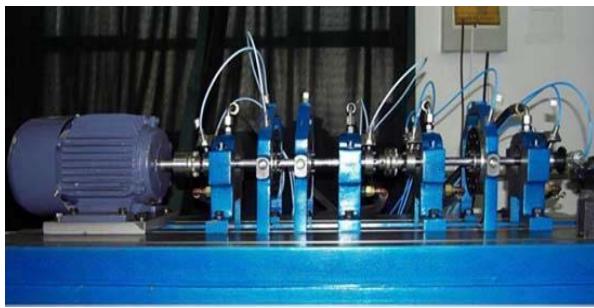


Fig.3 A test bench for rotor

Step 5 Initialization of GA.

Crossover rate P_c is 0.6. Mutation rate P_m is 0.1, and iterations are 100 in GA.

Step 6 Fitness Evaluation.

Fitness function of GA is $F=1/E$. Output error E of NN is:

$$E = \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^C (y_{ij} - O_{ij})^2 \quad (10)$$

Where N is number of test samples. y_{ij} is i output node ideal value of sample j.

If meet iteration times, new population will be moved to HPSO-NN. If not, the program will execute step 7.

Step 7 Inheritance.

Roulette and random selection was applied to individual fitness. Elitism is used to reserve optimal individuals. It makes the evolution process to high probability converge to the global optimal solution. In order to prevent premature, SA and crossover operator was used to individual $X_i(t)$ and $X_j(t)$:

$$\begin{cases} X_i(t+1) = aX_i(t) + (1-a)X_j(t) \\ X_j(t+1) = aX_j(t) + (1-a)X_i(t) \end{cases} \quad (11)$$

Where a is the random number in range (0,1) uniformly. t is iterations.

For results of crossover, fitness of $X_i(t)$ and $X_j(t+1)$ was compared according to Metropolis acceptance criteria. If fitness $X_i(t+1)$ is the optimal. $X_i(t)$ is replaced. Or for the random number r in range (0,1). If the under formula is workable.

$$e^{\frac{F_i(t+1)-F_i(t)}{T}} > r \quad (12)$$

$X_i(t+1)$ was accepted. Where $F_i(t+1)$ is fitness of $X_i(t+1)$. $F_i(t)$ is fitness of $X_i(t)$. T is SA temperature. SA temperature function is $T = T_0 \times \alpha^M$. T_0 is initial temperature of SA. α is temperature attenuation coefficient. M is the operation execution number of crossover. Initial temperature of SA T_0 is 100000. Temperature attenuation coefficient α is 0.95.

The same principle was applied to selection of $X_j(t)$ and $X_j(t+1)$. Some individuals are selected randomly to change gene value of using mutation probability P_m . Multipoint random mutations was used in this paper.

The final population as HPSO initial position vector was gotten when training reached iteration of terminate evolution.

HPSO-NN**Step 8 Initialization of PSO.**

Learning factor C_1 and C_2 of HPSO are 1.4962. w_0 is 0.9. Max. speed of particle is 0.2, iterations are 200.

Step 9 Modified Inertia Weight.

Inertia weight was calculated using the following formula [13]:

$$w = w_0 + r_1 \times e + r_2 \times a \quad (13)$$

Step 10 Crossover and Mutation, Evaluation Fitness, and Update Speed and Position.

Fitness function and population size selection of HPSO are same with GA. Particle fitness of PSO was sorted according to the current position fitness of every particle. Half outstanding particle fitness is chosen directly into the next generation. Same with GA, a crossover position is randomly generated for particle position vector through crossover of other half particles. Same number of offspring is gotten. Update is conducted after crossover finished. The fitness of calculated offspring was compare with corresponding father generation. Half outstanding particle fitness of offspring and parent was kept. Half outstanding particle in the kept particle and the original half particle was selected according to fitness sequence in order to keep populations number. Individual optimal value and global population optimal value of every particle was calculated. Speed and position of particle was updated according to formula (2) and (3).

Step 11 Max. Iterating Times or Min. Error Precision?

If training reached Max. iterating times or minimum error precision. The training process finished. If not, the program will move to step 9. [17-18].

III. NUMERICAL SIMULATION

Fig. 4 is iteration error of the three algorithms. Green dashed, blue dotted lines and red solid line respectively represent training process of GA, PSO and GA-HPSO. Before 70 times, GA and GA-HPSO are of obvious advantage. As the iterations As the increase of iteration times, GA-HPSO has an obvious superiority over GA and small advantage over PSO. Overall, PSO and GA have better convergence than GA-HPSO.

Testing samples are used to test the trained NN. Table 1 is the fault data classification results. Calculation formula of relative error is:

$$E_i = \frac{|O_i - y_i|}{|y_i|} \times 100\% \quad (14)$$

Where O_i and y_i are classification result and ideal output of testing sample i. Four kinds of algorithms are used to train NN. The trained NN is used to predict six kinds of typical fault of rotor. The forecasting results are

compared with idea output. The results show that GA-HPSO can rapidly and efficiently realize NN training. The novel algorithm shows high diagnostic accuracy and good classification effect in fault prediction.

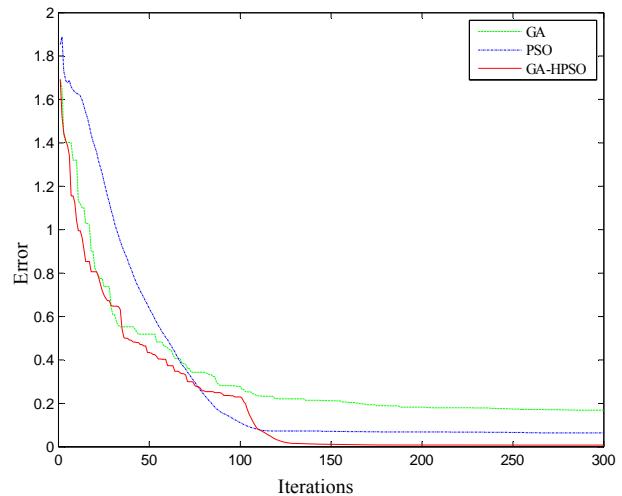


Fig.4 Iteration error of three kinds of algorithmic

TABLE I.
THE FAULT DATA CLASSIFICATION RESULTS

No.	Ideal Output	GA-HPSO Output	Relative error(%)	GA-BP Output	Relative error(%)	PSO-BP Output	Relative error(%)	BP Output	Relative error (%)	Fault Types
1	0.5	0.5228	4.56	0.5240	4.80	0.5246	4.92	0.5107	2.14	mass unbalanced
2	0.6	0.6130	2.16	0.6253	4.22	0.6262	4.37	0.6406	6.78	contact rubbing
3	0.4	0.3975	0.63	0.3998	0.05	0.4015	0.38	0.4096	2.40	rotor misalignment
4	0.7	0.7101	1.44	0.7090	1.29	0.7085	1.21	0.7162	2.31	loose
5	0.5	0.5246	4.92	0.5248	4.96	0.5254	5.08	0.5051	1.02	mass unbalanced
6	0.8	0.7980	0.25	0.8031	0.39	0.8054	0.68	0.8167	2.09	oil film whirl
7	0.8	0.7986	0.18	0.8055	0.69	0.8075	0.71	0.8097	1.21	oil film whirl
8	0.4	0.3984	0.40	0.3999	0.03	0.4017	0.43	0.4081	2.03	rotor misalignment
9	0.9	0.8945	0.61	0.9130	1.44	0.9175	1.94	0.8987	0.14	oil whip
10	0.4	0.3891	2.53	0.3867	3.33	0.3863	3.43	0.4008	0.20	rotor misalignment

IV. CONCLUSION

Four kinds of spectral entropy of fault diagnosis are the input variable of NN. The output values are six kinds typical fault of rotor. GA-HPSO are used to optimize weight and threshold of NN and compared with GA and PSO. The testing results show that the novel algorithm can realize quicker and more accurate fault diagnosis of rotor system than PSO and GA. GA-HPSO can give idea output results for multiple-diagnosis symptom. It shows that the novel algorithm is feasible and superior and supplies a new way and method for fault diagnosis.

ACKNOWLEDGMENT

This work supported by Natural Science Foundation (Grant No. 50875118), Scientific Research Foundation for the Returned Overseas Chinese Scholars of GANSU province (Grant No. 1002ZSB114) and Natural Science foundation of GANSU province (Grant No. 1112RJZA003).

REFERENCES

- [1] Xiu-yie Wei, Hong-xia Pan, and Qingfeng Ma, "Neural Network optimized by Particle Swarm Optimization and its application in fault diagnosis," *Journal of Vibration, Measurement & Diagnosis*, vol. 26, no. 2, pp. 133-137, 2006.
- [2] Xu ya-xiang, "Particle swarm algorithm and the application in neural network classification," *Xi'an: Xi'an Electronic Science and Technology University Press*, 2008.

- [3] K.Premalatha, and A.M.Natarajan, "Hybrid PSO and GA for Global Maximization," *Int.J.Open Problems Compt.Math*, vol. 2, no. 4, pp. 597-608, 2009.
- [4] Chang-cai Cui, Shi-wei Fu, Fugui Huang, and Bing Li, "Optimization algorithm GA-PSO and its application technologies in engineering," *Chinese Journal of Scientific Instrument*, vol. 28, no. 8, pp. 296-299, 2007.
- [5] Wen-Yi Wang, Guang-Jun Qin, and Ruo-Yu Wang, "Research on Genetic Algorithm Based on Particle Swarm Algorithm," *COMPUTER SCIENCE*, vol. 34, no. 8, pp. 145-147, 2007.
- [6] Jun Liu, Xiao-hong Qiu, Zhi-yong Wang, and Peng Yang, "Research on PSO algorithm in neural network generalization," *Computer Engineering and Applications*, vol. 45, no. 29, pp. 34-36, 2009.
- [7] Lin Gao, Hai-rong Sun, and Huai-shen Yang, "Identification method combining PSO and improved BP neural network," *Computer Engineering and Applications*, vol. 46, no. 21, pp. 37-39, 2010.
- [8] Fei Chen, Shu-hong Huang, Tao Yang, Wei Gao and Guo-qiang Huo, "Information Exergy Diagnosis Method of Vibration Faults of Rotating Machinery," *JOURNAL OF MECHANICAL ENGINEERING*, vol. 45, no. 11, pp. 65-71, 2009.
- [9] Cheng-zhi Cao, Mei Dong, Min Li, and Bo Zhou, "Design of Genetic BP Network Speed Identifier and Application in Direct Torque Control System," *Journal of System Simulation*, vol. 19, no. 4, pp. 925-927, 2007.
- [10] Gen, M. and Cheng, R., "Genetic Algorithms and Engineering Design," *John Wiley & Sons, Inc., Canada*, 1997.
- [11] Tad Gonsalves, and Kiyoshi Itoh, "Simulated Annealing In The Optimization Of Collaborative Systems Operation," *Journal of Integrated Design & Process Science*, vol. 10, no. 3, pp. 87-95, August 2006.
- [12] Yi Da, and Ge Xiurun, "An improved PSO-based ANN with simulated annealing technique," *Neurocomputing*, vol. 63, pp. 527-533, 2005.
- [13] Hong-Xia Pan, Jin-Ying Huang, Hong-Wei Mao, and Zhen-wang Liu, "Failure feature extraction study based on Particle Swarm Optimization," *Journal of vibration and shock*, vol. 27, no. 10, pp. 144-149, 2008.
- [14] B.Samanta, and C.Nataraj, "Use of particle swarm optimization for machinery fault detection," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 2, pp. 308-316, March 2009.
- [15] X H Shi,Y.C.Liang,H.P.Lee,C.Lu, and L.M.Wang, "An improved GA and a novel PSO-GA-based hybrid algorithm," *Information Processing Letters*, vol. 93, pp. 255-261, 2005.
- [16] Xu Yi-shan, Zeng Bi, Yin Xiu-wen, and Lu Bo-sheng, "BP Neural Network and its applications based on improved PSO," *Computer Engineering and Applications*, vol. 45, no. 35, pp. 233-236, 2009.
- [17] Amir Hosein Zamanian, and Abdolreza Ohadi, "Gear Fault Diagnosis Based on Gaussian Correlation of Vibrations Signals and Wavelet Coefficients", *Applied Soft Computing*, In Press, July 2011.
- [18] Jiuzhong Zhang, and Xueming Ding, "A Multi-Swarm Self-Adaptive and Cooperative Particle Swarm Optimization," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 6, pp. 958-967, September 2011.



Bin Peng was born in Yining, Xinjiang Province, China in 1976. He received his M.S. degree in Mechanical Engineering from Lanzhou University of Technology, China, in 2003, and his Ph.D. in 2007. Currently he is an associate professor of Mechanical Engineering at Lanzhou University of Technology, China. His primary research interest is Fault Diagnosis, Scroll Machinery and CAD/CAM/CAE.



Chang-an Hu was born in Linyi, Shandong Province, China in 1986. He received his B.S. degree in Mechanical Engineering from Linyi University, China, in 2009. Currently he is a graduate student of Mechanical Engineering at Lanzhou University of Technology, China. His primary research interest is Intelligent Evolution Algorithm.



Rong-zhen Zhao was born in Zaozhuang, Shandong Province, China in 1960. She received her M.S. degree in Mechanical Engineering from XI'AN JIAOTONG UNIVERSITY, China, in 1996, and her Ph.D. in 2006. Currently she is a professor of Mechanical Engineering at Lanzhou University of Technology, China. Her primary research interest is Fault Diagnosis and Mechanical System Dynamics.

Yu-qiao Zheng was born in Zhuanglang, Gansu Province, China in 1977. She received her M.S. degree in Mechanical Engineering from Lanzhou University of Technology, China, in 2009. Currently she is a lecturer of Mechanical Engineering at Lanzhou University of Technology, China. Her primary research interest is Fault Diagnosis.

A Novel Approach to Hardware/Software Partitioning for Reconfigurable Embedded Systems

Linhai Cui

School of Software, Harbin University of Science and Technology, Harbin, China
Email:cuilinhai@hrbust.edu.cn

Abstract—Hardware/software partition is a crucial point in the design of a reconfigurable embedded system. Reconfigurable computing is a promising approach to overcome the traditional trade-off between flexibility and performance in the design of computer architectures which adapt their hardware to each application to achieve a high performance of dedicated hardware. In this paper, some hardware and software partitioning algorithms were analyzed and summarized first, then a innovative algorithm for task partition and scheduling is proposed based on new features of reconfigurable hardware such as dynamic reconfiguration and the delay of reconfiguration. In the proposed algorithm, a large-scale application is decomposed into multiple sub-tasks of suitable granularity and each sub-task has constraint relationship with each other. And a directed acyclic graph (DAG) which presents the relationship between tasks was drawn according to the execution order of tasks. Then the specific application presented in the DAG is mapped to the hardware and software platform by a strategy called GATS which combine the Genetic Algorithm and the Tabu Search algorithm together. The shortest time of assignment and task execution order can be found by the priority-based scheduling method. The experimental results show that the method is of high performance and can effectively mapping the application task to the reconfigurable system.

Index Terms—Reconfigurable embedded system, Task scheduling, Hardware/software partitioning, Genetic algorithm; Tabu search algorithm

I. INTRODUCTION

Most modern electronic systems are composed of both hardware and software. Embedded system is some combination of computer hardware and software to perform a particular function. It can be found in many applications such as automobiles, telecommunication systems, intelligent home devices, medical equipments, and in systems for the military.

Comparing to the hardware parts, the software parts are much easier and faster to develop and modify. Thus, software is less expensive in terms of the development cost and time. Hardware, however, provides better performance. For this reason, an embedded system designer's goal is to minimize the weighted sum of the software delay, Hardware area, and power consumption.

There are two basic implementations of embedded system. One is hardware and the other is software. The hardware method achieves system functionality through the design of dedicated hardware logic circuits while the software method is based on microprocessor software to complete the system functions through the design of the program.

The main task of hardware/software partition is to assign the system functions to the target structure on the software and hardware domain under the condition of meeting the design constraints, and its essence is a kind of combination optimization problem. It includes the following three aspects: first, processing unit allocation, i.e. to determine the type and number of required software and hardware processing unit; second, task assignment, i.e. to assign tasks and communication to the target structure in the processing unit and communication resources to execute and to meet the performance and cost constraints; third, task scheduling, i.e. to determine the order of execution and the start time of the assigned task and communication in each processing unit to meet the dependency relationship of control and data between the system tasks. The solution space is a huge multi-dimensional non-contiguous space, so it is difficult to solve if taking the solution quality and solution time into account. Therefore, only the execution time, cost, power and other major overhead are considered when studying the hardware and software partition to reduce the difficulty of solving the whole problem by simplifying the model of the target structure.

Most embedded systems use CPU (Central Processing Unit) + ASIC (Application Specific Integrated Circuit) structure. For the systems which only have a hardware processing unit (ASIC) and a software processing unit (CPU), it is relatively simple to partition the system, and it is called binary partitioning. For the systems which have multi-processing units, the hardware processing units and software processing units may not be the same, and it may be more complicated to partition the system, such problems are called multi-way partitioning.

In addition to the CPU + ASIC structure, FPGA (Field Programmable Gate Array)-based reconfigurable hardware system has been developed. There are two kinds of reconfiguration concern to the time when

reconfiguration take place. One is static reconfiguration and the other is dynamic reconfiguration. More and more studies have been focused on dynamic reconfigurations. And a growing number of embedded systems employ dynamic reconfigurable architecture. In order to use the dynamic reconfiguration efficiently, one needs a support of operating systems to manage both software and hardware. Therefore, the structure of the traditional CPU + ASIC method is no longer suitable to apply to software and hardware reconfigurable systems.

Reconfigurable hardware components such as FPGAs are used more and more in embedded systems, since such components offer a sufficient capacity for a complete SoC (System on a Chip) or even NoC (Network on a Chip). The advantage of a reconfigurable hardware platform over pure software reconfiguration is that it can provide a system implementation flexibility to be adaptable to new functional requirements while meeting constraints for critical system parameters such as data throughput rates and latencies.

When doing partition, we need not only to assign the system tasks to the software or hardware domain, but also need to divide the tasks which is possibly assigned on the reconfigurable devices into different segments which is not overlapping in time. Measures should be also taken to reduce the delay caused by the reconfiguration when designing a reconfigurable embedded system.

In this paper, we proposed a portioning algorithm called GATS which employs the advantages of traditional algorithms such as Genetic and Tabu, and a task schedule approach by using DAG.

II. RELATED WORK

With the development of integrated circuit technology, embedded system is moving towards small size, mobile, portable, light in weight, low power consumption, more complex and so on. Traditional design approach has become a bottleneck restricting the development of embedded systems. Software design and hardware design are required to be integrated closely and coordinated with each other. This leads to the development of a new design theory - hardware and software co-design.

A. Evolution of Hardware/software Partitioning

The studies of hardware/software co-design began in the early 1990s, the idea of hardware/software co-design is formally proposed in the first International Workshop on Hardware/Software Codesign (CODES), held in 1992. Then many famous universities set up research group on embedded systems engaged in software and hardware co-design theory and research. Some EDA vendors have also introduced some tools supporting hardware and software co-design.

SOS system [1], developed by Prakash and Parker from the University of Southern California, is the first hardware and software co-design systems in the world. The system can schedule tasks on multiple processors, but it was slow, not suitable for large-scale systems.

The COSYMA (Co-synthesis for Embedded

Architecture) [2] system, developed by the German Technical University of Braunschweig, is mainly restricted to a single processor and a single ASIC system. Its partition method is mainly for software to optimize the calculation through co-processors. The main drawback of COSYMA is that the processor and the coprocessor can not work concurrently [3].

The Corsair system [4], developed by Frank Slomka, is an embedded system design environment suitable for multi-processor and multi ASIC structure. The system generates the system model by using tabu search algorithm. But it uses static method to assess the system performance when the system model is generated, so it can not evaluate complex systems.

In 1997, Eles proposed to achieve the hardware and software partition by using simulated annealing and tabu search algorithm. He described a model called condition task graph using list scheduling algorithm to realize the structure for each processing unit to form the scheduling table and as a basis for selection of software and hardware. Experiments result shows that the tabu search algorithm is more suitable for hardware and software partition compared to simulated annealing, [5].

In 1999, Henkel introduced the IP-based low-power embedded systems hardware and software partitioning algorithm. The idea is to reduce the system's idle time to reduce the power consumption [6]. In 2002, Theerayod Wiangtong and others compared and analyzed the three heuristic algorithms of hardware and software co-design and found that tabu search algorithm is proved to be superior in hardware and software partition to the genetic algorithm and simulated annealing algorithm [7]. In 2006, Michalis D Galanis proposed a hybrid reconfigurable system [8].

B. Research on Task Scheduling

Scheduling problem is a kind of combinational optimization problem and is applied in many computer and communications fields. It has a close relationship with the algorithm design, complexity theory.

The research related to task scheduling was first proposed by Liu and Layland [9] which ignores some implementation details. It is the basis of many real-time task models, and extends up to the processor environment with task scheduling and feasibility analysis in algorithm design.

Reference [10] proposed a grouping appropriate algorithm, which is a non-dynamic scheduling algorithm for periodic tasks. Due to the use of the grouping strategy and appropriate scheduling policy, the utilization of platform resource and processor is increasing in recent years.

Hsu Heng Ruey from China Taiwan National University researched real-time periodic task scheduling problem by using dynamic voltage scheduling technique for a given energy constraints [11]. In recent 20 years, the research on the parallel task scheduling on multiprocessor represented in directed acyclic graph DAG (Directed Acyclic Graph) has been developed rapidly. DAG-based task scheduling is to map the distribution of tasks to processors and coordinate the

implementation. Under the condition of meeting constraints, the overall execution time, power consumption, area, and other indicators of the task are best. It is NP-complete problem.

Becchi and Crowley thought that task management is the key to raising computing performance of a multi-processor platform, and they developed a run-time monitoring program to capture the dynamic behavior of the process, allowing the process to migrate between multiple processors. Experiments show that to use dynamic process allocation method on a heterogeneous multi-processor platform can significantly improve overall system performance [12]. In recent years, some researchers began to use some new methods to solve the multiprocessor scheduling problems such as genetic algorithm for multi-processor task scheduling [13].

The introduction of the new method of calculation improves the solution accuracy. But the efficiency of the algorithm need to be further improved. For the task scheduling on CPU + FPGA structure, more research focused on how to place the hardware tasks on the FPGA dynamically. However, less corresponding research are made in task allocation, task migration and other issues of mixed task scheduling [14].

III. RECONFIGURABLE SYSTEM BASED ON FPGA

With the emergence of programmable device, especially field-programmable gate array (FPGA), the reconfigurable technology is developed rapidly in embedded applications. The development of reconfigurable technology makes the traditional boundaries over hardware and software blurred.

The so-called reconfigurable means that in an information processing system under control of software, if the system can be reformed into a different information processing systems to adapt to different application requirements by using reusable resources, the information processing system is called reconfigurable [15].

By using reconfigurable technology, the system can be realized in software and hardware in the case only a little more resources are needed. On one hand, the calculation task can be accomplished by building a dedicated hardware circuit on FPGA, similar to the ASIC. On the other hand, different tasks can be optimized by building different circuit on FPGA.

Reconfigurable Systems based on Large-scale programmable device, FPGA, perform reconfigurable of circuitry at runtime by using the features of can be repeated programming and configuration of FPGA. It can dynamically change the circuit structure while a real-time electronic systems work. Its essence is to achieve the time-sharing reuse of all or part of the internal FPGA logic resources. It can make the logic circuits which are discrete in time works in order on the same FPGA.

A. Dynamic Reconfigurable Technology

For dynamic reconfigurable, a special study group RAW introduced the following description in 2005 [16]:

The characteristics of dynamic reconfigurable are that the hardware architecture or devices can quickly change (while the system running) its functions and connections.

As shown in Figure 1, the three key issues for the research on dynamic reconfigurable system are the hardware platform, the mapping from specific application to the hardware platform, and the controls needed during the running of the system.

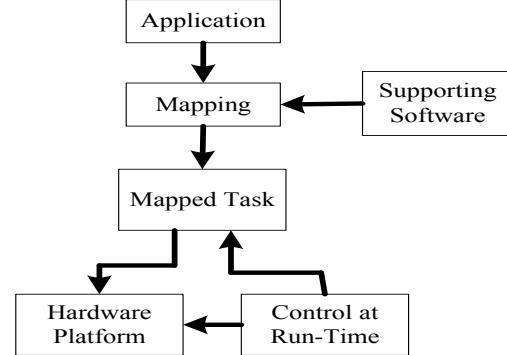


Figure1. Research contained in dynamic reconfiguration system

According to the reconfiguration ways of the reconfigurable logic, the hardware supporting dynamic reconfiguration can be divided into context configure devices and part reconfigurable configure device [17]. Typically, the variable and time-consuming parts of the system are implemented in hardware and the complex controls and data structures are implemented in software [18].

B. Unified Management of Hardware and Software

The emphasis of unified management of hardware and software of dynamically reconfigurable system is focused on the management of the hardware tasks. The hardware modules are converted into hardware tasks, and to be managed under the operating system, and then the unified management by the extended operating system is implemented.

In order to convert hardware modules into the hardware tasks, certain constraints need to be imposed on the tasks first to enable it to response to the basic communication and control primitives in the operating system making the user call the hardware tasks as normal software tasks.

After the completion of the hardware tasks building, the expanded operating system will be able to manage them. The next step is how to manage that task, that is when the hardware can be downloaded into a piece of programmable logic resources.

In short, the difference between hardware tasks and software tasks need to be fully taken into accounts when they are managed under the operating system. Some measures must be taken to reduce the preparation time such as by pre-configured configuration, scheduling, etc.

IV. PARTITION STRATEGY

Hardware/software co-design is to give an algorithm which can automatically search for the best compromise point between the hardware and software under certain

constraints and to produce the actual system architecture.

A. Dynamically Reconfigurable System Modeling

There are significant differences in performance and cost between software and hardware implementation. Hardware and software partition is one of the key issues in co-design. Its goal is to maximize resource utilization, and to minimize application execution time under the constraints to meet the time and shared resources conflict conditions constraints.

A typical dynamic reconfiguration system is shown in Figure 2. It consists of microprocessor, configuration controller, reconfigurable hardware (FPGA), memory and configuration file memory. To rescue the microprocessor from the task of configure FPGA, and to make them perform parallel computing, an additional configuration controller is added to configure the FPGA.

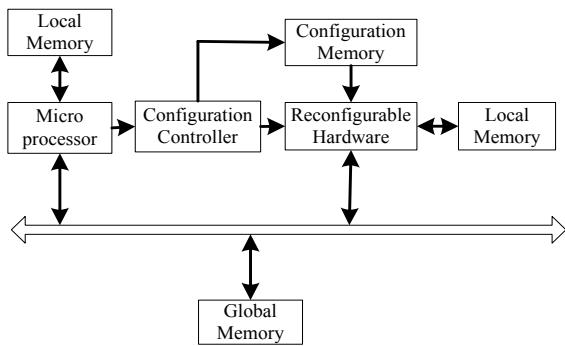


Figure2. Dynamic reconfiguration system

When the microprocessor executes a hardware configuration instruction, data is passed to the configuration controller, and the latter retrieve the corresponding configuration data from the configuration memory, and download it to the FPGA to complete the configuration. Microprocessor and the reconfigurable hardware communicate by sharing memory, so the two tasks located on the microprocessor and reconfigurable hardware not only need time for data transfer and communication, but also need data reading and writing time.

The calculation model must be considered as an important element in the hardware and software partition. Different levels of abstraction of the system forms the different calculation model, and the partition can be performed in a different granularity. These models have different features and application areas, and they can be divided into the following categories: finite state machine model, data flow diagrams, Petri nets, data / control flow graph, task flow diagram.

Task flow diagram, also called a directed acyclic graph (DAG), is the behavior level description of the system. Its purpose is to describe the controls between the tasks in the reconfigurable system, data relationship and the cost information for each task. It has nothing to do with the system architecture. It can be expressed as a triple $G = (V, R, E)$, as shown in Figure 3, the V on behalf of the task node set, R is the edge connecting node, the node E on behalf of its right to inter-communication.

Each node in the diagram represents a system task (or

a function module), including its software, hardware, cost information, the edge represents the data flow or control relationship between tasks, and its weight is on behalf of the communication overhead between two tasks.

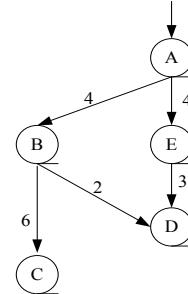


Figure3. Task flow diagram

B. Hardware and Software Partition

Functional partition between hardware and software and its implementation vary with the applications. The solution is to find the time / space mapping from the application described with the task flow diagram to the reconfigurable system consist of microprocessor and reconfigurable hardware.

There are three main methods to solve the hardware and software partition problem. The first one is planning method which uses planning to partition the problem and to get the optimal solution. Its drawback is the high computational complexity and large memory overhead. The second method is construction approach. The construct method compares the pros and cons of the solution layer by layer first, and then the optimal solution will be drawn from each comparison group to obtain the optimal solution, such as poly-clustering technology. The advantage is that you can find the optimal solution or near optimal solution more efficiently. The third one is search methods. It includes local search method and oriented random search method. By searching its neighborhood and replacing the current solution to achieve the optimization.

Oriented random search methods include tabu search, simulated annealing, genetic algorithms, and etc. The main idea is to use some guidance rules to guide the search within entire solution space for good solutions. Randomly oriented search method does not rely on objective information, and can be used to solve complex issue. At present, most researchers have adopted oriented random search method.

C. Traditional Partitioning Algorithm

Genetic Algorithm (GA) is an algorithm based on Darwin's biological evolution theory. It simulates the biological mechanisms of natural selection and genetic search heuristic. Each element in the solution space is encoded and then divided optimal solution space into groups by iteration to find the optimal solution. Crossover and mutation operator are the two most important components of GA hardware and software partition which are repeatedly applied to the solution of the problem encoding and form chromosomes.

As a global optimization search algorithm, genetic

algorithm is simple, universal, robust, and has wide applications. It can meet the requirements of multi-partition, and it has become the key technology to deal with traditional search methods to solve difficult, complex and nonlinear problems. But it also has many shortcomings, such as the genetic algorithms search the solution space in parallel way. This makes it have strong global search ability, but the local optimization ability is poor. The premature convergence phenomena may occur.

Tabu search (TS) is a algorithm simulating human intelligence process. It is the expansion of the local field search, and is a global step by step optimization algorithm. A flexible memory structure and the corresponding Tabu search criteria was introduced to avoid circuitous search.

Fields, tabu table, tabu length, candidate solutions and amnesty criteria are the key factors in tabu search algorithm design. There are two tables in the tabu search algorithm which impact the performance of search algorithm: the length of taboo objects. Compared with traditional optimization algorithms, tabu search algorithm has flexible memory and amnesty guidelines, and in the search process, it can accept inferior solutions, with a strong climbing ability, it can jump out of local search optimal solution, turn to the other regions of solution space, leading to better probability of a global optimal solution. It is a strong local search global iterative optimization algorithm.

However, the tabu search algorithms have significant deficiencies, such as: strong dependence on the initial solution, a good initial solution can make the tabu search in the solution space to search for good solutions, and poor initial solution will reduce convergence speed of tabu search. It is not efficient to search the solution space in individual, serial way. Its global search capability is not strong.

D. Proposed Partitioning Algorithm

After analyzing the genetic (GA) and Tabu Search (TS) algorithm, a hybrid approach based on GA and TS strategy, called GATS was proposed. It can map the specific application to a software and hardware platforms under the reconfigurable system resource constraints and other conditions.

The main idea of GATS is to make the TS as a mutation operator of GA, TSM. It employ the strong memory function and the climbing ability of TS to overcome the weaknesses of poor climbing ability of GA. And it remains the advantage of multi starting point maintaining in the GA.

Genetic Algorithm can not be directly used in the solution of the problem space. So what you need to do is to encode the solution by using chromosome. Encoding is the most important thing required to apply genetic algorithms to solve the problem, and it also is a critical step in designing genetic algorithm. The encoding method will directly affect the solution quality, restrict the choice of genetic manipulation. For a single microprocessor and single reconfigurable hardware architecture, the binary encoding makes the encoding and decoding simple, and also makes the crossover

operation and mutation operation easy to implement.

Each task is represented as a binary genes, gene is mapped to the microprocessor if its value is 0, and it is mapped to the reconfigurable hardware if it is 1. The chromosome is a vector whose length is N standing for the number of the tasks. The coding chart of tasks is shown in Figure 4. This encoding method makes the crossover and mutation operation easy to use and does not produce invalid chromosomes.

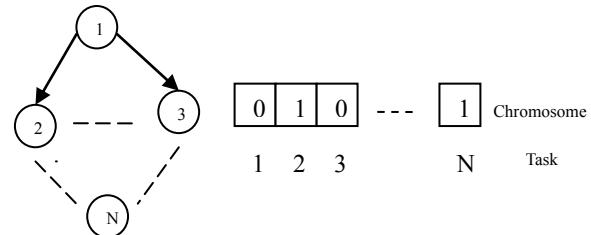


Figure4. Coding Chart of Tasks

For reconfigurable hardware, the above logic resources are part dynamically reconfigurable at runtime. So that for any a different chromosome, the number of different modules in reconfigurable hardware, the process of reconfiguration, the time of reconfiguration, and the calculation overlap will change according to the assigned tasks on reconfigurable hardware that is genetically different.

The choice of fitness function can be the first target that is based on algorithm to optimize the overall system running time T and the cost to construct a generalized objective function C , then by scaling the objective function to be generalized. Running time is converted to the general penalty term in the objective function. Note that the choice of penalty term should ensure that the results meet the constraint requirements, and partition results must also ensure the full utilization.

In addition, costs and the values of run-time may vary on the scope and magnitude. Normalization is needed. Equation (3-1) and (3-2) were used to introduce the normalization factor σ_c and σ_t .

$$\sigma_c = CostHw - CostSw \quad (3-1)$$

$$\sigma_t = \max(TimeSw - TimeReq, TimeReq - TimeHw) \quad (3-2)$$

Where $CostHW$ stands for the cost when all the nodes are implemented in hardware, and $costSW$ stands for the cost when all the nodes are implemented in software. $TimeHW$ and $TimeSW$ stand for hardware execution time and software execution time respectively. $TimeReq$ is the constraint time provided by the designer, and its value is between $TimeHW$ and $TimeSW$. So the generalized objective function can be defined as:

$$COBJ = \alpha \exp\left(\frac{T - TimeReq}{M_i \sigma_t}\right) \left| \frac{T - TimeReq}{\sigma_t} \right| + (1 - \alpha) \frac{C}{\sigma_c} \quad (3-3)$$

And the fitness function is defined as:

$$Fitness = \frac{1}{1 + COBJ} \quad (3-4)$$

The C and T in formula (3-3) stand for the partition cost and execution time respectively. And α here is used to adjust the relative weight of C and T . Since

system performance determines whether the design is available, so α equals 0.6.

V. TASK SCHEDULING ALGORITHM

One of the key technologies of traditional operating system is task scheduling. In dynamically reconfigurable systems, especially in large-scale applications, the hardware task can not be configured to the reconfigurable device at one-time. Then scheduling becomes more important, and the scheduling algorithm have a direct impact on the system performance.

Task scheduling has two main purposes. One is to fully use the resources on the device and the second is to optimize the device configuration sequence. Scheduling should be done to have the tasks whose execution order are more close or simultaneously at one-time schedule to the device. To optimize the configuration sequence can reduce the configuration time overhead caused by the configuration process, and it can reduce the configuration time if the configuration time is too long to the system bottleneck effect.

Scheduling on reconfigurable hardware is also a constrained layout problem. What you need to do is not only to find out the schedule start time, but also to identify the layout position of tasks in reconfigurable hardware under certain conditions and resource constraints.

In summary, the purpose of this section is to find the shortest time of assignment in entire task flow diagram and the task execution order based on the partition result by using genetic and tabu search algorithm.

A. Scheduling Algorithms based on DAG

In practice, for any DAG graph in which the weight values of the nodes and edges can be any value, heuristics method is the first choice for solving those DAG scheduling problems. To sum up, the current DAG scheduling algorithm can be divided into four categories: List Scheduling algorithm, Clustering scheduling algorithm, Scheduling algorithm based on Task Duplication, and random search technology.

The basic idea of list scheduling algorithm is to sort the priority of node, to construct a scheduling list, so that all the ready tasks are in the schedule list, and then the highest priority task is selected from a list and put it into idle computing resources began to run.

The basic idea of Clustering scheduling algorithm is that if there are infinite numbers of processors, the DAG task graph nodes are taken as a cluster when starting scheduled, and merger all of these clusters without increasing the overall task completion time until there is no cluster can be merged finally.

The basic idea of the scheduling algorithm based on task duplication is to duplicate the precursor mission when the processor is idle to avoid some of data transmission from the precursor task, thus to decrease the gap of waits time for the processor.

Random search technique is mainly driven by a random choice to search for the problem solution. The

searching results are better than other algorithms, but its scheduling length is longer, so it is not employed so much.

B. Configuration Prefetch Scheduling Algorithms

Dynamically reconfigurable hardware architecture or device is characterized by changing its functionality and connectivity rapidly. The computing time of a system employ dynamic reconfiguration can be divided into work time and preparation time. The real effective operation time is the time required for the module and the communication time between modules. Preparation time is the delay caused by the configuration between the function switching.

Now the bottleneck of a dynamically reconfigurable system is that the preparation time is too long. The preparation time can be shortened by hiding the critical software configuration time in addition to improving the speed of reconfigurable devices.

The basic idea is that to configure the successor node in advance when node in the DAG to be scheduled for execution. And save the configuration node that need to be configured but because of the FPGA configuration port is occupied and can not start immediately by using a configuration wait queue.

VI. EXPERIMENTAL RESULTS

In order to verify the hardware/software partitioning algorithm presented here and the effectiveness of the configuration scheduling strategy, we perform the following software simulation.

First, we use the TGFF tool (Windows Version) to randomly generate the task flow graph in which 30, 40, 50, 60, 70, 80 nodes are included. Each node contains hardware and software implementation costs, hardware resources area, and other information, reconfiguration time and the area occupied by the specific tasks related to the number of resources.

For the same computing tasks, a reconfigurable hardware implementation is usually 10 times faster than using a microprocessor[23], it is assumed that for each task on the processor, the average execution time of reconfigurable hardware in the average execution time of 10 times. Simulation environment for hardware and software test are Inter 1GHz processor, 512MB RAM, Linux operating system, GNU compiler. And assuming that the target system consists of a single processor and Virtex II series xc2v1000 FPGA which has 1280 CLBs. Table1 shows the comparisons between the three algorithms used to get the fitness value.

Table1. The best fitness value comparison of GA, TS and GATS

	30	40	50	60	70	80
TS	0.853600	0.852139	0.853721	0.860228	0.863175	0.864006
GA	0.848012	0.848369	0.842773	0.846343	0.847218	0.845983
GATS	0.922487	0.924224	0.920106	0.914286	0.928246	0.949677

The relationship between the fitness values is shown in Figure5. It is clear that the values get from GATS are greater than that of the genetic algorithm GA and tabu

search algorithm TS. It shows that the GATS algorithm has the advantage of multi-start point and a strong hill-climbing ability. Although the running time of GATS algorithm is longer than the GA, TS algorithm, the accuracy is high. So the GATS algorithm can be applied in the applications demanding high accuracy.

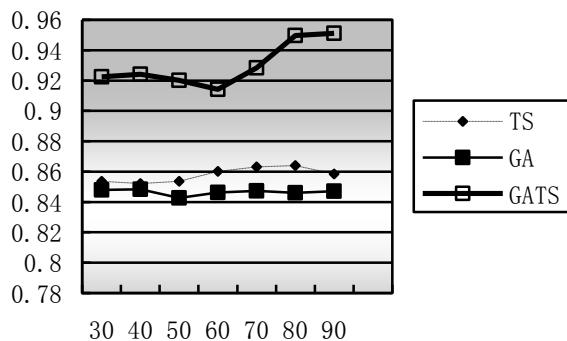


Figure 5. Fitness Value/Nodes Curve

Table 2 shows the results using these three algorithms in different scale applications, in which S and NS stand for the scheduling strategies with and without configuration prefetch respectively.

Table 2 Data comparison of GA, TS and GATS

Task	GA		TS		GATS	
	Node	NS	S	NS	S	NS
30	2071	1644	2058	1950	2056	1538
40	2845	2241	2835	2743	2809	2076
50	3680	3087	3573	3402	3362	2458
60	4316	3085	4311	4001	4310	2770
70	4944	3324	4806	4522	4776	3014

VII. CONCLUSIONS

This paper focuses on the hardware/software partition technologies in reconfigurable embedded systems. The characteristics of reconfigurable systems and the key issues involved in dynamic reconfigurable technology are analyzed in detail. A reconfigurable hardware architecture model consists of microprocessor, configuration controller, reconfigurable hardware (FPGA), memory, and configuration file memory is proposed. And it gives a directed acyclic graph DAG for reconfigurable embedded systems modeling. After comparing the advantages and disadvantages of GA and TS algorithms, a mixture of GA and TS strategy, called GATS algorithm is proposed. The GATS approach drawing on the strengths of genetic algorithms and tabu search algorithm respectively, and achieved good results.

Task scheduling algorithms based on DAG models such as configuration prefetch, priority-based scheduling algorithm, especially for the CPU + FPGA structure, are proposed. The scheduling algorithms are used to evaluate the system partitioned in GATS algorithm. The results show that it effectively reduces the reconfiguration time and the overall application execution time.

REFERENCES

- [1] Parkash S, Parker A C. SOS: Synthesis of Application-Specific Heterogeneous Multiprocessor Systems. Journal of Parallel and Distributed Computing, Vol.16, 1992, pp: 338-351.
- [2] J.Henkel and R.Ernst. High-Level Estimation Techniques for Usage in Hardware/Software Co-Design. In IEEE/ACM Proc of Asia and South Pacific Design Automation Conference, Yokohama, Japan, February 1998: 353-360.
- [3] JERNST, R, J HENKEL, and T. BENNER. Hardware-Software Co-Synthesis for Micro-Controllers[J], IEEE Design & Test of Computer, 1993, 10(4): 64-75.
- [4] Frank Slomka, Matthias Dorfel Ralf Munzenberger, Richard Hofmann. Hardware/Software Codesign and Rapid Prototyping of Embedded Systems. IEEE Design & Test of Computers, 2000, 17(2): 28-38.
- [5] P ELES, Z PENG, K KUCHCINSKI, et al. System Level Hardware/Software Partitioning Based on Simulated Annealing and Tabu Search [J]. Design Automation for Embedded Systems, 1997, 2(5): 5-32.
- [6] HENKEL J. A Low Power Hardware/Software Partitioning Approach for Core-Based Embedded System[C]. In: Proceedings of the 36th ACM/IEEE Conference on Design Automation Conference, 1999, 122-127.
- [7] THEERAYOD WIANGTONG, PETER Y.K CHEUNG, WAYNE LUK. Comparing Three Heuristic Search Methods for Functional Partitioning in Hardware-Software Codesign[J]. Design Automation for Embedded Systems, 2002(6): 425-449.
- [8] MICHALIS D GALANIS, ATHANASIOS MILIDONIS, GEORGE THEODORIDIS. A Method for Partitioning Applications in Hybrid Reconfigurable Architectures[J]. Des Automation Embedded System, 2006(10): 27-47.
- [9] Liu C , Layland J. Scheduling algorithms for multiprogramming in a hard real-time environment [J]. Journal of the ACM,1973,20(1):4-61.
- [10] R. Gupta and G. De Micheli. System-level synthesis using re-programmable components. Proceedings of EDAC, IEEE Press, 1992, 2-7.
- [11] Institute of Electrical and Electronics Engineers, New York, USA, IEEE Standard VHDL Language Reference Manual (IEEE Std 1076-1987), 1987.
- [12] J. Iyoda. PARTS: A Support Tool to Hardware/Software Partitioning. Master thesis, Federal University of Pernambuco, Brazil, May 2000 (in Portuguese).
- [13] L. Silva. An Algebraic Approach Hardware/Software Partitioning. PhD. thesis, Federal University of Pernambuco, Recife, Brazil, July 2000.
- [14] C.Steiger, H.Walder and M.Platzner. Operating Systems for Reconfigurable Embedded Platforms: Online Scheduling of Real-Time Tasks[J].IEEE Transactions on Computers, 2004, 53(11): 1393-1407
- [15] R.B. Hughes and G. Musgrave. The lambda approach to system verification. Hardware/Software Codesign, Kluwer Academic Publisher, 1996, 427-451.
- [16] A.Dehon. Comparing computing machines.Configurable Computing: Technology and Applications of Proc.Of SPIE, 1998, 3526:124-133
- [17] J. Iyoda, A. Sampaio, and L. Silva. ParTS: A partitioning transformation system. Proceedings of FM99(World Congress on Formal Methods), Vol. 1709, Lecture Notes in Computer Science, 1999, 1400-1419.
- [18] A. Kalavade and E. Lee. The extended partitioning problem: Hardware/software mapping, scheduling and implementation-bin selection. Design Automation for Embedded Systems, Vol. 2, No. 2, 1997, 125-163.
- [19] J. Madsen, J. Groge, P.V. Knudsen, M.E. Petersen, and A. Haxthausen. Lycos: The lyngby co-synthesis system. Design Automation of Embedded Systems, Vol. 2, No. 2, 1997, 195-235.
- [20] R. Niemann and P. Marwedel. An algorithm for hardware/software partitioning using mixed integer linear

- Programming. Design Automation of Embedded Systems, Vol. 2, 1997, No. 2, 165–193.
- [21] D. Saha, R.S. Mitra, and A. Basu. Hardware/software partitioning using genetic algorithm. Proceedings. of 10th International Conference on VLSI Design, India, 1997, 155–160.
- [22] A. Sampaio. An Algebraic Approach to Compiler Design. Vol. 4 of Algebraic Methodology and Software Technology (AMAST) Series in Computing, World Scientific, 1997.
- [23] A. Dehon. Comparing computing machines. Configurable Computing: Technology and Applications of Proc. Of SPIE, 1998, 3526:124-133
- [24] L. Silva, A. Sampaio, and G. Jones. Serialising parallel processes in a hardware/software partitioning context. Proceedings of Formal Method Europe (FME) 2001, Vol. 2021, Lecture Notes in Computer Science, 2001, 344–363.
- [25] DICK R P, RHODES D L, WOLF W. TGFF: Task graphs for free. Proc. Int. Workshop Hardware/Software Co-design [M]. Mar. 1998: 97–101.

Linhai Cui was born in HeiLongJiang, China, in 1961. He received the B.S. degree in computer science from the Xi'an JiaoTong University, China, in 1983. In 2006, he received the M.S. degree in electrical engineering from the Northwestern Polytechnic University, USA. He is an Associate Professor at the Harbin University of Science and Technology, China.

A Study on the Control Methods Based on 3-DOF Helicopter Model

Junshan Gao

School of Automation/Harbin University of Science and Technology, Harbin, China
Email: junshangao@hrbust.edu.cn

Xinghu Xu, Chen He

School of Automation/Harbin University of Science and Technology, Harbin, China
Email: xxh5g@163.com

Abstract—The 3 degree of freedom (3-DOF) helicopter system is a typical higher-order times, instability, multi-variable, nonlinear and strong coupling control system. This paper presents the analysis of the mathematical model and the basic principle of PID control and fuzzy PID control, which are based on 3-DOF helicopter systems. In allusion to balance control of 3-DOF helicopter system, PID controller and fuzzy PID controller are designed respectively to control the model. The MATLAB simulation results demonstrate the control effects of both controllers achieve the requirements in this system, and fuzzy self-tuning PID shows more advantages.

Index Terms—helicopter system, Fuzzy PID, matlab, pitching axis, lateral axis, revolving shafts

I. INTRODUCTION

The 3 degree of freedom (3-DOF) helicopter model is a typical multi-input and multi-output system with high-order, which contains the properties of strong channel coupling and high nonlinearity. Since the motion equations of its three axes in elevation, pitch, and travel are always used to simulate the dynamic characteristics of real helicopter, 3-DOF helicopter model becomes an efficient tool for teaching and researching in control method [1], [2].

Many efforts have been demonstrated it is difficult to illustrate the dynamic characteristic of the helicopter, because of its extremely complex and particular flying state. Normally, its dynamic characteristics will be correspondingly varied with flying altitude and flying state, all of which are nonlinear and multivariable coupling. Consequently, it is probably impossible to achieve formulations of the helicopter as well as its accurate models. It seems that the helicopter models used in engineering tend to be processed simply. In this dissertation, a 3-DOF helicopter model as the object of study, so as to researching the attitude control and simplifying the complexity of research [3].

According to the dynamic trait of pitch axis, elevation axis and travel axis, the system mathematic model was established, the model offered basic for the PID control and theoretical support for the controller improvement. Finally, the fuzzy control was used to control the system. The control results of the simulation experiments proved that the using of the fuzzy control worked better than the pure PID controller. Compared with the traditional PID control, it had a remarkable improvement. Both the PID controller and the fuzzy controller had their respective advantages. Based on their excellent merits, PID and fuzzy controller were combined together as one controller which fit together with switch and weighted value. It optimizes the flying gesture of the helicopter [4], [5].

II. SYSTEM MODELING

A 3-DOF helicopter control system introduced in this paper, is inherently unstable with MIMO, nonlinear and high degree of coupling. This system consists of some basic components, such as a pedestal, an equalizer bar, counterbalance and blades. Based on pedestal for fulcrum, the equalizer bar can be driven to pitch and turn. Counterbalance and blades are separately installed on both ends of the equalizer bar. Accordingly, the equalizer bar due to the lift force which the blades produced to pitch (based on pedestal for fulcrum). Under the similar condition, equalizer bar turns by velocity different of two blades pitching (based on pedestal for axis). In order to measure data of revolving shafts and pitch axis of equalizer bar and lateral axis of propellers, encoders are installed respectively on revolving shafts and pitching axis of equalizer bar and lateral axis of propellers. Two blades are driven by two Direct Current Brushless Motors, blades are proposed power simultaneously. Hence, models are established for three axes (degree of freedom) as the characteristics of the system [6].

A. Pitching Axis

Structure of pitching axis as Fig.1,

Manuscript received May15, 2011; revised September28, 2011; accepted January8, 2012 .

Corresponding author: Xinghu Xu

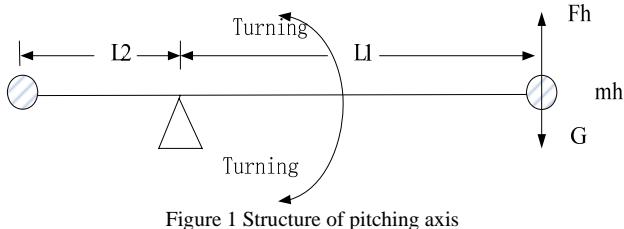


Figure 1 Structure of pitching axis

Torque of pitching axis is generated by two propeller-motors producing the direct lift F_1 and F_2 . Helicopter increases when lift F_h is greater than gravity G ; conversely it drops. Assuming that helicopter is hung in the air and the degree of pitch is zero; equations can be given as follows:

$$\begin{aligned} J_e \ddot{\varepsilon} &= l_1 F_h - l_1 G = l_1 (F_1 + F_2) - l_1 G \\ J_e \ddot{\varepsilon} &= K c l_1 (V_1 + V_2) - T_g = K_c l_1 V_s - T_g \end{aligned} \quad (1)$$

In equation (1), J_e is the rotational inertia of the pitching axis; V_1 and V_2 are voltages of two motors; K_c is the lift constant for blades motor; l_1 is the distance from fulcrum to motors; l_2 is the distance from fulcrum to counterbalance; T_g is the effective gravity torque produced by gravity of pitching axis; m_h and m_b are respectively quality of blades and counterbalance of helicopters; $\ddot{\varepsilon}$ is the rotation acceleration of pitching axis.

If gravity disturbance torque is ignored, the transfer function of this part of the system as following equation:

$$\frac{\varepsilon(s)}{V_s} = \frac{K_c l_1}{J_e s^2} \quad (2)$$

Lateral Axis

Structure of Lateral axis as Fig.2,

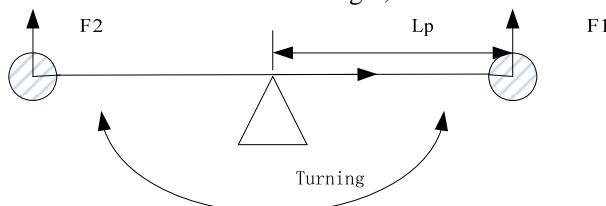


Figure 2 Structure of lateral axis

From above image, lateral axis controlled by lift produced by two blades. If the lift created by F_1 is greater than that which is created by F_2 , blades will tilt, and then there will be a lateral force to make helicopter turn around the base, i.e.

We can get equation as follows:

$$J_p \ddot{p} = F_1 l_p - F_2 l_p = K_c l_p (V_1 - V_2) = K_c l_p V_d \quad (3)$$

In equation (3), J_p is the rotational inertia of the lateral axis, the distance from fulcrum to motor, and P the rotation acceleration of the lateral axis.

The transfer function of this part of the system can be given by formula (4) as follows:

$$\frac{p(s)}{V_d} = \frac{K_c l_p}{J_p S^2} \quad (4)$$

B. Revolving Shafts

Structure of revolving shafts as Fig.3,

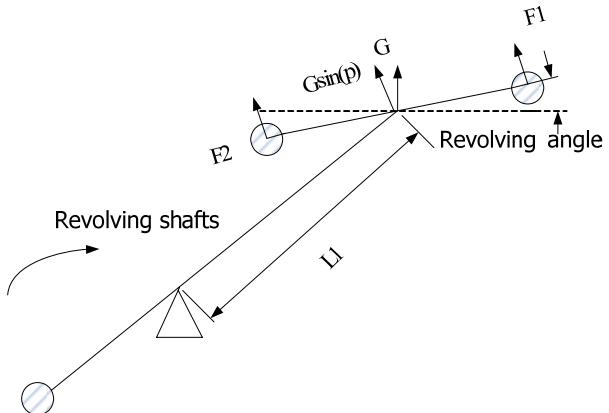


Figure 3 Structure of revolving shafts

Its dynamics equation is as follows:

$$J_r \dot{r} = -G \sin(p) l_1 \quad (5)$$

In equation (5), r is the rotating speed, rad/sec as one unit. $\sin(p)$ is the sine value of lateral angle p . Generally the angle is small, approximately considered as $\sin(p) \approx p$ here. The transfer function of this part of the system can be got by formula (6) as follows:

$$\frac{r(s)}{P(s)} = -\frac{G l_1}{J_r S} \quad (6)$$

III. CONTROLLER DESIGN

Helicopter control requires the ability to produce moments and forces on the vehicle to produce equilibrium and to change the helicopter's velocity, position and orientation [7]. The problem of helicopter control has received much attention and especially during the last two decade [8]-[10]. The usage of the traditional PI, PD controllers is not satisfied because the helicopter parameters are very dependent on the operating point. These controllers only work well in a very small area around that set point. Also, when dealing with multivariable systems, one of the major concerns are the cross-couplings of the system [11].

A. PID Controller Design

Structure of PID controller as Fig.4,

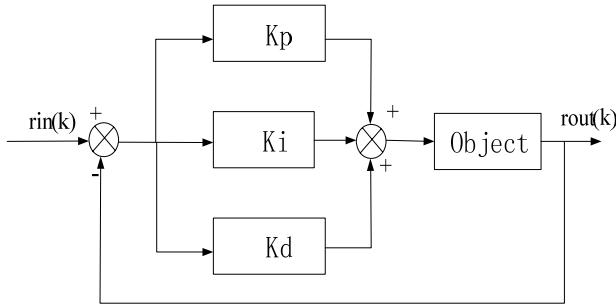


Figure 4 Structure of PID controller

PID controller is a linear controller, the control rules as follows [12]:

$$u(t) = k_p(error(t) + K_i \int_0^t error(t)dt + K_d \frac{derror(t)}{dt}) \quad (7)$$

Where, error(t) is the deviation composed by given value $rin(t)$ and actual output value $yout(t)$, K_p , K_i , K_d , respectively known as proportional coefficient, integral coefficient and differential coefficient. The effect of the three links of PID control is obvious:

- a) Proportional link: the deviation signal is reflected proportionally, and the controller would produce effect to reduce deviation if deviation produced;
- b) Integrating link: mainly used to eliminate static errors, and improve indiscrimination degree of the system;
- c) Differential modulus, to reflect the rates of deviation signal change, and to lead in an effective early modify signal in system before error signal becomes too large, to speed up the responsiveness of the system, and to reduce the setting time.

Then, three parameters operate convenient in practical control [13].

1) Controller of Pitching Axis

According to equation (7), a PD controller is designed as follows:

$$V_s = K_{ep}(\varepsilon - \varepsilon_c) + k_{ed}\dot{\varepsilon} \quad (8)$$

In equation (8), ε represents actual angle of pitch, and ε_c represents the pitch angle of expectation.

The transfer function of this part of the system can be given as follows:

$$\frac{\varepsilon(s)}{\varepsilon_c(s)} = \frac{-k_c k_{ep} l_1 / J_e}{s^2 - k_c k_{ep} l_1 / J_e s - k_c k_{ep} l_1 / J_e} \quad (9)$$

In equation (9), the denominator can be expressed as

follows:

$$s^2 + 2\xi\omega_0 + \omega_0^2 \quad (10)$$

Peak time as $t_p = \frac{\pi}{\omega_0 \sqrt{(1-\xi^2)}}$, by choosing expected peak time t_p and damping ratio, to determine the k_{ep} and k_{ed} to satisfy expectations response.

2) Controller of Lateral Axis

The rotation speed of the helicopter can be controlled by changing the size of slant angle of lateral axis. According to equation (7), we design a PD controller as follows,

$$V_d = K_{pp}(p - p_c) + k_{pd}\dot{p} \quad (11)$$

The transfer function of this part of the system can be got as follows,

$$\frac{p(s)}{p_c(s)} = \frac{-k_c k_{pp} l_p / J_e}{s^2 - k_c k_{pd} l_p / J_p s - k_c k_{pp} l_p / J_p} \quad (12)$$

Then, as the controller of pitching axis, we can determine the k_{pp} and k_{pd} by choosing expected peak time and damping ratio to satisfy expectations response.

3) Controller of revolving shafts

According to equation (7), we design a controller to get expected Lateral Angle,

$$P_c = K_{rp}(r - r_c) + k_{ri}\int(r - r_c) \quad (13)$$

The transfer function of this part of the system can be get as follows:

$$\frac{r(s)}{r_c(s)} = \frac{-k_{rp} G_l s + k_{ri} G_l / J_t}{s^2 - k_{rp} G_l / J_t s - k_{ri} G_l / J_t} \quad (14)$$

This controller can also be designed according to the second-order controller, by choosing expected peak time t_p and damping ratio, to determine the k_{rp} and k_{rd} to satisfy expectations response.

B. FUZZY-PID Controller Design

This fuzzy self-adaptive PID controller used the signals of the error e and the error varying ec as inputs, modifying PID parameters at any time through the fuzzy control rules, and then puts control signal into actuators to adjust controlled object [14]. The structure of FUZZY-PID controller is shown in Fig.5.

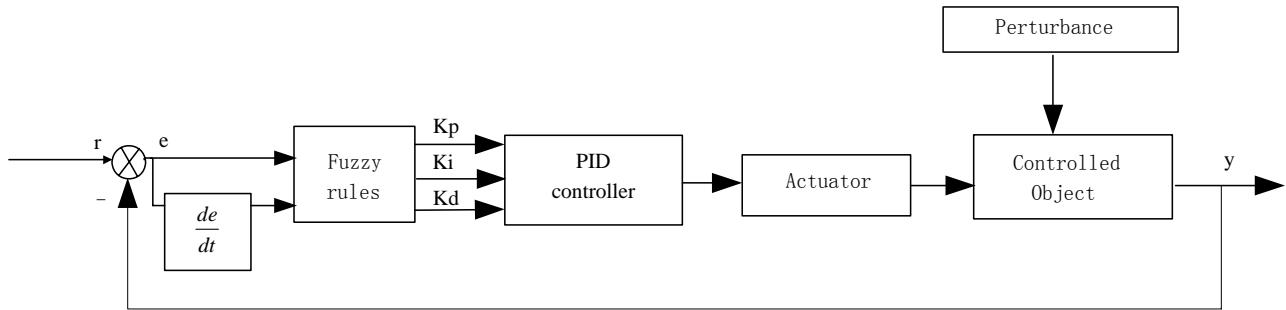


Figure 5 Structure of fuzzy self-adaptive PID controller

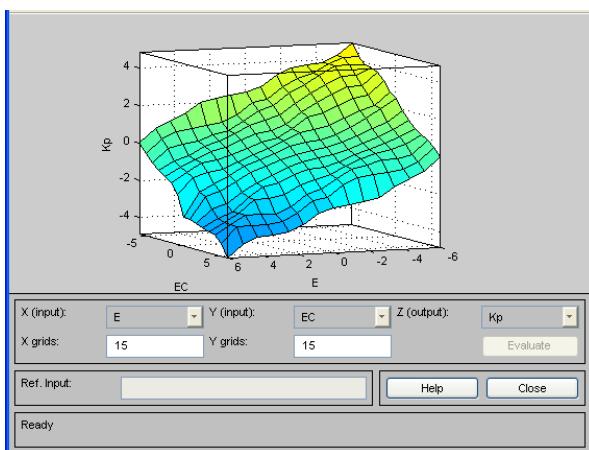
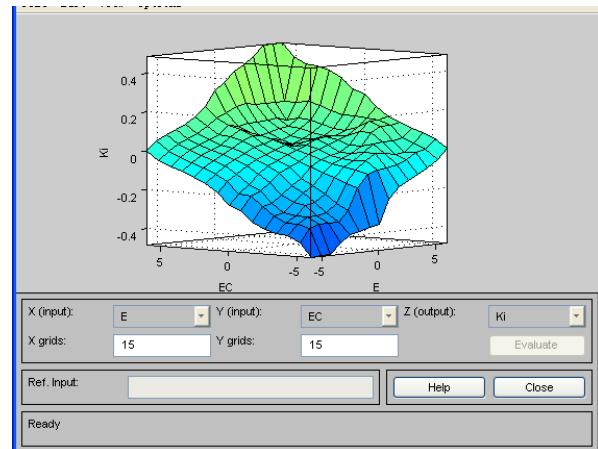
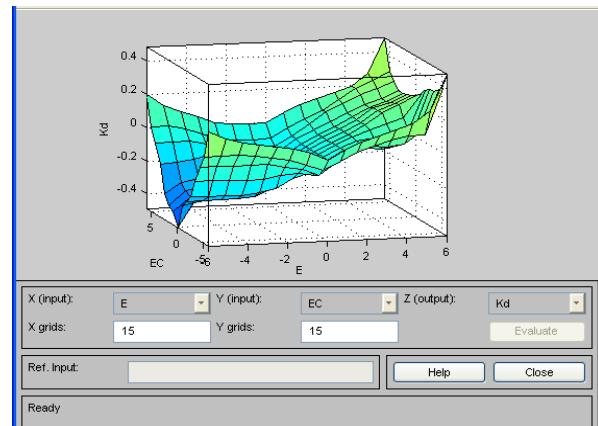
In view of the different error e and error rate, different PID control parameters of K_i , K_d , K_p are demanded, and we set rules as follows.

1) When $|e|$ is big, for the system to have good tracking performance, we should take large K_p and small K_d ; meanwhile, in order to avoid large overshoot in system response, we should set a limit to the integral effect, usually we take $K_i = 0$.

2) When $|e|$ and $|ec|$ are median size, for the system to have a small overshoot, we should take small K_p . In this case, K_d is bigger impact on the system, we should take a small value, K_i should be proper.

3) When $|e|$ is small, for the system to have a good stability, we should take big K_i and K_d ; meanwhile, in order to avoid appearing oscillation within the set value, and considering anti-interference performance of the system ,when $|ec|$ is large, K_d should be taken a smaller value, and vice versa[15].

Three-dimensional diagram of rules of K_p , K_i , K_d as Fig. 6, Fig. 7 and Fig.8.

Figure 6 Three-dimensional diagram of rules of K_p Figure 7 Three-dimensional diagram of rules of K_i Figure 8 Three-dimensional diagram of rules of K_d

The fuzzy self-adaptive PID controller in this paper, all of the linguistic values of inputs and outputs are divided into seven linguistic values: NB, NM, NS, O, PS, PM and PB. K_i , K_d , K_p parameters of the controller process on-line self-tuning, its structure consists of two parts, the routine PID controller section and fuzzy reasoning calibration parameters section. In on-line operation process, control system does online self-adjusting and fuzzy self-adjusting for K_i , K_d , K_p , through processing, look-up table and calculation about results of fuzzy logic rules. Computational formula as equation (15),

$$\begin{cases} K_p = K_p^* + \{e_i, e_{ci}\} q_p \\ K_i = K_i^* + \{e_i, e_{ci}\} q_i \\ K_d = K_d^* + \{e_i, e_{ci}\} q_d \end{cases} \quad (15)$$

In equation (15), K_p , K_i and K_d are initial-value of three control parameters of self-adaptive fuzzy PID controller, K_p, K_i, K_d are the adjusted parameters of PID, q_p, q_i and q_d are corrective factors of PID, $\{e_i, e_{ci}\}$ is error e and error rate ec corresponding to output values of the fuzzy control rules table. In process of establishing the fuzzy rule base, we should consider that input and output variables of the fuzzy controller are accurate quantity, and fuzzy reasoning is the quantity of fuzzy signal [15].

In order to adapt to more systems, membership functions of this paper selected smoothing continuous Z type membership function at negative boundary, and smoothing continuous S membership function at positive boundary, and we choose triangular type membership function which has higher delicacy in the middle part. The universe of error e , error rate ec and the output P is $\{-6 \dots 6\}$, and the universe of the output I and D is $\{-0.6 \dots 0.6\}$. Membership function diagram as Fig. 9, Fig. 10 and Fig. 11.

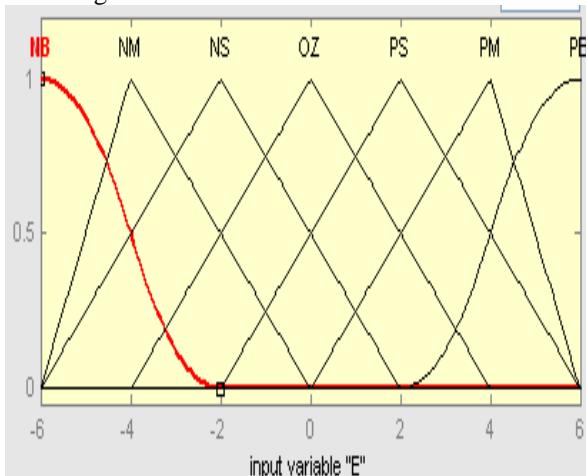


Figure 9 Membership function diagram of E

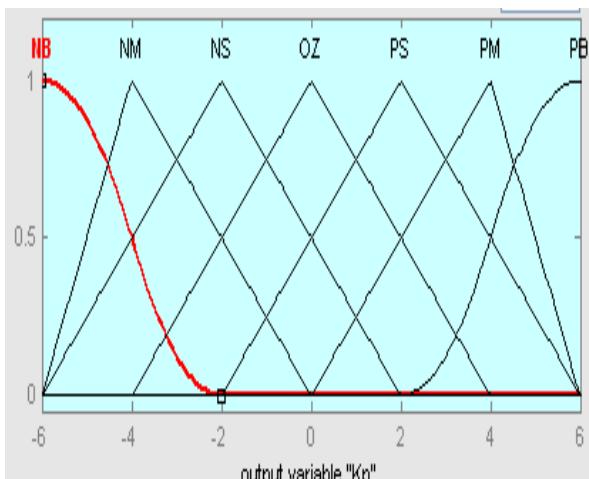


Figure 10 Membership function diagram of Kp

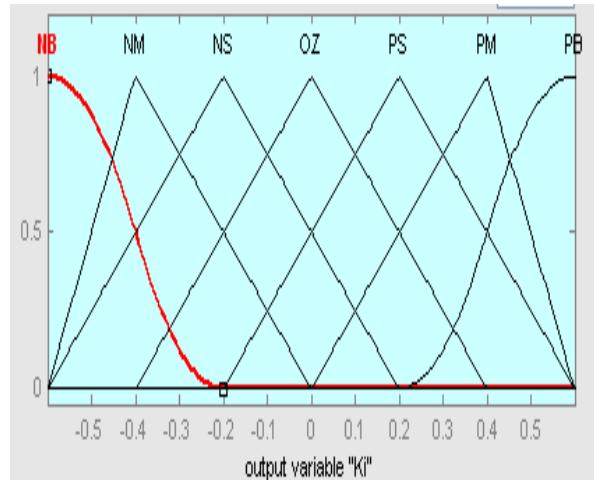


Figure 11 Membership function diagram of Ki

And in this paper , rules of Fuzzy PID controller as Table I

TABLE I

CONTROL RULES Of $\Delta K_p, \Delta K_i, \Delta K_d$

Control rules of ΔK_p :

ΔK_p	NB	NM	NS	OZ	PS	PM	PB
ec							
e							

NB	PB	PB	PM	PM	PS	ZO	ZO
NM	PB	PB	PM	PS	PS	ZO	NS
NS	PM	PM	PM	PS	ZO	NS	NS
ZO	PM	PM	PS	ZO	NS	NM	NM
PS	PS	PS	ZO	NS	NS	NM	NM
PM	PS	ZO	NS	NM	NM	NM	NB
PB	ZO	ZO	NM	NM	NM	NB	NB

Control rules of ΔK_i and ΔK_d :

NB	NB	NB	NM	NM	NS	ZO	ZO
NM	NB	NB	NM	NS	NS	ZO	ZO
NS	NM	NM	NS	NS	ZO	PS	PS
ZO	NM	NM	NS	ZO	PS	PM	PM
PS	NM	NS	ZO	PS	PS	PM	PB
PM	ZO	ZO	PS	PS	PM	PB	PB
PB	ZO	ZO	PS	PM	PM	PB	PB

NB	PS	NS	NB	NB	NB	NM	PS
NM	PS	NS	NB	NM	NM	NS	ZO
NS	ZO	NS	NM	NM	NS	NS	ZO
ZO	ZO	NS	NS	NS	NS	NS	ZO
PS	ZO						
PM	PB	NS	PS	PS	PS	PS	PB
PB	PB	PM	PM	PM	PS	PS	PB

Using fuzzy toolbox in MATLAB, Mandani reasoning methods is adopted to taking And method as

Min, and Or method as Max according based on fuzzy control rules and fuzzy membership functions. Implication method is Min, Aggregation method is Max and considering defuzzification method as centroid. According to all linguistic values in rule tables, its forms can be used (i.e. If...then) in order to complete fuzzy reasoning section, and prepare simulation in next step in simulink, And relations are used between outputs and inputs to edit rules [16].

FIS Editor in matlab of Fuzzy PID controller as Fig.12.

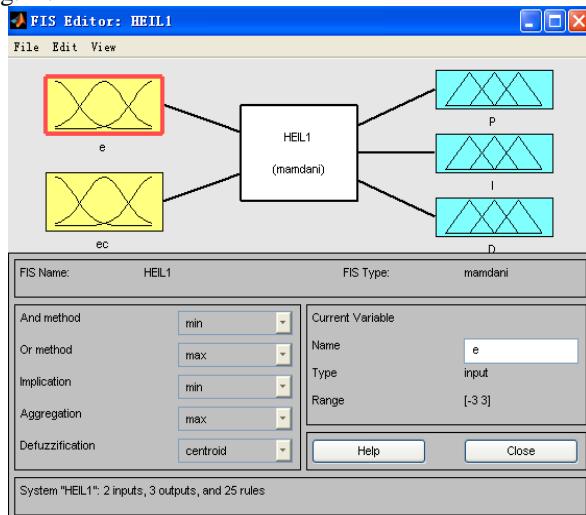


Figure 12 FIS Editor

IV. THE SIMULATION RESULTS AND ANALYSIS

According to transfer function of controlled object and relationship of every link which are got in section 2 , we build the controlled object model in matlab/simulink as Fig.13.

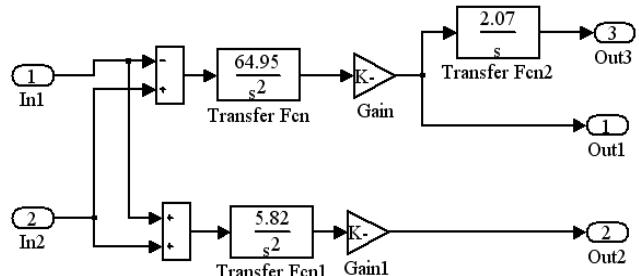


Figure 13 Structure of controlled object model

According to Section 3, we establish subsystem of PID controller, subsystem of fuzzy PID controller and full figure of system as Fig. 14, Fig. 15 and Fig.16.

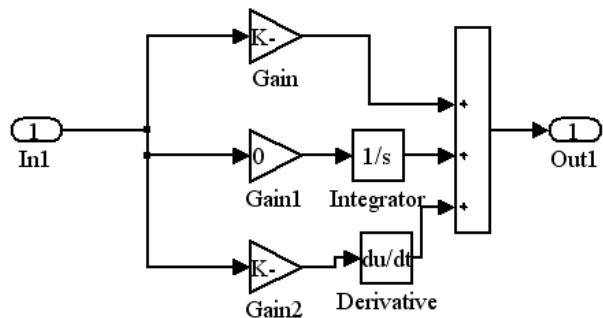


Figure 14 Structure of PID controller

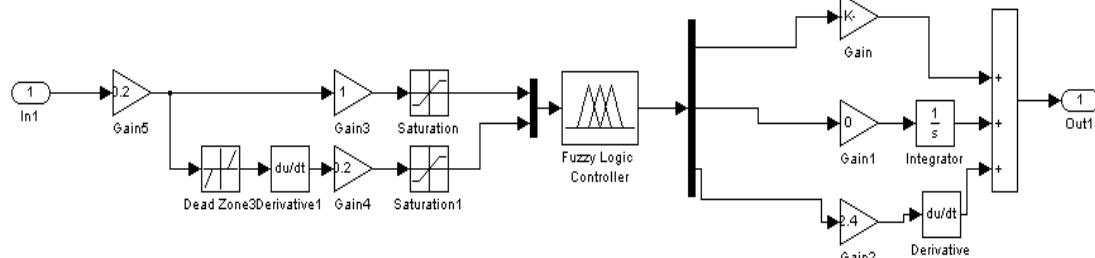


Figure 15 Structure of fuzzy PID controller subsystem

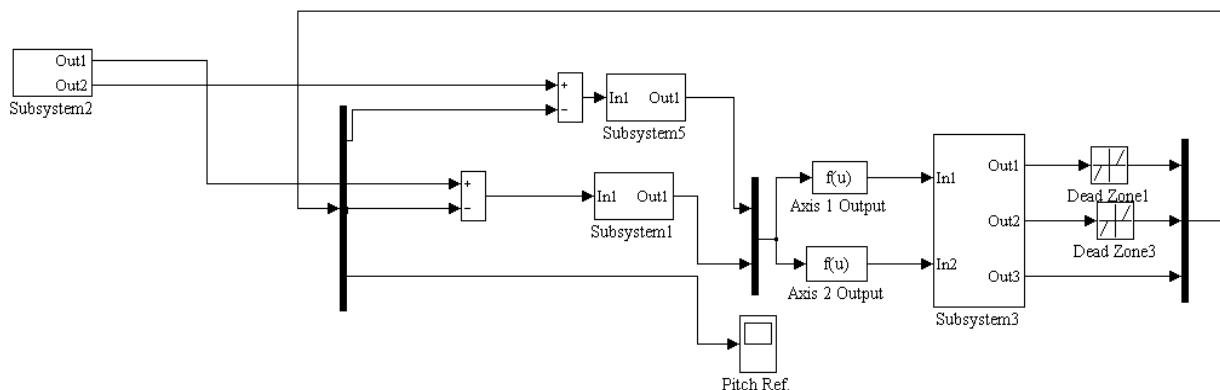


Figure 16 Full Figure of system

The set value of input module is: input 1 is set as setting value of the pitch, and input 2 is set as setting value of the revolving shafts. The expected control effect is: the helicopter raise at first falls at 10 seconds, rises again at 20 seconds, and falls back to the initial lift at 30 seconds. It starts to revolve from 40 seconds, counterclockwise rotation at first, and clockwise rotation at 50 seconds [17].

According to the intended target, in simulation process, we set input 1 initial as the unit step, and the set value reduced to 0 at 10 seconds, added to 2 at 20 seconds, and reduced to unit step at 30 seconds; we set input 2 initial as 0, and it reduced to -1 at 40 seconds, added to 0.5 at 50 seconds, then kept uniform. We add a pulse disturbance angle of pitch at the 42 seconds to simulate collision disturbance of the environment. Response curve of simulation figures without delay link shown in Fig.17 and Fig.18.

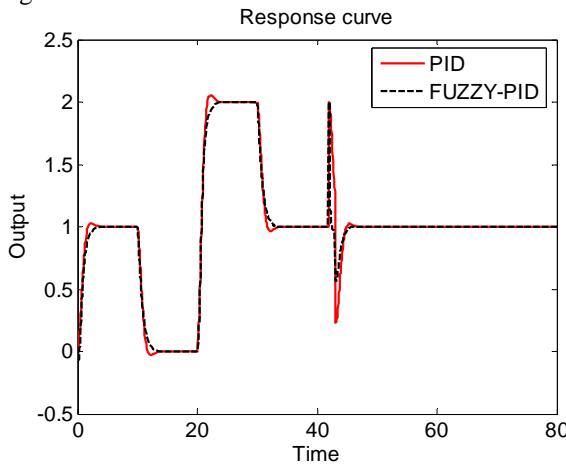


Figure 17 Response curve of angle of pitch without delay

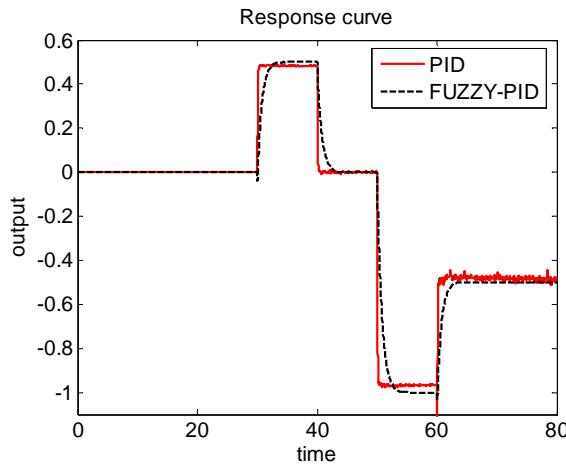


Figure 18 Response curve of lateral angle without delay

Through the simulation curve, we can get two groups of system performance data of control modes, shown in Table II .

TABLE II

SYSTEM PERFORMANCE DATA OF CONTROL MODES

	Setting time	Overshoot	Static errors of pitching angle	Static errors of lateral angle
PID	3.395	4.8 %	0	3%
Fuzzy-PI D	2.682	0	0	0

Add a delay link about 0.02s into model, then, response curve of angle of pitch with delay and response curve of lateral angle with delay shown in Fig.19 and Fig.20.

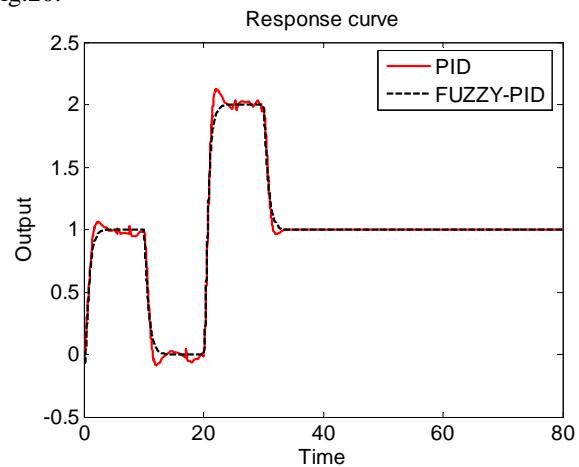


Figure 19 Response curve of angle of pitch with delay

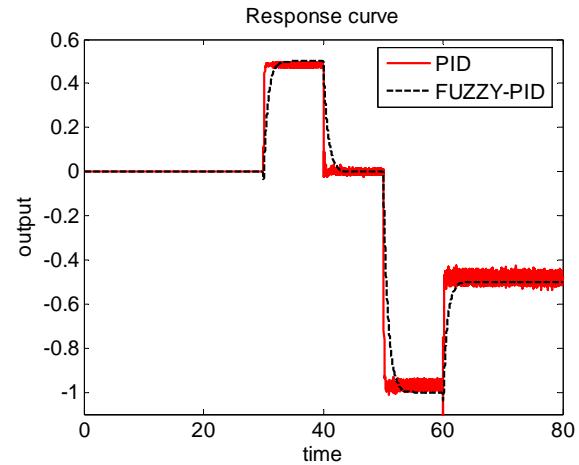


Figure 20 Response curve of lateral angle with delay

It can be seen from above tables that both of PID control method and fuzzy PID control method can achieve the purpose of remaining the system stable. However, PID control had disadvantages in the rapidity and accuracy. Also, PID control appeared overshoot, even appeared static error at response curve of lateral angle. Furthermore, system performance of PID control method was significantly worse after adding the time delay process, especially smaller

irregular shocks appeared. Correspondingly, system of fuzzy PID control method is not influenced almost.

V . CONCLUSION

The paper focus on analyzing 3-dof helicopter model, and designing the PID controller and the PID parameters self-adjusting controller, which depend on fuzzy inference functions to achieve.

Simulation results demonstrate that both of the controllers can be applied into 3-dof helicopter system to keep stable. Admittedly, the accuracy improved and the speed met the requirements of 3-dof helicopter model. In addition, the fuzzy PID control method has better adaptability for parameters change and environmental change, so it is applicable to environment with a delay and other actual existence but working trouble. With comparison, fuzzy PID controller could probably develop the characteristics of stronger robustness, better dynamic response, more quickly regulating time and smaller overshoot of fuzzy control. As a result, fuzzy PID controller performances better than PID controller dose.

REFERENCES

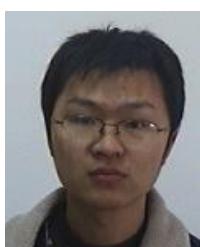
- [1] Ge Jinlai, Zhang Chenghui, and Cui Naxin. Fuzzy Self-tuning PID Controller in the 3-DOF Helicopter Experimental System. *Information and Control*. Vol.39,pp. 342-347.June, 2010
- [2] Liang Li-min, Yin Ping-lin. Optimal Design of the TRMS Based on MATLAB. *Microcomputer Development*. Vol.14, No.4. pp.25-30. Apr,2004.
- [3] Xiaoxiao Zhao. The 3 Dof Helicopter Control System Based on PID Contmller. *Journal of Shandong Electric Power College*. Vol.12 No.4, pp. 52-55.
- [4] Mostafa A. Hamood, Rini Akmeliaawati, Ari Legowo. ultiple-surface sliding mode control for 3DOF helicopter. 2011 4th International Conference on Mechatronics (ICOM). Page(s): 1-5, 2011 .
- [5] Takamuku, K.; Ishitobi, M.; Kumada, T.; Nishi, M.; Kunimatsu, S.; Redesign implementation of a nonlinear sampled-data controller for a 3-DOF model helicopter. *SICE Annual Conference (SICE)*, 2011 Proceedings of Publication Year: 2011 , Page(s): 1014 – 1018.
- [6] Wang Xiu-yan, Zhao Chang-li, and Li Zong-shuai. Study on Simulation of 3-DOF Helicopter Model Control Based on System Decomposition. Vol.27 No.6. December 2009.
- [7] R. W. Prouty, *Helicopter Performance, Stability, and Control*, Krieger Publishing Co., Inc., 1995.
- [8] A. Isidori and C. I. Byrnes, "Output regulation of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 35, pp. 131-140, 1990.
- [9] J. K. Vi, Y. Ma, and S. S. Sastry, "Nonlinear Control Of A Helicopter Based Unmanned Aerial Vehicle Model."
- [10] T. J. Koo and S. Sastry, "Output tracking control design of a helicopter model based on approximate linearization," in Proc. 37th IEEE Conference on Decision and Control, Tampa, FL, 1998, pp. 3635-3640.
- [11] Velagic, J., Osmic, N.; Identification and control of 2DOF nonlinear helicopter model using intelligent methods. *Systems Man and Cybernetics (SMC)*, 2010 IEEE International Conference on. Page(s): 2267 – 2275, 2010 .
- [12] Jinkun Liu. Advanced PID control Base on MATLAB simulation. 2nd edition. Beijing: Publishing House of Electronics Industry, 2004.
- [13] Qiu Li, Zeng Gui-e, and Zhu Xue-feng. A Comparative Study of PID Turning Methods. *Industry Control and Application*. Vol. 24, No.11. pp.27-31.
- [14] Zhou Liyin, Zhao Guoshu. Application of fuzzy-PID control algorithm in uniform velocity temperature control system of resistance furnace. V01.29 No.2. pp. 405-409. Feb.2008.
- [15] Shiyong Li. *Fuzzy control.neural control and intelligent control theory*. Heilongjiang: Harbin Institute of Technology Press,1998
- [16] Yang Li. *Computer control and simulation technology*. 2nd edition. Beijing: ChinaWaterPower Press, 2006.
- [17] Yue Xin-cheng, Yang Ying, and Geng Zhi-yong. No Steady-state Error Tracking Control of 3-DOF Experimental Helicopter System. *Journal of System Simulation*. Vol.19, NO.18. pp.4279-4283. Sep. 2007.



Junshan Gao born in 1962, he is professor of Harbin University of Science and Technology in Harbin, China. Master Tutor, Doctor of Engineering. He graduated from Harbin University of Science and Technology in Harbin, China. His research fields are intelligent control, chaos theory, Power Electronic Technology.

He completed four research tasks , such as Heilongjiang natural science fund item, Harbin Reserve leader fund item and so on. He is author of more than 20 papers published in international peer reviewed journals.

Prof. Gao is Councilor of Harbin Control Academy.



Xinghu Xu born in 1986, Bachelor of Engineering is a graduate of North China University of Technology in Beijing, China. His research fields are intelligent control and advanced detection technology. He completed some parts of research tasks guidance of tutor.



Chen He born in 1987, Bachelor of Engineering is a graduate of Harbin University of Science and Technology in Harbin, China. His research fields are intelligent control and advanced detection technology.

High Level Synthesis using Learning Automata Genetic Algorithm

Huijing Yang

School of Software, Harbin University of Science and Technology, Harbin, China

Email: yhj833@gmail.com

Chunying Wang

School of Software, Harbin University of Science and Technology, Harbin, China

Email: spring_hrbust@yahoo.com.cn

Ning Du

Department of Software, Harbin University of Scienc&Technology, Harbin, China.

Email: dn.duning@gmail.com

Abstract—High-level synthesis consists of many interdependent tasks such as scheduling, allocation and binding. All tasks in high-level synthesis are NP-complete and the design objectives are in conflict for nature, most of the already proposed approaches are not efficient in the exploration of the design space and not effective in the identification of different trade-offs. For these reasons, genetic algorithms can be considered as good candidates to tackle such difficult explorations. A new algorithm that named Learning Automata Genetic Algorithm (LAGA) is used in this paper to perform scheduling and allocation concurrently. This algorithm is based on the Genetic Algorithm, the difference is that the Learning Process is added to the Genetic Algorithm. This strategy can complete the scheduling and the allocation effectively in the high-level synthesis under certain time and resource constraints. This algorithm is implemented in C language and is tested finally on a number of DSP benchmarks, and the test results then are compared with those obtained from four other different techniques which are commonly used in high-level synthesis. The experimental results show that the high-level synthesis using the LAGA algorithm is very effective, especially under the area constraint.

Index Terms—Genetic algorithms, Learning Automata, High-level synthesis (HLS), Scheduling, design space exploration

I. INTRODUCTION

There is a growing consensus among VLSI designers that one of the most effective methods to handle the complexity of today's system-on-chip (SoC) designs is to use computer-aided design (CAD) techniques. CAD techniques start with an abstract behavioral or algorithmic description of a circuit and automatically synthesize a structural description of a digital circuit that realizes the behavior. The behavioral description consists of computational operations (additions, multiplications, comparisons, logical operations) and control operations (conditional statements, loops, and procedure calls) [1].

The structural description maps the operations and data transfers onto functional units in a data path and a control unit that coordinates the flow of data between various functional units of the data path. The data path include hardware units (ALUs, multipliers, logical gates), storage units (registers, registers files, RAM, ROM), and interconnect units (multiplexers, buses) that are connected together to realize the specified behavior. This structural description is called a register transfer (RT)-level description. Once an RT-level design of a circuit is obtained, it can be transformed into a logic gate level netlist through logic synthesis, then into a layout via layout synthesis, and finally fabricated into an integrated circuit. Fig. 1 illustrates a typical high-level synthesis flow used for creating a chip design, starting from an abstract algorithmic specification [2].

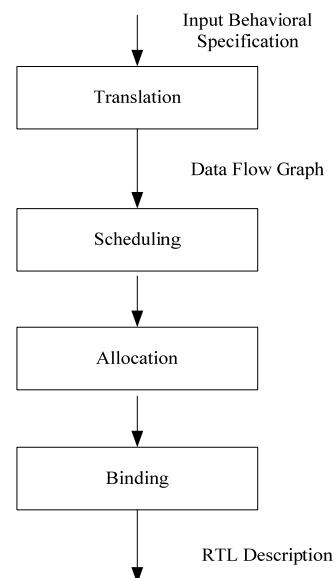


Figure 1. Interdependence of subtasks in high level synthesis

High-level synthesis (HLS) is the process of translating a behavioral description into a hardware implementation at register transfer level [3]. The design specification is usually written as a behavioral description, in a language such as C. The behavioral description is first compiled into an internal representation (such as data flow graphs - DFGs), which are then mapped to the functional units that are selected from the resource library to meet design goals (such as power, area, and performance). This process of transforming a behavioral description into a synthesizable structural description affords a methodology of automatically synthesizing a realizable digital circuit from an abstract algorithmic specification of the design, thus considerably reducing the design cycle time. VLSI designs are multiobjective by nature, since they have to trade-off several conflicting design objectives such as chip area, circuit delays, and power dissipation. The shorter design time using behavioral synthesis allows one to examine many alternative circuit realizations during the design process [4]. Often, the structural specification is divided into a data path comprised of the functional and storage units and a control unit that coordinates the flow of data between the data path elements [5]. Due to the division, high-level synthesis is traditionally divided into data path synthesis and controller synthesis[6]. The primarily focus will be on data path synthesis.

Datapath synthesis can be modeled as the process of searching a complex multidimensional space represented by the set of possible schedules, allocations, and bindings that can realize a given behavioral specification.

As modern VLSI and SoC designs become more complex, a major problem is the extremely large number of possible schedule and allocation combinations that must be examined in order to select a design that meets constraints and is optimal [7]. This process, called design space exploration, is further compounded by the need for shortening design times due to time-to-market pressures. Since an exhaustive search could be prohibitive and an ad hoc design exploration could be inefficient, designers often select a conservative architecture after some experimentation, which often results in a suboptimal design. Given this scenario, there is an acute need for techniques that automate the efficient exploration the large space in a reasonable time, during high-level synthesis of datapaths [8].

Searching a complex space of problem solutions often involves a tradeoff between two apparently conflicting objectives: exploiting the best solutions currently available and robustly exploring the design space. Genetic algorithms (GAS) manage this tradeoff in an intelligent way. GAS have recently been applied successfully to optimization problems in diverse fields, such as standard cell placement [9], searching and machine learning [10] and data path synthesis [11].

The Genetic Algorithm begins with a randomly selected population, and through recurrence of the production of the generation, looks for the best chromosome [12]. The aim of the Genetic Algorithm is to find the best chromosome. The position of genes in each

chromosome, in the Genetic Algorithm is random. If we should select the appropriate position of genes, it would be possible to appropriate the nearly optimal answer in fewer generations [13]. The Genetic Algorithm, in fact chooses the best chromosome from among the existing ones, and the positions of the genes of chromosomes are totally random [14]. If it were possible to find the optimal place of the genes of chromosomes, we would be able to find the ideal answer in fewer generations. Through utilizing the advantage of both methods, the proposed algorithm tries to achieve the optimal answer in fewer generations.

In the LAGA algorithm each chromosome is equal to an automaton and each gene equal to an action of an automaton.

In this paper, we use LAGA performing subtasks of scheduling in high-level synthesis, and to trade-off conflicting design objectives the process of scheduling based on DFG with weights. And we set weights for DFG according to the constraints of resource.

The paper is organized as follows. Section II provides a brief review of the related work, with particular attention to the evolutionary approaches. Then, Section III describe in details methodology of LAGA, while Section IV presents the results of the experimental. Section V we summarize the paper and draw conclusions based on our experimental results.

II. RELATED WORK

A. High Level Synthesis Methods

A large number of scheduling and allocation techniques have been developed for HLS over the past two decades. It is well-known that there is a strong interdependence between the HLS subtasks, and there is no clear consensus on their order of execution [18]. Such decision often has a large impact on the quality of solutions found and most of the early HLS systems performed those two subtasks separately, obtaining poor results. In literature, the high level synthesis techniques can be classified into four categories: constructive approaches, iterative transformational approaches, exact approaches and non-deterministic approaches. The constructive approaches operate on one operation or resource at a time until all elements are considered. Important algorithms following this approach, for example for scheduling, include common as-soon-as-possible (ASAP) and as-late-as-possible (ALAP) scheduling, list based scheduling [15], force-directed scheduling [16] and path-based scheduling [17]. The iterative transformational approaches perform continuous refinements to the set of solutions while exact approaches [18] exploit a mathematical formulation of the problem to find the optimal solution, but the execution time of these algorithms grows exponentially with the number of variables and the number of inequalities. Therefore, these methods are impractical for large designs. Several high-level synthesis systems use nondeterministic approaches, and in particular GAs, to perform some or all of the synthesis subtasks. Most of them consider two phases and

problems separately like in [19] where GAs are used to schedule the operation while in [20] they are used to allocate and bind a scheduled graph. In the last years, the design of algorithms for DSE is becoming crucial to consider the effects of all the HLS subtasks.

B. Genetic Algorithm

Among the optimization methods inspired by the living nature, genetic algorithm, which is based on the principles of natural evolution, is considered one the best and most sophisticated [21]. Genetic Algorithm is a non-classic and random search optimization method that deals with the function itself, not its derivations, and is based on the theory of the survival of the fittest, inspire by Darwin's evolution theory, and natural genetics [22]. In this method, search begins from several points in solution space simultaneously and through point to point search. The variables of target function are evaluated and, finally, the point which the most or the least absolute is introduced as the optimal point [23]. Optimization is the most important and function of the Genetic Algorithm. In common optimization methods target function must necessarily be coherent and consistent [24]. In Genetic Algorithm, however, a consistent and devisable function is not needed. In accordance with Genetic Algorithm, a sample from among all decision variables, that affects the function, is regarded as a member, and a certain number of these samples, makes up a set of members[25]. In this method, a set of the population of variables is used in the process of search. As a result, as the chance of creating better variables is boosted, the possibility of finding the absolute or general optimal point is heightened. This quality is specifically suitable for functions sudden changes and possessing several situational optimal points. Complete information concerning Genetic Algorithm is brought in [26].

C. Learning Automata

A learning automaton is an abstract model that randomly selects an action from a set of the finite actions and applies it to the environment. The environment evaluates the selected action and informs the result of its evaluation, by a boosted signal, to the learning automata. By using the selected action and boosted signal, the learning automata results its internal situation and then selects its next action.

We can present the environment by $E=\{\alpha, \beta, c\}$ in which $\alpha=\{\alpha_1, \alpha_2, \dots, \alpha_r\}$ is the set of inputs, $\beta=\{\beta_1, \beta_2, \dots, \beta_r\}$ is the set of outputs and $c=\{c_1, c_2, \dots, c_r\}$ is the set of penalty possibilities. When β is a two-member set, the environment is P type. In such environment, $\beta_1=1$ is considered penalty and $\beta_2=0$ reward. In a type Q environment, β set of processes an infinite number of members. c_i is the possibility of a action's being penalized. Learning automata are directed into two groups: those with fixed structures, those with variable structures [27].

III. SCHEDULE USING LAGA

LAGA algorithm is constructed of two phase. In first phase with use of Genetic Algorithm, the result is optimized and in second phases the obtained results from Genetic algorithm improved using learning automata. In first phase the genetic Algorithm is endeavor to optimize the chromosomes and then the obtained chromosomes are putting into learning automata. Then Learning Automata is focus on Chromosome Genes and finding the most suitable place of Genes in Chromosomes. Figure 2 shows the flowchart of proposed algorithm.

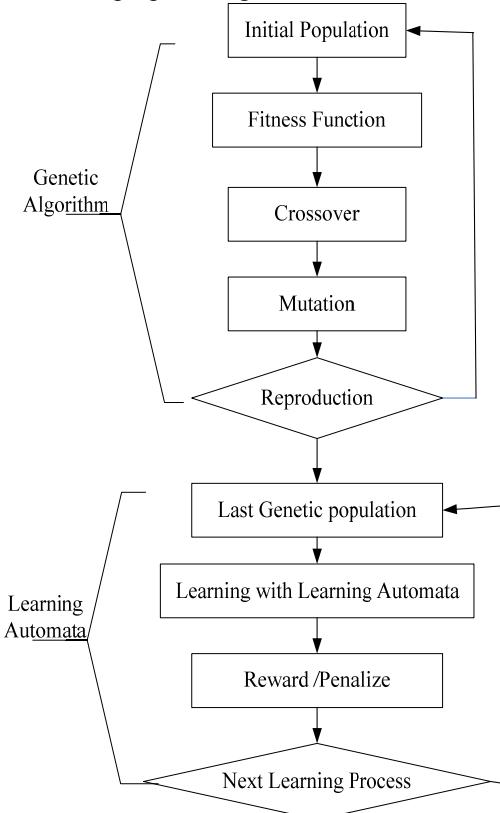


Figure 2. The Steps of proposed algorithm is described in below

A. Initial Population

At first, P (number of population) random generated chromosome and then all tasks are allocated to genes of chromosomes. Then a number is allocated randomly to all genes. The random allocated number to genes includes 2 concepts:

1. Task priority.
2. Processor's number, which executes the task.

After allocating random numbers to chromosome genes, the values of the genes are interchanged with the number of $2/N$ load.

We are going to perform, the shown graph in Figure 3, referring to Figure 4, you can observe how tasks are allocated to interior status in order. Since there are two processors in the system, so odd numbers indicate P1 processor and even numbers indicate P2 processor. If the system includes more than two processors, processor's number will consist of the result of the allocated number to the number of all processors.

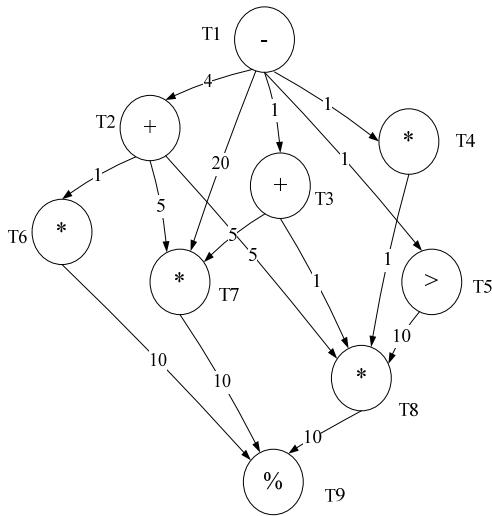


Figure 3. Operator task graph

B. Task Execution on Processors

When implementing a program on parallel processors, the data dependence between tasks should also be taken into consideration. In fact, a task cannot be implemented unless all of its parent tasks are implemented. In this section, how tasks are implemented on processors has been described [28]. Each task will be implemented in its relevant processor with regard to the automata of figure 4. From among all ready tasks, one, priority is higher than other tasks is implemented. In case, priorities of two tasks are the same as such other, one is randomly chosen for implementation. Ready task is one whose all parent tasks have been implemented.

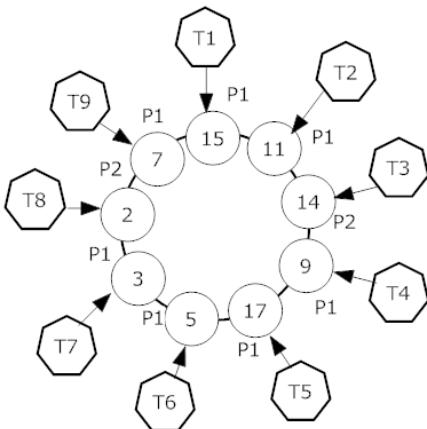


Figure 4. Example of automata for Figure 3 operator task graph

For example, by considering the automata of figure 4 and the task graph of figure 3 in the first phase, it becomes evident that only T1 task is ready. In the first phase, then, T1 will be executing. After the execution of T1 task, T2, T3, T4, T5 tasks will be on the ready. In this phase, priority of T5 task is higher than other ready tasks, so it is executed. In the same fashion, all tasks are implemented on their own specific processors [29].

With regard to Figure 4 automata, tasks T1, T5, T2, T4, T6, T7, T9 will be run on P1 processor and T3, T8 on P2

processor as well. Tasks execution order will be as TABLE 1.

TABLE I.
TASK EXECUTION ORDER

J	S'	v(j)	J*	TS
0	{1}	v(1)-15	1	S-{1}
2	{2,3,4,5}	v(2)-11,v(3)-14,v(9)-9,v(5)-17	2	S-{1,5}
1	{2,3,4}	v(2)-11,v(3)-14,v(9)-9	3	S-{1,5,3}
3	{2,4}	v(2)-11, v(9)-9	4	S-{1,5,3,2}
5	{4,6,7}	v(4)-9,v(6)-5,v(7)-3	5	S-{1,5,3,2,4}
4	{6,7,8}	v(6)-5,v(7)-3,v(8)-2	6	S-{1,5,3,2,4,6}
7	{7,8}	v(7)-3,v(8)-2	7	S-{1,5,3,2,4,6,7}
9	{8}	v(8)-2	8	S-{1,5,3,2,4,6,7,8}
6	{9}	v(9)-7	9	S-{1,5,3,2,4,6,7,8,9}

All tasks, now, will be executed according to table2. It shows task execution order on processors in details.

TABLE II.
DISPLAY OF TASK EXECUTION ORDER IN DETAILS

J*	S	pi	tj=ej+pj
1	P1-{1},P2-{}	1	T1 - 0+2-2
2	P1-{1,5},P2-{} P1-{1,5},P2-{3}	1 2	T5 - 2+5-7 T3 - 3+3-6
3	P1-{1,5,2},P2-{3}	1	T2 - 7+3-10
4	P1-{1,5,2,4},P2-{3}	1	T4 - 10+4-14
5	P1-{1,5,2,4,6},P2-{3}	1	T6 - 14+4-18
7	P1-{1,5,2,4,6,7},P2-{3}	1	T7 - 18+4-22
8	P1-{1,5,2,4,6,7},P2-{3,8}	2	T8 - 17+4-21
9	P1-{1,5,2,4,6,7,9},P2-{3,8}	1	T9 - 31+1-32

C. Fitness Function

In Genetic Algorithm, fitness function determines whether chromosomes are going to stay alive. In the problem of task scheduling, the object is to find a short makes pan. Analysis function for scheduling problem is:

$$\text{eval}(v_k) = 1/f^k, \quad k=12, \dots, \text{pop size}$$

f^k : the makes pan resulting from k^{th} chromosome.

D. Crossover Operator

Crossover is a technique which produces off-springs when two parents mate together. The parents are selected by binary tournament selection method [30]. In this paper, a novel method for combining chromosomes has been put forward. The combination method used in this paper is a two-point one. First two points are randomly chosen as subclasses, and then their contents and orders are analyzed. For instance, the subclass chosen from 1 v , has a weight order of 1-2-3-4. This weight order is used for changing the subclass chosen by v2 . Thus, the 6-13-15-11 is changed to 15-13-11-6 and changes with the weight order of v1 subclass.

WMX algorithm is not one, which changes only the contents of two points selected from two chromosomes, but it also changes the contents of classes according to weight priorities. WMA is comprised of three steps.

Step1: random substring selection for two chromosomes. In figure 5, an example of the step1 of

combining chromosome by using WMA algorithm is displayed.

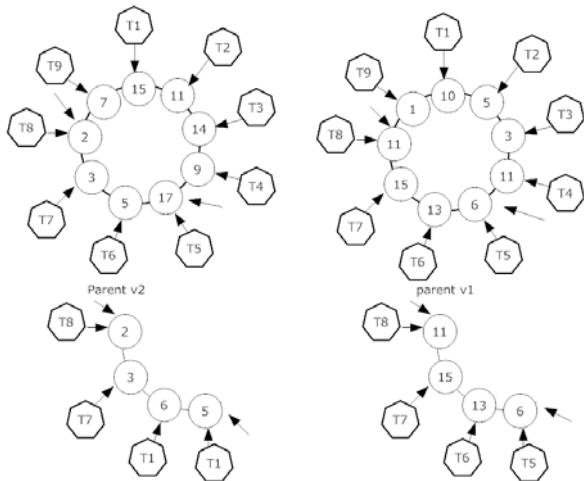


Figure 5. Step1 of WMA algorithm

Step2: defined genes mapping relation. Such as TABLE III.

Genes Mapping Relation

4	2	1	3	1	2	3	4	17	5	3	2
2	5	17	3	17	5	3	2	6	13	15	11
1	2	3	4	4	2	1	3	6	13	15	11
15	13	11	6	6	13	15	11				

Figure 6. Step2 of WMA algorithm

Step3: creating two new offspring generation, the result of above example indicated in Figure6.

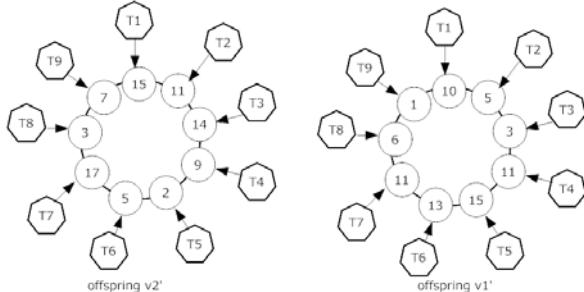


Figure 7. Step3 of WMA algorithm

E. Mutation Operator

For operating mutation, two Genes are randomly selected from a chromosome and their amounts are changed with each other. The manner of swapping actions values of a chromosome is departed into three step.

Step1: automata action status after allocating values randomly.

Step2: Selection of two random actions.

Step3: The output for automata after swapping actions randomly.

F. Selection Operator

Selection operator in this paper is as follows: In each step of new population production, $(1-p)$ % of

chromosome, which has least amount of FT, are selected and enter the new population directly. The rest of the population is, than produced through combining chromosomes.

G. Reward and Penalize Operators

Since, each chromosome is presented as a learning automaton, in each automaton, after considering the fitness of a gene (either processor or action), which is selected on a random basis, that gene, is duly penalized or rewarded. As a result of rewarding or penalizing a gene, its position in the boundary position of an action, its punishment leads to a change in its action and, in consequence, creation of a new makes pan. Departing on the type of learning automata, reward and penalize operator will be different.

Reward action occurs when the fitness of a task is smaller than threshold.

Fitness of t_i is: x/y

x : is the sum of connection cost of all parent and offspring nodes of t_j node so that. $[\sum c(t_i, t_j)]$ if($p_{ti} \neq p_{tj}$)
 p_{ti} : A processor that i t task is performed on it. p_{tj} : A processor that j t task is performed on it.

y : is the sum of costs of all parent and offspring nodes of t_i node. $\Sigma c(t_i, t_j)$

$c(t_i, t_j)$: Communication cost between t_j and t_i tasks.

Threshold rate is equal $T/Ntaks$

T : Consist of a number of related tasks to t_i task that is executed on a processor which t_i task is run in it.

$Ntaks$: The number of all graph tasks.

The more fitness level of t_j task tends to zero of the connection cost between processors tends towards zero too. If, therefore, the fitness level of a t_i task is equal to zero, It turns out that all related tasks of t_i are performed on the same processors. T has a direct relation with x ; as T increases x decreases and vice versa.

In case the fitness level of a task is lower or equal to the threshold amount, then the head of the task gets penalized. Two positions are possible when penalizing a head:

1. The head might be in a position other than frontier position. In the case, penalizing makes it less important. How the head of task T_7 is penalizes, is shown in Figure 8.

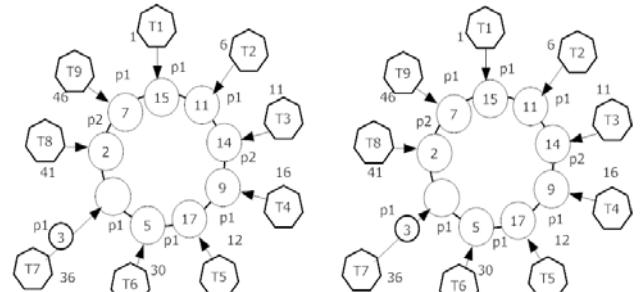


Figure 8. T7 task penalizing.

2. The head might be in frontier position. In that case, we look for a head in the graph that has the greatest reduction in the amount of FT when processors (the numbers attributed to heads) are changed. Now if the

found head is in the frontier position, the positions of the two heads are changed with each other and if otherwise, i.e. if the found head is not in the frontier position,

First the found head should be moved to its frontier position and then change occurs. Figure 9 shows how T8 task is penalized

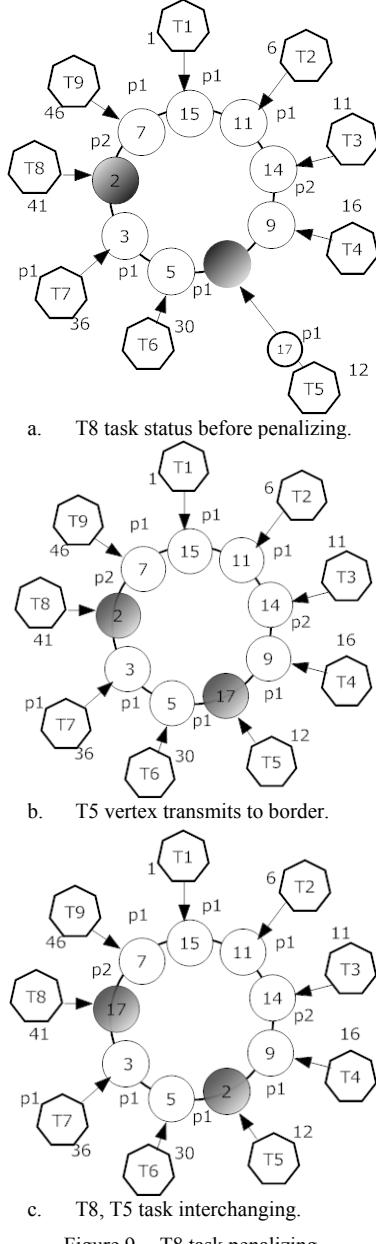


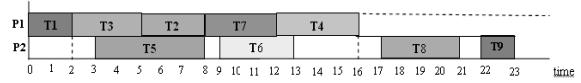
Figure 9. T8 task penalizing.

In this paper, the performance of the proposed algorithm is compared with well-known definite and indefinite algorithms. At first the proposed algorithm is simulated and evaluated on homogeneous platforms and then evaluated on heterogeneous platforms. Parameters that are used in the hybrid algorithm are shown in TABLE III.

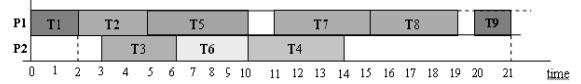
TABLE III.
HYBRID ALGORITHM PARAMETERS

Algorithm	Memory Depth	Mutation Rate	Crossover Rate	Iteration	Population
GA	-	0.2	0.7	40	50
LAGA	4	0.2	0.7	20	50

First by observing the task graph in figure 6, results obtained from various algorithms and the proposed algorithm is displayed in figure 10.



a. Acquired Gantt chart of GA Algorithm



b. Acquired Gantt chart of LGA Algorithm

Figure 10. Comparison of proposed algorithm with other algorithms (for indicated graph in Figure 3).

IV. EXPERIMENTAL RESULTS

The proposed LAGA-based high-level synthesis system has been implemented in the C language. The extended high level synthesis tool accepts ANSI C programs and generates RTL specifications in verilog.

To perform a qualitative assessment of algorithm of LAGA in the high-level synthesis, it was tested on a number of DSP benchmarks drawn from high-level synthesis literature.

For all the benchmarks tested, the synthesized designs were assumed to operate with a clock period of 20 ns. We used a 0.35- CMOS module library, where ALUs, multipliers, registers, and multiplexers are implemented as hard macro cells (cells having fixed aspect ratio and pin locations). The ALUs have a propagations delay of 6.5 ns, and multipliers have a propagations delay of 15 ns. We assume that the area cost and delay of a pipelined multiplier are the same as those of a nonpipelined multiplier, respectively.

In all the experiments, the size of the GA population was set to 100, the crossover probability was 0.90, and the mutation probability set to 0.20. Each of the GA runs was stopped after 10000 fitness evaluations. Since GA algorithms are stochastic algorithms, ten independent runs with different random number seeds were performed for each of the benchmark problem instances, and the best solution found by the GA in each of the ten runs was recorded.

We compared our results with those obtained from four different scheduling techniques commonly used in high-level synthesis, namely, the GA scheduling , ALAP scheduling , force-directed (FDS) scheduling , and simultaneous scheduling, allocation, and binding (SAM) technique . These scheduling algorithms were tested in a traditional high-level synthesis framework that performs the three synthesis subtasks of scheduling, allocation, and binding independently. The goal of this comparison was twofold: 1) to verify the performance gains from concurrently performing scheduling and allocation, over a traditional synthesis flow that carries out these subtasks independently and 2) to use the performance of these scheduling techniques as a baseline to compare our results. The same benchmarks are used for comparing the results.

For each of the benchmark problems, a design-space exploration was performed by setting different values to the weights (and) corresponding area constraints. Results are shown in TABLE IV and TABLE V.

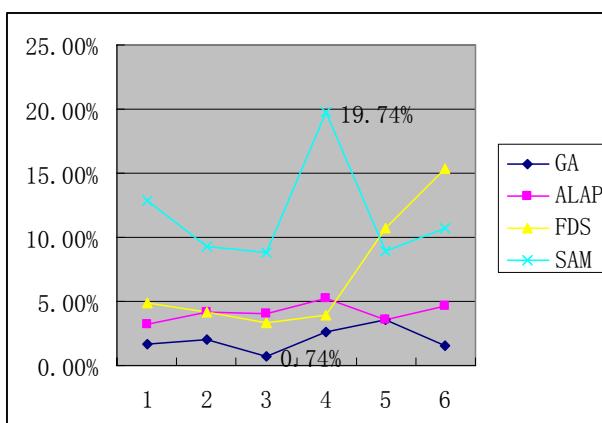
In TABLE IV, the chip latency in bold indicate the Fastest designs for each areas value. From the table, it can be seen that LAGA algorithm finds better solutions than those of the other four scheduling techniques, for all the benchmarks tested.

TABLE IV.
COMPARISON OF OUR LAGA-BASED METHOD WITH OTHER
SCHEDULING ALGORITHMS

Benchmark Example	Area Constraint	Chip Latency(ns)				
		LAGA	GA	ALAP	FDS	SAM
IIR	6.0mm ²	62	63	65	66	70
FIR	5.0 mm ²	97	99	101	101	106
EWF	4.5 mm ²	272	274	283	281	296
ARF	10.0 mm ²	76	78	80	79	91
DCT	12 mm ²	56	58	58	62	61
FDCT	18 mm ²	65	66	68	75	72

TABLE V shows the improvement of the LAGA-based solutions over those of the other four scheduling techniques. The average improvements range from 0.74% (compared to the GA method) to 19.74% (compared to the FDS method).

TABLE V.
INCREASE PROPORTION OF THE LAGA-BASED SOLUTIONS
OVER OTHERS



The experimental results show that the LAGA algorithm to complete scheduling and allocation concurrently in high level synthesis is very effective, especially under area constraint. It can obtain a better delay performance than other algorithms.

V. CONCLUSION

In this paper, a new method LAGA is used in high level synthesis to deal with scheduling and allocation simultaneously. It can produce area and performance optimized designs. This algorithm utilizes Genetic Algorithm and Learning Automata methods sequentially to search for the mode space. It can find the Solutions quickly by using Genetic Algorithm and Learning Automata sequentially in search process.

The method is simulated on a number of DSP benchmarks. It can succeed in obtaining optimal solutions. The same problems have been also solved in a general way. The experimental results indicate that LAGA algorithm is very effective in high-level synthesis, especially under the area constraint.

REFERENCES

- [1] I. Das. A preference ordering among various Pareto optimal alternatives. Structural and Multidisciplinary Optimization, 18(1):30–35, Aug. 1999.
- [2] M. Palesi and T. Givargis, “Multi-objective design space exploration using genetic algorithms,” in CODES ’02: Proceedings of the tenth international symposium on Hardware/software codesign. Estes Park, Colorado: ACM, pp. 67–72. 2002,
- [3] M. Laumanns, L. Thiele, and E. Zitzler, “An efficient, adaptive parameter variation scheme for metaheuristics based on the epsilon-constraint method,” European Journal of Operational Research, vol. 169, no. 3, pp. 932–942, Mar. 2006.
- [4] R. J. Cloutier and D. E. Thomas, “The combination of scheduling, allocation and mapping in a single algorithm,” in Proc. 27th Design Automation Conf., pp. 71–76, Jun. 1990.
- [5] X. Yao and T. Higuchi, “Promises and challenges of evolvable hardware,” IEEE Trans. Syst., Man, Cybern., pt. C, vol. 29, no. 1, pp. 87–97, Feb. 1999.
- [6] S.M. Logesh, D.S. Harish Ram, M.C. Bhuvaneswari, Multi-objective Optimization of Power, Area and Delay during High-Level Synthesis of DFG’s - A Genetic Algorithm Approach, Electronics Computer Technology (ICECT), 2011 3rd International Conference on, pages 108 – 112, 8-10 April 2011.
- [7] J. C. Gallagher, S. Vigraham, and G. Kramer, “A family of compact genetic algorithms for intrinsic evolvable hardware,” IEEE Trans. Evol. Comput., vol. 8, no. 2, pp. 1–126, Apr. 2004.
- [8] G. Ascia, V. Catania, and M. Palesi, “A GA-based design space exploration framework for parameterized system-on-a-chip platform,” IEEE Trans. Evol. Comput., vol. 8, no. 4, pp. 329–346, Aug. 2004.
- [9] Fujimoto, N.; Hagihara, K.; A comparison among grid scheduling algorithms for independent coarse-grained tasks. International Symposium on Applications and the Internet Workshops, SAINT 2004, Pages: 674 - 680, 2004.
- [10] Zhang, Yaping; Peng, Haoyu; Wang, Zonghui; Shi, Jiaoying; Research and Implementation of Distributed Simulation and Parallel Rendering System Based on Grid. 2010 First International Conference on Networking and Distributed Computing (ICNDC), Pages: 3 - 7, 2010.
- [11] Yanfang Fu; Yan Fan; Scheduling Method of Resource for Simulation System Based on Grid. Second International Conference on Computer Modeling and Simulation, ICCMS '10, Pages: 226 - 229, 2010.
- [12] Liu, Hanbing; Su, Hongyi; Zhan, Shouyi; Chai, Xundong; Zhang, Yabin; Hou, Baocun; Guo, Linqin; Fan, Shuai. Resource scheduling based on dynamic dependence injection in virtualization-based simulation grid. 14th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Pages: 396 - 401 , 2010.
- [13] Huang Hai, Tian Lei, Wu Wei, Sun Songlin, Jing Xiaojun. Genetic algorithm for scheduling of interactive tasks in simulation grid. Proceedings of 2010 WASE International

- Conference on Information Engineering, ICIE 2010, v 1, Pages 30-33, 2010.
- [14] Wei Hongtao. Task Management and Scheduling Methods for Grid-Computing-based Simulation, Doctor's thesis, National University of Defense Technology, 2005.
- [15] B. M. Pangle and D. D. Gajski, "Slicer: A state synthesizer for intelligent silicon compilation," in Proc. Int. Conf. Computer-Aided Des., pp. 536–541, 1987.
- [16] P. G. Paulin and J. P. Knight, "Force-directed scheduling for the behavioral synthesis of ASICs," IEEE Trans. Comput.-Aided Des., vol. 8, no. 6, pp. 661–679, 1989.
- [17] A. C. Parker, J. T. Pizarro, and M. Mlinar, "Maha: A program for datapath synthesis," in Proc. 23rd ACM/IEEE Design Automation Conf., pp. 461–466., 1986,
- [18] R. Camposano, "Path-based scheduling for synthesis," IEEE Trans. Comput.-Aided Des., vol. 10, pp. 85–93, 1991.
- [19] R. K. Brayton, R. Camposano, G. De Micheli, R. Otten, and J. van Eijndhoven, "The Yorktown silicon compiler system," in Silicon Compilation, D. Gajski, Ed. Reading, MA: Addison-Wesley, pp. 204–310, 1988.
- [20] S. G. Ara ' ujo, A. C. Mesquita, and A. Pedroza. Optimized Datapath Design by Evolutionary Computation. In Proceedings of the 3rd IEEE International Workshop on System-on-Chip for Real-Time Applications (IWSOC'03), 30 June - 2 July 2003, Calgary, Alberta, Canada, pages6–9, 2003.
- [21] F. Ferrandi, P. L. Lanzi, G. Palermo, C. Pilato, D. Sciuto, and A. Tumeo. An evolutionary approach to area-time optimization of FPGA designs. Embedded Computer Systems: Architectures, Modeling and Simulation, 2007. IC-SAMOS 2007. International Conference on, pages 145–152, 16-19 July 2007.
- [22] S. Gupta, N. Dutt, R. Gupta, and A. Nicolau. Spark: a high-level synthesis framework for applying parallelizing compiler transformations. VLSI Design, 2003. Proceedings. 16th International Conference on, pages 461–466, 4-8 Jan. 2003.
- [23] V. Krishnan and S. Katkoori. A genetic algorithm for the design space exploration of datapaths during high-level synthesis'. IEEE Trans. Evolutionary Computation, 10(3):213–229, 2006.
- [24] C. Mandal, P. Chakrabarti, and S. Ghose. Design space exploration for data path synthesis. Proc. of the 10-th International Conference on VLSI Design, pages 166–170, 1996.
- [25] C. Mandal, P. P. Chakrabarti, and S. Ghose. GABIND: a GA approach to allocation and binding for the high-level synthesis of data paths. IEEE Trans. Very Large Scale Integr. Syst., 8(6):747–750, 2000.
- [26] M. Palesi and T. Givargis. Multi-objective design space exploration using genetic algorithms. In Proc. CODES'02 Workshop, pages 67–72, New York, NY, USA, 2002. ACM.
- [27] P. G. Paulin and J. P. Knight. Force-directed scheduling in automated data path synthesis. In Design Automation Conference, 1987.
- [28] C. Pilato, G. Palermo, A. Tumeo, F. Ferrandi, P. L. Lanzi, and D. Sciuto. Fitness inheritance in evolutionary and multi-objective highlevel synthesis. In IEEE Proceedings of CEC 2007 - Congress on Evolutionary Computation, 2007.
- [29] G. Papa, J. silc, Scheduling algorithms based on genetic approach, Proc. 4'h Conference on Neural networks and their applications, Zakopane, Poland, pp. 469-474, May 1999.
- [30] G. Papa, J. silc, The Use of Genetic Algorithm in the Integrated Circuit Design, Proc. 2nd Intl. Workshop on Design, Test and Applications WDTA'99, Dubrovnik, Croatia , pp. 73-76 June 1999.
- [31] G. Papa, J. silc, Using Simulated Annealing and Genetic Algorithm in the Automated Synthesis of Digital Systems, In Mastorakis (ed.): Recent advances in circuits and systems, World Scientific, pp. 377-381, 1998.

Huijing Yang received the B.S. degree in Computer Science and Technology from Heilongjiang University in 2002. And she received the M.S. degree in Computer Software and Theory from Harbin Engineering University in 2005. Her primary research area focused on IC design and Computer Architecture. Currently she is a Lecturer of the School of Software, Harbin University of Science and Technology, Harbin, China.

Chunying Wang was born in Jilin Province, China, in 1977 .She received the B.S. degree in Computer Science and Technology from Northeast Forestry University, China in 2000, and M.S. degree in Computer Software and Theory from Technology from Harbin University of Science and Technology, China in 2003. She has been working in Harbin University of Science and Technology from 2005. Currently, she is a in school of Software. Her major research interests include wireless network and network security.

Ning Du was born in Heilongjiang Province, China, in 1979 .She received the B.S. degree in Computer Science and Technology from Harbin University of Science and Technology, China in 2003 , and M.S. degree in Computer Software and Theory from Technology from Harbin University of Science and Technology, China in 2006. She has been working in Harbin University of Science and Technology from 2006. Currently, she is a in college of Software. Her major research interests include wireless network and network security.

Application of Particle Swarm Optimization in Figuring out Non-differentiable Point of Function

Mo Yuanbin

1 School of statistics and mathematics, Zhongnan University of economics and law, Wuhan 430073,China

2 Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis, Guangxi University for Nationalities ,Nanning, 530006 China

Email: moyuanbing@263.net

Zhao Xin-quan and Xiang Shu-jian

School of statistics and mathematics, Zhongnan University of economics and law, Wuhan 430073,China

Email: zxq556@163.com, xiangsj63@yahoo.com.cn

Abstract—Owing to the importance of non-differentiable point in a function for economy, engineering and theoretical analysis, this paper brings forward a novel methodology to figure out the non-differentiable point of function which is based on adjusting the models for global extremum and local extremum of particle swarm optimazation (PSO). The algorithm takes the difference between left and right difference quotients as the adaptive value of the particle. By them, it defines local extremum and global extremum for PSO and makes particle close to non-differentiable point of target function, and then figures the point out. The validity of the algorithm is verified by the result of numerical calculation.

Index Terms—Non-differentiable Point; Particle Swarm Optimization (PSO); Difference Quotient

I. INTRODUCTION

The non-differentiable point of function is importantly applied in analyzing the characteristics of function, economy and engineering practices. It will be used in figuring out the extremum of function. Therefore, it is actually and theoretically significant to figure out the point of function within certain interval. However, only few scholars pay attention to how to figure out non-differentiable point of function within certain interval until now.

It is not easy to figure out non-differentiable point of function. According to definition, the point has to be estimated one by one. That is obviously not an operable method. And there is no Iterative formula can be used to calculate non-differentiable point as used in other mathematic problems.

In this paper, by adjusting Particle Swarm Optimization (PSO)^[1,3-10] properly, the algorithm can be more suitable to figure out non-differentiable point of function.

And there are examples in the following text which shows the validity of the algorithm.

II. NON-DIFFERENTIABLE POINT OF FUNCTION

Derivable point ^[2] x_0 of function is the point that the follow limit is existed.

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \quad (1)$$

Non-differentiable point of function is the point that the limit is not existed.

According to the above definition, the non-differentiable point of function has to be figured out one by one to testify the existence of the limit. In this way, it is actually not feasible to figure out the point.

Particle Swarm Optimization (PSO) is a kind of optimized algorithm designed after the simulation of foraging behavior of birds. It imitates the group biological behavior of birds and the like. In it every organism is called as particle. PSO means to find out the optimum solution among a group of particles in n-dimensional space. The position of every particle means the vector for one solution. As an agent, the particle may memorize the optimum solution found by itself and acquire the optimum solution experienced by the whole group of particles to direct its movement and gradually iterate the most optimum solution.

Take the position of No.i particle in generation-j as $\mathbf{x}_i(j)$, the optimum solutions found out by itself and the whole group of particles are individual extremum (Pbest) $\mathbf{p}_{id}(j)$ and global extremum (Gbest) $\mathbf{p}_{gd}(j)$ respectively, its movement speed vector $\mathbf{v}_i(j)$, then No.i particle will move to the following position in generation-(j+1)

$$\mathbf{v}_i(j+1) = \omega\phi_1\mathbf{v}_i(j) + \eta_1\phi_2(\mathbf{p}_{id}(j) - \mathbf{x}_i(j)) + \eta_2\phi_3(\mathbf{p}_{gd}(j) - \mathbf{x}_i(j)) \quad (1)$$

$$\mathbf{x}_i(j+1) = \mathbf{x}_i(j) + \mathbf{v}_i(j+1) \quad (2)$$

Among which ω, η_1, η_2 are control parameters all taking values from $[0, 1]$, and ϕ_1, ϕ_2, ϕ_3 are random numbers in the interval $[0, 1]$. The movement of every particle may be illustrated as follows

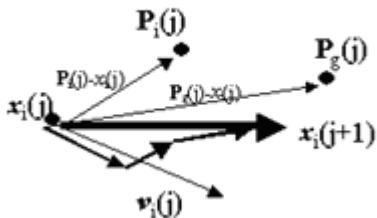


Fig.1 Complex Movement of Particle

III. FIGURE OUT NON-DIFFERENTIABLE POINT OF FUNCTION BY PSO

PSO adjusts itself in each step by contrasting the $p_{id}(j)$ and $p_{gd}(j)$ of its former step. In this way, the optimum value of function will be figured out after adjustment one step by one step. Owing to this thought, it is possible to figure out the solution for the certain characteristic by means of PSO iteration, if provided that $p_{id}(j)$ and $p_{gd}(j)$ have certain characteristic. Since the non-differentiable point of function is a point with certain characteristic, so it can be figured out by PSO.

For function $y = f(x), x \in [a, b]$

The key to figure out the non-differentiable point by PSO is how to confirm the global extremum and local extremum in PSO. Different global extremum and local extremum will bring forward different algorithm in figuring out non-differentiable point by PSO. According to the definition of derivative of function (1); if a function is not an odd function and is differentiable at certain point, the absolute value of left and right difference quotients at this point is relatively small;

$$p(x_0) = \left| \frac{f(x_0 + h) - f(x_0)}{h} - \frac{f(x_0 - h) - f(x_0)}{-h} \right| \quad (3)$$

If the function is at the non-differentiable point, the absolute value of left and right difference quotients at this point is relatively big.

If $p(x_0) = 0$ appears in the calculation of formula (3), it means $f(x)$ is an odd function within the local field of x_0 . If it is differentiable at x_0 , that

$$p(x_0) = \left| \frac{f(x_0 + h) - f(x_0)}{h} - \frac{\frac{f(x_0 + \frac{h}{2}) - f(x_0)}{\frac{h}{2}} - f(x_0)}{\frac{h}{2}} \right| \quad (4)$$

is relatively small; if the function is at the non-differentiable point, the absolute value is relatively large.

In this paper, the absolute value of the difference between left and right difference quotients of every particle is used as the adaptive of this particle. The individual extremum and global extremum are same with general algorithm. In this way, PSO is used to figure out the non-differentiable point of function.

A. Flow of Algorithm

The basic procedure in figuring out mean value by PSO is:

Step1 $k = 1$ given $h, M > 0$

Given m initial points like x_1, x_2, \dots, x_m and initial speed $v_i (i = 1, 2, \dots, m)$,

Step2 Figure out respective adaptive values by formula (4) or formula (3) according to $f(x)$ whether or not it is odd function.

Find out the maximum among $p(x_1), p(x_2), \dots, p(x_m)$ and take it as $f(p_{gd})$. Accordingly, the current global mean value point is p_{gd} . Local mean value is $p_{id} = x_i$, calculate $f(p_{id})$.

Step3 Update according to iteration formula (1), (2).

Step4 Confirm the local mean value.

On the assumption that after the update of Step2, the m points and their speeds are x_1, x_2, \dots, x_m and $v_i (i = 1, 2, \dots, m)$ respectively, calculate

$$p(x_1), p(x_2), \dots, p(x_m)$$

Compare $p(x_i), p(p_{id})$, and take the maximum as the new $p(p_{id})$, the local mean value point is p_{id} .

Step5 Confirm the global mean value

Compare $f(p_{id}), (i = 1, 2, \dots, m)$ and $f(p_{gd})$, take the maximum as the new $f(p_{gd})$, the global mean value point p_{gd} and the global mean value $f(p_{gd})$ are acquired.

Step6 $k = k + 1$, judge if it conforms to the condition of conclusion. If it does, it may conclude. If not, return to Step2.

Step7 When the final maximum is bigger than M , it is considered that the function is non-differentiable at $x = x_{gd}$. When it is smaller than M , it is considered that the function is differentiable at every point within the interval $[a, b]$.

B. Performance Test of Arithmetic

Test: to test following functions in different types,

$$(1) \quad y = |x|, \quad |x| \leq 20$$

$$(2) \quad y = \begin{cases} 1 & 0 \leq x \leq 20 \\ 0 & -20 \leq x < 0 \end{cases}$$

$$(3) \quad y = \begin{cases} \sin \frac{1}{x} & -20 \leq x \leq 20, \text{ and } x \neq 0 \\ 0 & x = 0 \end{cases}$$

$$(4) \quad y = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$(5) \quad y = \begin{cases} 1 - \cos x & x \geq 0 \\ x & x < 0 \end{cases}$$

$$(6) \quad y = x^{\frac{2}{3}}, \quad |x| \leq 20$$

$$(7) \quad y = |\ln|x-1||, \quad |x| \leq 0.5$$

The initial population of the algorithm is 100 and iteration generation number is 100.

Example 1 Parameter $h = 0.1$, after 200 times of calculation, the curve of the average of the acquired non-differentiable points was given in figure2 and the non-differentiable points was listed in the appendix A.

Example 2 Parameter $h = 0.01$, after 200 times of calculation, the curve of the average of the acquired non-differentiable points was given in figure3 and the non-differentiable points was listed in the appendix B.

Example 3 Parameter $h = 0.001$, after 200 times of calculation, the curve of the average of the acquired non-differentiable points was given in figure 4 and the non-differentiable points was listed in the appendix C.

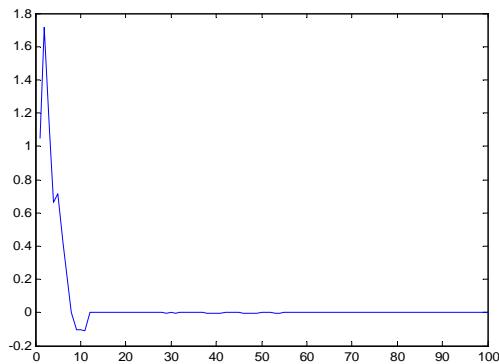


Fig.2 The curve of the relation the non-differentiable point and steps of example 1

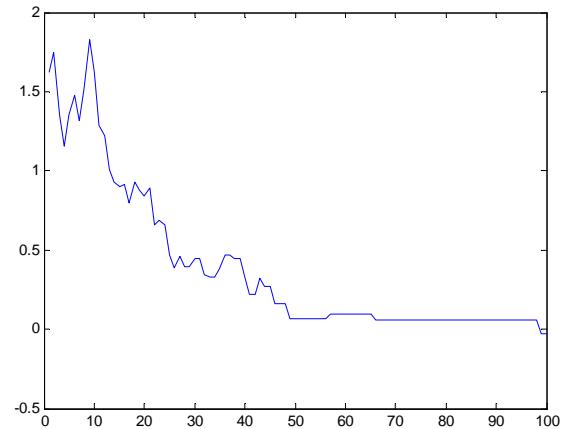


Fig.3 The curve of the relation the non-differentiable point and steps of example 2

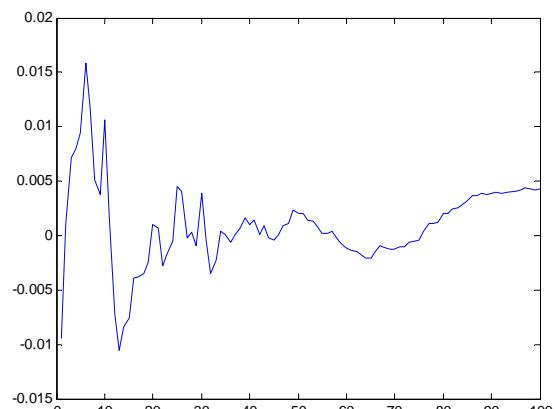


Fig.4 The curve of the relation the non-differentiable point and steps of example 3

Example 4 Parameter $h = 0.1$, after 200 times of calculation, the curve of the average of the acquired non-differentiable points was given in figure5 and the non-differentiable points was listed in the appendix D.

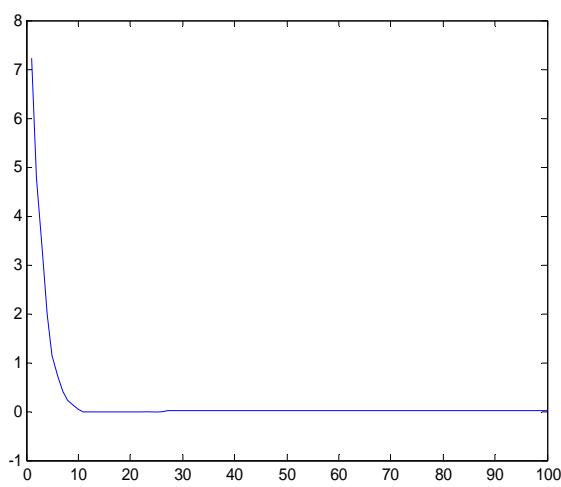


Fig.5 The curve of the relation the non-differentiable point and steps of example 4

Example 5 Parameter $h = 0.1$, after 200 times of calculation, the curve of the average of the acquired non-differentiable points was given in figure 6 and the non-differentiable points was listed in the appendix E.

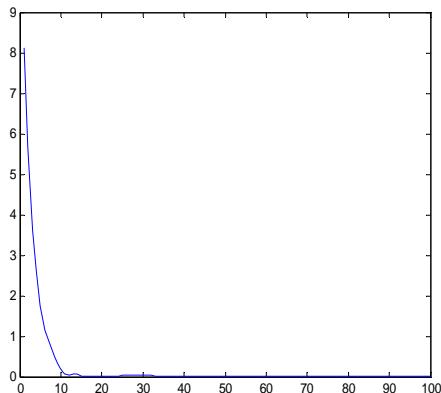


Fig.6 the curve of the relation the non-differentiable point and steps of example 5

Example 6 Parameter $h = 0.1$, after 200 times of calculation, the curve of the average of the acquired non-differentiable points was given in figure7 and the non-differentiable points was listed in the appendix F.

Example7 Parameter $h = 0.1$, after 200 times of calculation, the curve of the average of the acquired non-differentiable points was given in figure 8 and the non-differentiable points was listed in the appendix G.

From the above results, it is obvious that the algorithm can figure out non-differentiable points of function relatively and correctly within certain interval. This is because the algorithm makes proper definition for local extremum and global extremum in PSO, which may

move closely to the non-differentiable points of function and finally figure them out.

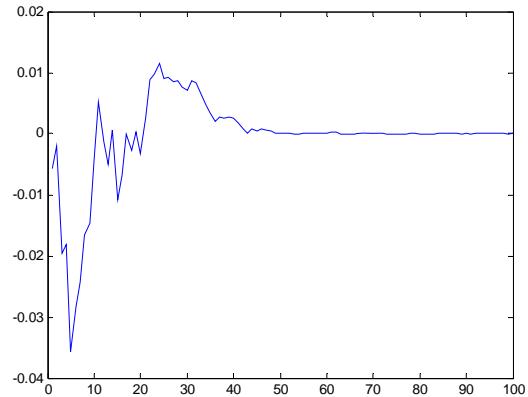


Fig.7 The curve of the relation the non-differentiable point and steps of example 6

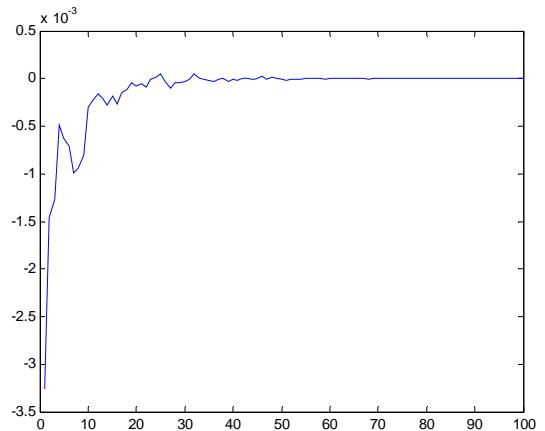


Fig.8 the curve of the relation the non-differentiable point and steps of example 7

C. Analysis of Time and Capability of Algorithms

The calculation of the algorithm is same with the calculation of standard PSO without additional calculation. Therefore, it is same with PSO in complicity of calculation. Through the above test, good result is figured out. This is because PSO is a kind of global search arithmetic. During the course of iteration, the main trend is always moving forward the optimum value. As a result, after global mean value and local mean value are defined in this paper, the main trend of the algorithm moves forward the non-differentiable points of function and finally figure out the non-differentiable points of function.

Performance analysis: The results of the above showed that the algorithm has strong searching property for figuring out Non-differentiable Point of function if the function exist the non-differentiable Point. It is because that the algorithm can fully take advantage of the idea of the PSO. By the definition of the non-differentiable point of the function, it gave extremum and global extremum. From the extremum, we can see that it fully use the

change rate in the neighborhood of one point of the function. So, in the course of the iteration, the extremum point closer to the non-differentiable by step and step, and finally find out the non-differentiable.

Convergence analysis: Since the algorithm is based on PSO, and no other performance was added to it. Then, if the non-differentiable is exists, the algorithm will be convergence to it

Time Complexity analysis: The algorithm is based on PSO, and only advance the extremum and global extremum definition of the non-differentiable, in order to make PSO can use to find out the non-differentiable, so the time complexity of this algorithm is same to the PSO's, and their run-time is similar.

IV. APPLICATION OF ALGORITHM

We all know that it is important to solving optimization problems, so lots of algorithm was advanced to solve it. Especially, evolution algorithm provides more possibility and advantage for optimization algorithm to solve optimization problems. But evolution optimization algorithm is easy to be trapped into local minima in optimizing it, and lots of improve tactics^[11-39] was put forward to improve evolution algorithm. Because, we know that the non-differentiable point is quietly possible the minima point, so it is entirely to improve one optimization algorithm by leading searching non-differentiable points mechanism to it to enhance the searching of the algorithm.

For example, we can improve Particle Swarm Optimization by added non-differentiable points searching to it (PSOS), and the principle is as follows::

$$\min f(\mathbf{x})$$

Step 1: Randomly produce k particles $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ in solution space, and given $h > 0$

Step 2 Figure out respective adaptive values by formula (4) or formula (3) according to $f(\mathbf{x})$ whether or not it is odd function. And find out the maximum among $p(\mathbf{x}_1), p(\mathbf{x}_2), \dots, p(\mathbf{x}_m)$, let it is $p(\mathbf{x}_0)$

Step 3: Compute the current global optimum \mathbf{p}_g and local optimum \mathbf{p}_i , substitutes iterative equations with the values

$$\begin{aligned} v_i(t+1) = & \alpha v_i(t) + c_1 \phi_i (\mathbf{p}_i(t) - \mathbf{x}_i(t)) + c_2 \phi_g (\mathbf{p}_g(t) - \mathbf{x}_i(t)) \\ & + c_3 \phi_i (\mathbf{x}_{0i} - \mathbf{x}_i(t)) \end{aligned}$$

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + v_i(t+1) \quad (i = 1, 2, \dots, k)$$

Compute $f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_k)$

Step 4: Return to Step 2, repeat the process until a stop criterion is met.

One benchmark functions are given to examine the performance of PSOS.

Test1 30 dimensions Griewank function

$$f(\mathbf{x}) = \frac{1}{4000} \sum_{i=1}^{30} x_i^2 - \prod_{i=1}^{30} \cos(x_i / \sqrt{i}) + 1$$

$$|x_i| \leq 600.$$

The best result is $\min f(\mathbf{x}) = f(0, 0, \dots, 0)$.

Test 2 Hartman's Function

$$f(\mathbf{x}) = -\sum_{i=1}^4 c_i \exp[-\sum_{j=1}^6 a_{ij} (x_j - p_{ij})^2],$$

where $0 \leq x_j \leq 1$, $c = (1 \ 1.2 \ 3 \ 3.2)$,

$$(p_{ij}) = \begin{bmatrix} 0.1312 & 0.1696 & 0.5569 & 0.0124 & 0.8283 & 0.5886 \\ 0.2329 & 0.4135 & 0.8307 & 0.3736 & 0.1004 & 0.9991 \\ 0.2348 & 0.1415 & 0.3522 & 0.2883 & 0.3047 & 0.6650 \\ 0.4047 & 0.8828 & 0.8732 & 0.5743 & 0.1091 & 0.0381 \end{bmatrix}$$

$$A = (a_{ij}) = \begin{bmatrix} 10 & 3 & 17 & 3.5 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{bmatrix}$$

$$\min(f(\mathbf{x})) = f(0.201, 0.15, 0.477, 0.275, 0.311, 0.657) = -3.32.$$

PSOS and PSO was run 10 times, and the final optimal results showed that PSOS is better than PSO

V. FURTHER EXTENSION OF ALGORITHM

By using this algorithm, the non-partially differentiable point of multi-function can also be figured out. Meanwhile, the maximum and minimum of function may be calculated in combination of this algorithm.

VI. CONCLUSION

In this paper, a novel method of constructing local extremum and global extremum in PSO is given according to the need of figuring out the non-differentiable point of function. By this method, PSO can be used to calculate the non-differentiable point of function at a certain interval. The results in the paper show the algorithm is practical.

APPENDIX A THE DATE OF EXAMPLE 1

0.0002	-0.0000	-0.0000	-0.0000	0.0000	0.0001
0.0011	0.0000	-0.0000	0.0045	-0.0000	0.0000
0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0000

APPENDIX B THE DATE OF EXAMPLE 2

0.0050	-0.8458	0.0043	0.0037	0.0022	0.0050
-0.0041	0.0011	0.0013	0.0035	0.0028	0.0024
0.0029	0.0048	0.0049	0.0009	0.0001	-0.0016

APPENDIX C THE DATE OF EXAMPLE 3

-0.0033	-0.0054	-0.0033	0.0075	0.0045	0.0075	-
-0.0036	0.0075	0.0075	0.0075	-0.0036	0.0075	
0.0075	0.0075	0.0075	0.0075	0.0075	0.0075

APPENDIX D THE DATE OF EXAMPLE 4

0.0001	0.0000	0.0049	0.0191	0.0016	0.0000
0.0392	0.0056	0.0148	0.0000	0.0000	0.0164
0.0137	-0.0216	0.0000	0.0088	0.0037	0.0000 ...

APPENDIX E THE DATE OF EXAMPLE 5

-0.0001	0.0212	0.0002	0.0056	0.0000	0.0178
0.0000	0.0000	0.0176	0.0001	0.0062	0.0000
0.0208	0.0000	0.0038	0.0110	0.0014	0.0356

APPENDIX F THE DATE OF EXAMPLE 6

1.0e-003 *					
-0.0083	-0.0244	-0.0227	-0.0107	-0.0638	0.1712
0.0216	-0.0200	-0.0085	-0.0088	-0.0437	-0.0067
0.0050	-0.0089	-0.0229	-0.0516	-0.0194	0.1974...

APPENDIX G THE DATE OF EXAMPLE 7

1.0E-005 *					
-0.1851	0.0515	-0.2821	-0.0893	0.0207	0.0264
0.1997	-0.1132	-0.3145	0.0292	-0.0787	0.0647
0.1891	0.0655	-0.2352	0.3584	-0.1287	-0.3723....

REFERENCES

- [1] KENNEDY J, EBERHART R C. Particle swarm optimization [A]. Proceedings of IEEE International Conference on Neural Networks. Piscataway NJ: IEEE , 1995: 1942-1948.
- [2] Mathematical Analysis(M), Department of Mathematics, East China Normal University. China Higher Education Press,2009.
- [3] Lu Yan-ping, Wang Sheng-rui. Li Shao-zi,etal. Particle swarm optimizer for variable weighting in clustering high-dimensional data [J]. Machine Learning, 2011,82:43-70
- [4] Kulkarni R V, Venayananamoorthy G K. Particle swarm optimization in wireless-sensor networks: a brief survey [J]. IEEE Transaction on Systems, Man and Cybernetics- Part C: Applications and Reviews, 2011:262-267.
- [5] Chen Guo-long, Guo Wen-zhong, Chen Yu-zhanong. A PSO -based intelligent decision algorithm for VLSI floor planning [J]. Soft Computing, 2010,14(12), 1329-1337
- [6] Liu H, Cai Z x, Wang Y . Hybridizing particle Swarm optimization with differential evolution for constrained numerical and engineering optimization. Applied Soft Computing, 2010, 10 (2): 629-640 .
- [7] Kim P, Lee J. A n integrated method of particle swam optimization and differential evolution. Journal of Mechanical Science and Technology, 2009, 23(2): 426 — 434.
- [8] Onwubolu, G. C. and Clerc, M., "Optimal path for automated drilling operations by a new heuristic approach using particle swarm optimization," *International Journal of Production Research*, vol. 4 pp. 473-491, 2004.
- [9] Rana S, Jasola S , Kumar R. A review on particle swarm optimization algorithms and their applications to data clustering. Artificial intelligence Review, 2011,35 (3): 211-222.
- [10] Glover F. Tabu search-part I. ORSA Journal on Computing, 1989, 1(3): 190-206.
- [11] Zhang C S, Ning J X, Lu S, Ou yang D T , Ding T N . A novel hybrid differential evolution and particle swarm optimization algorithm for unconstrained optimization. Operations Research Letters, 2009,37(2): 117-122.
- [12] Zhan Z H,Zhang J. Discrete particle swarm optimization for multiple destination routing problems [C]. Evo Workshops, 2009:117-122.
- [13] Kadirkamanathan V,Selvarajah K,Fleming P J.Stability analysis of the particle dynamics in particle swarm optimizer [J]. IEEE Transaction on Evolutionary Computation,2006,10(3):245-255.
- [14] Noel MM,Jeannette T C.Simulation of a newhybrid particle swarm optimization algorithm. Proceedings of the Thirty-Sixth Southeastern Symposium on System Theory,2004:150-153
- [15] Wachowiak MP,Smolkova R,Zheng Y,et al.An approach to Multimodal Biomedical Image Registration Utilizing Particle Swarm Optimization. IEEE Transac -tions on Evolutionary Computation, 2004, 8 (3) :289- 301
- [16] Victoire TAA,Jeyakumar AE.Hybrid PSO- SQP for economic dispatch with valve- point effect.Electric Power Systems Research,2004,71(1):51- 59.
- [17] Shi X,Lu Y,Zhou C,et al. Hybrid evolutionary algorithms based on PSOand GA.Proceedings of IEEE Congress on EvolutionaryComputation(CEC),Canberra, Australia, 2003:2393- 2399.
- [18] LIU Wei; CAI Qian-feng1; LIU Hai-lin. Particle swarm optimization algorithm based on parametric equation method to handle equality constraints[J]. Computer Engineering and Design. 2008 , 29 (3) ,697-699
- [19] CAO Chun-hong; ZHANG Bin; LI Wen-hui. The Research Based on the Composite Particle Swarm Optimization Algorithm in the Geometric Constraint Solving. Journal of Image and Graphics; 2007, 12(4),713-717.
- [20] WANG Jin-hua; YIN Ze-yong. Solving constrained discrete optimization problem with Particle Swarm Optimizer. Computer Engineering and Applications ; 2008,44(3), 242-244.
- [21] Wei Jing-xuan; Wang Yu-ping. Fuzzy Particle Swarm Optimization for Constrained Optimization Problems . Journal of Electronics & Information Technology; 2008, 30(5), 1218-1221.
- [22] WANG Fang; DING Haili; GAO Chengxiu. Advanced Particle Swarm Optimization for the Vehicle Routing Problem with Stochastic Demands. Journal of Wuhan University (Natural Science Edition); 2007, 53(1), 41-44.
- [23] LI Hong-mei; SUN Jun; XU Wen-bo. Empirical study based on Quantum-behaved Particle Swarm Optimization stochastic programming algorithm. Computer Engineering and Applications; 2007, 43(24), 185-189.
- [24] ZHAO Pei-xin; WANG Hong; MA Jian-hua. Particle Swarm Optimization Algorithm for Stochastic Loader Problem. Computer Engineering and Applications; 2006(19), 215-218.
- [25] LU Lin; TAN Qing-mei. Hybrid particle swarm optimization algorithm for stochastic vehicle routing problem. Systems Engineering and Electronics; 2006, 28(2), 244-247.
- [26] SUN Kai ; GE Xin-sheng. Optimal control of stretching process od solar array on spacecraft using particle swarm optimizatin algorithm. Engineering Mechanics ; 2007, 24(9), 188-192.)
- [27] Li Hong; Zhang Bing. Application of Particle Swarm Optimization in Balance Control for the Acrobot. Microcomputer Information; 2006,22,(3-1),80-82.

- [28] GUAN ShengTao; CHU Ji Zheng; SHAO Shuai. Application of nonlinear model predictive control based on particle swarm optimization. Journal of Beijing University of Chemical Technology (Natural Science Edition); 2007,34 (6),653-656.
- [29] MA Chang-xi; QIAN Yong-sheng; WANG Chun-lei. Urban loop-road traffic-coordination-control system based on TWP-PSO. Journal of Computer Applications; 2007, 27(11), 2640-2642.
- [30] MO Yuan-bin; CHEN De-zhao; HU Shang-xu. Particle Swarm Optimization for Multi-Objective Process System Optimization Problems. Journal of Chemical Engineering of Chinese Universities; 2008, 22(1), 94-99.
- [31] Schutte, J. F. and Groenwold, A. A., "Sizing design of truss structures using particle swarms," *Structural and Multidisciplinary Optimization*, 2003, vol. 25, no. 4, pp. 261-269.
- [32] HE Yijun; YU Huanjun; CHENG Biao; CHEN Dezhao. Multi-objective particle swarm optimization approach to solution of fed-batch bioreactor dynamic multi-objective optimization. *Journal of Chemical Industry and Engineering(China)*; 2007, 58(5), 1262-1270.
- [33] ZHANG Wen-ming; DONG Zeng-chuan; ZHU Cheng-tao; QIAN Wei. Automatic calibration of hydrologic model based on multi-objective particle swarm optimization method. *Journal of Hydraulic Engineering*; 2008, 39(5), 528-534.
- [34] PENG Zhi-ping; CHEN Ke. A Multi-Objective Particle Swarm Optimization Algorithm for Solving Negotiation Deadlock. *Acta Electronica Sinica*; 2007, 35(8), 1452-1457.
- [35] JIN Xin-Lei; MA Long-Hua; WU Tie-Jun; QIAN Ji-Xin. Convergence Analysis of the Particle Swarm Optimization Based on Stochastic Processes. *Acta Automatica Sinica*; 2007 33 (12) 1263-1268.
- [36] CUI Hong-mei; ZHU Qing-bao. Convergence analysis and parameter selection in particle swarm optimization. *Computer Engineering and Applications*; 2007, 43(23), 89-92.
- [37] LIU Hong-bo; WANG Xiu-kun; TAN Guo-zhen. Convergence Analysis of Particle Swarm Optimization and Its Improved Algorithm Based on Chaos. *Control and Decision*; 2006, 21(6), 636-641
- [38] LUO Jin-yan; JIANG Zhong-liang. Stability Analysis of the Particle Swarm Optimization Algorithm. *Journal of Jimei University (Natural Science)*; 2007, 12(4) 376-379. 2007, 12(4) 376-379
- [39] Correa E S, Freitas A A, Johnson G G. A new discrete particle swarm algorithm applied to attribute selection in a bioinformatics data set [C]// *Genetic and Evolutionary Computation Conference*. New York: ACM, 2006:35-42

Yuanbin Mo received the B.S. degree from the Guangxi University, Nanning, China, in 1991 and the M.S. degree from Guizhou University, Guiyang, China, in 2001 and the Ph.D. degree from Zhejiang University, Hangzhou, China, in 2007. Currently, he is a Associate Professor at Guangxi University for Nationalities, Nanning, China. His research interest covers computational intelligence and complex network theory.

A Robust Watermarking Against Shearing Based on Improved S-Radon Transformation

Deng Minghui^{1,2}

¹School of Astronautics, Harbin Institute of Technology

²Department of Engineering, Northeast Agricultural University, Harbin, China

Email: markdmh@163.com

Zeng Qingshuang*

School of Astronautics, Harbin Institute of Technology

Harbin, China

Email: zqshuang2000@yahoo.com.cn

Zhou Xiuli

Department of Engineering, Northeast Agricultural University

Harbin, China

Email: zhoushiuli@neau.edu.cn

Abstract—In this paper, a robust image watermarking method in two-dimensional space/spatial-frequency distributions domain is proposed which is robust against geometric distortion. This watermarking is detected by a linear frequency change. The one-dimensional S transformation and radon transformation are used to detect the watermark. The chirp signals are used as watermarks and this type of signals is resistant to all stationary filtering methods and exhibits geometrical symmetry. In the two-dimensional Radon-Wigner transformation domain, the chirp signals used as watermarks change only its position in space/spatial-frequency distribution, after applying linear geometrical attack, such as scale rotation and cropping. But the two-dimensional Radon-Wigner transformation needs too much difficult computing. So the image is put into a series of 1D signal by choosing scalable local time windows. The watermark embedded in the 1D improved S-Radon transformation domains. The watermark thus generated is invisible and performs well in StirMark test and is robust to geometrical attacks. Compared with other watermarking algorithms, this algorithm is more robust, especially against geometric distortion, while having excellent frequency properties.

Index Terms—Digital watermarking, improved S-Radon transform, Geometrical attack

I. INTRODUCTION

With the arrival of the information era and the broad application of E-business, there is a growing importance to protect the security of messages. As an important branch in the field of the research on the message cryptic technique, the digital watermarking technique is an efficient way to the authentication of content and copyright. This technique authenticate and protect the data by imbed watermark in the original data. The watermark imbedded can be a passage, some marks or

images. The traditional encryption can only assure the message security when being visited and the security of both parts when in a single-phase communication mode, but to the public messages transformed in the multi phase mode a new technique and mechanism is needed. As a potential method to solve the problem, digital watermarking technique is being widely concerned, and it is becoming the top research in the international academic field.

Digital watermark is a special mark cryptic in the multi-media products. Digital watermark should have three basic characteristics: Insensitive, that is the imbed watermark can't destroy the digital products, and we can feel the exist of the watermark neither visual nor aural; robustness, that is under the usual signal processing (compressed, rejected or effected by noise) and geometric transmitting (translated, flexed or rotated), It can assure that the watermark can't be destroyed. The imbedded watermark can be done in time-space frequency, and it can also be done in the transformable domain. The first method is easy to be carried out, but the protecting from the attack to signal processing can't be done perfectly. However, the watermarking method under transformable domain is better. The robustness must be better in the efficient digital watermark in [1]-[7].

In this paper we put forward a robust digital image watermarking based on S-Radon transform. In Srdjan Stankovic's paper, a watermarking algorithm in the space/spatial domain using two-dimensional Radon-Wigner distribution is introduced. This algorithm uses of the Radon-Wigner transform to detect the watermark and the two-dimensional chirp signals are used as watermarks. In the two-dimensional Radon-Wigner transformation domain, the chirp signals used as watermarks change only its position in space/spatial-frequency distribution, after applying linear geometrical attack, such as scale rotation

and cropping. Compared with other watermarking algorithms, this algorithm is more robust, especially against geometric distortion, while having excellent frequency properties. But the 2D Radon-Wigner transformation needs much difficult computing and can be impossible in reality. So we introduce a algorithm based on 1D S transform. In this algorithm, the chirp signals used as watermarks are inserted in the image and the image is put into a series of 1D signal by choosing scalable local time windows. By using improved S-Radon transformation on the 1D image signal series, the watermark is detected. The shearing attack can break watermarks in one part of space support district, but watermarks in another one part of space support district still can not be destroyed. Synthesizing each supporting space, the watermark extracted still can be clear and the algorithm achieves the robustness to the shearing attacks.

II. THE PRINCIPLE OF S TRANSFORM

As a linear time-frequency analysis, S transform in [8] has some features similar to the nature of time-frequency domain of the Fourier transform and the wavelet transform. For example, it is a reversible transformation of non-destructive and its inverse transform can perfectly reconstruct the original signal. So the time-frequency invariant feature ensures that the invariant features with the transformation signal for a specific application. One-dimensional S transform is:

$$S(\tau, f) = \int_{-\infty}^{+\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(t-\tau)^2 f^2}{2}} e^{-i2\pi ft} dt \quad (1)$$

where t, τ are t domain variables and f is frequency domain variable. One-dimensional signal $h(t)$ is mapped from the one-dimensional time domain to the two-dimensional time-frequency plane through S transform. One-dimensional S inverse transform is

$$h(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} S(\tau, f) d\tau e^{i2\pi ft} df \quad (2)$$

If S transform is local spectrum, the Fourier spectrum can be received by computing the average local spectrum through the whole time domain. So S transformation:

$$\int_{-\infty}^{+\infty} S(\tau, f) d\tau = H(f) \quad (3)$$

$H(f)$ is the Fourier transform of $h(t)$. $h(t)$ can be deduced from $S(\tau, f)$. S transform is the general Fourier transform of non-stationary time series.

S transform is the linear computation of time series $h(t)$. The transformed noise can often influence time-frequency resolution ratio. If signal $x(t)$ is equal to the sum of original data $s(t)$ and noise $n(t)$.

$$x(t) = s(t) + n(t) \quad (4)$$

After S transform:

$$S\{x(t)\} = S\{s(t)\} + S\{n(t)\} \quad (5)$$

The S transform can not creat cross terms and overwhelmingly increase the time-frequency resolution

ratio. If the S transform of $h(t)$ is $S(\tau, f)$, the S transform of $h(t - r)$ is $S(\tau - r, f)e^{-i2\pi fr}$.

From equation 3, we can know that S transform has direct connection with Fourier transform. S transform and S inverse transform are a lossless reversible procedure. S transform will not create the cross terms and has nice time-frequency energy centralization quality.

Two-dimensional S transform is based on one-dimensional S transform to develop, that formula to transform is:

$$S(x, y, k_x, k_y) = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} h(x', y') \frac{|k_x|}{\sqrt{2\pi}} e^{-(x'-x)^2/2} e^{-i2\pi k_x x'} dx' \right] \frac{|k_y|}{\sqrt{2\pi}} e^{-(y'-y)^2/2} e^{-i2\pi k_y y'} dy' \quad (6)$$

(6)

where $h(x', y')$ is the two-dimensional image and (x', y') is the space domain variables. After transformation, S transform spectrum contains 4 variables (x, y, k_x, k_y) . (x, y) are the variables in space domain and (k_x, k_y) are the variables in frequency domain, also known as the wavelength. Similar to the two-dimensional Fourier transform, the nature of the two-dimensional S transform can be seen as a cascade of two one-dimensional S transformations.

Reference to the fast Fourier transform, the $h(n)(n = 0, 1, L, N)$ is the corresponding $h(t)$ discrete time series and sampling time interval is T. The discrete Fourier transform is:

$$H\left(\frac{1}{NT}\right) = \frac{1}{N} \sum_{k=0}^{N-1} h(kT) e^{-\frac{i2\pi nk}{N}} \quad (7)$$

The discrete S transformation of time series $h(t)$ is as follows:

$$S[jT, \frac{n}{NT}] = \sum_{m=0}^{N-1} H\left[\frac{m+n}{NT}\right] e^{-\frac{2\pi^2 m^2}{n^2}} e^{-\frac{i2\pi mn}{N}} \quad (8)$$

When $n = 0$ (equivalent to zero frequency), discrete form of expression is:

$$S[jT, 0] = \frac{1}{N} \sum_{m=0}^{N-1} h\left[\frac{m}{NT}\right] \quad (9)$$

Equ.9 ensures that the time series of anti-transformation can be accurate. Of course, discrete S transformation has been limited by sampling and the length and will have a border effect in time and frequency domain.

Discrete S inverse transformation is to obtain by calculating the discrete Fourier transform. When n is not equal to 0, the summation of S matrix ($S[n, m]$) along the line is:

$$S[jT, \frac{n}{NT}] = \sum_{j=0}^{N-1} \sum_{m=0}^{N-1} H\left[\frac{m+n}{NT}\right] e^{-\frac{2\pi^2 m^2}{n^2}} e^{-\frac{i2\pi mn}{N}} \quad (10)$$

Equ.10 can be turned into:

$$S[jT, \frac{n}{NT}] = \sum_{m=0}^{N-1} H\left[\frac{m+n}{NT}\right] e^{-\frac{2\pi^2 m^2}{N^2}} \sum_{j=0}^{N-1} e^{-\frac{i2\pi nj}{N}} \quad (11)$$

The average of $S[jT, \frac{n}{NT}]$ is:

$$S[jT, \frac{n}{NT}] = \sum_{m=0}^{N-1} N\delta_{m,0} H\left[\frac{m+n}{NT}\right] e^{-\frac{2\pi^2 m^2}{N^2}} \quad (12)$$

$$\frac{1}{N} S[jT, \frac{n}{NT}] = H\left[\frac{n}{NT}\right] \quad (13)$$

Therefore, discrete inverse S transformation is:

$$h[kT] = \frac{1}{N} \sum_{n=0}^{N-1} \left\{ \sum_{j=0}^{N-1} S\left[\frac{n}{NT}, jT\right] \right\} e^{-\frac{i2\pi nj}{N}} \quad (14)$$

When n is equal to zero, the width of the Gaussian function is zero. Zero frequency is the average of time series and is constant. $S[jT, \frac{n}{NT}]$ is the average of $h[kT]$ when the value of n reduced to zero. That is, every value along zero n value is replaced by this value. In this way, it ensures that S transformation is completely reversible.

The generalized S-transform is given by

$$S(\tau, f, \beta) = \int_{-\infty}^{+\infty} h(t) \omega(t, f, \beta) e^{-i2\pi ft} dt \quad (15)$$

where ω is the window function of the S-transform and β denotes the set of parameters that determine the shape and property of the window function. The window satisfies the normalized condition

$$\int_{-\infty}^{+\infty} \omega(t, f, \beta) dt = 1 \quad (16)$$

The alternative expression of (15) by using the convolution theorem through the Fourier transform can be written as

$$S(\tau, f, \beta) = \int_{-\infty}^{+\infty} X(\alpha + f) W(\alpha, f, \beta) e^{i2\pi\alpha\tau} d\alpha \quad (17)$$

Where:

$$X(\alpha + f) = \int_{-\infty}^{+\infty} h(t) e^{-i2\pi(\alpha+f)t} dt \quad (18)$$

And

$$W(\alpha, f, \beta) = \int_{-\infty}^{+\infty} \omega(t, f, \beta) e^{-i2\pi\alpha t} dt \quad (19)$$

The variable α and f in the above expression have the same units. In this scheme we retain the window function as the same Gaussian window because it satisfies the minimum value of the uncertainty principle. We have introduced an additional parameter (β) into the Gaussian window where its width varies with frequency as follows

$$\sigma(f) = \frac{\delta}{|f|} \quad (20)$$

Hence the generalized S-transform becomes

$$S(\tau, f, \delta) = \int_{-\infty}^{+\infty} h(t) \frac{|f|}{\sqrt{2\pi\delta}} e^{-\frac{-(\tau-t)^2 f^2}{2\delta^2}} e^{-i2\pi ft} dt \quad (21)$$

Where the Gaussian window becomes

$$\omega(t, f, \delta) = \frac{|f|}{\sqrt{2\pi\delta}} e^{-\frac{-t^2 f^2}{2\delta^2}} \quad (22)$$

And its frequency domain representation is

$$W(\alpha, f, \delta) = e^{-\frac{2\pi^2 \sigma^2 \delta^2}{f^2}} \quad (23)$$

III THE PRINCIPLE OF RADON TRANSFORM

The 2D Radon transformation is the projection of the image intensity along a radial line oriented at a specific angle [9]. Radon expresses the fact that reconstructing an image, using projections obtained by rotational scanning is feasible. His theorem is the following: The value of a 2-D function at an arbitrary point is uniquely obtained by the integrals along the lines of all directions passing the point. The Radon transformation shows the relationship between the 2-D object and its projections [10].

The Radon Transformation is a fundamental tool which is used in various applications such as radar imaging, geophysical imaging, nondestructive testing and medical imaging [11]. Many publications exploit the Radon Transformation. Meneses-Fabian et al. [12] describe a novel technique for obtaining border-enhanced topographic images of a slice belonging to a phase object. Vitezslav [13] examines fast implementations of the inverse Radon transform for filtered back projection on computer graphic cards. Sandberg et al. [14] describe a novel algorithm for topographic reconstruction of 3-D biological data obtained by a transmission electron microscope. Milanfar [15] exploits the shift property of Radon transformation to image processing. Barva et al. [16] present a method for automatic electrode localization in soft tissue from radio-frequency signal, by exploiting a property of the Radon Transform. Challenor et al. [17] generalize the two dimensional Radon transform to three dimensions and use it to study atmospheric and ocean dynamics phenomena.

Figure 2 illustrates several 1D projections from different angles of an image consisting of three spots in the 2D domain. In some of the projections, only two spots are shown. This reveals the importance of the selection of the “correct” projections for image reconstruction.

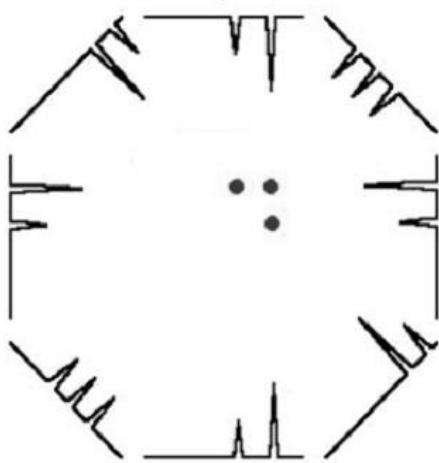


Figure1. Different projections of a three-dot image example

Suppose a 2-D function $f(x, y)$ (Fig. 3). Integrating along the line, whose normal vector is in θ direction, results in the $g(s, \theta)$ function which is the projection of the 2D function $f(x, y)$ on the axis s of θ direction. When s is zero, the g function has the value $g(0, \theta)$ which is obtained by the integration along the line passing the origin of (x, y) -coordinate. The points on the line whose normal vector is in θ direction and passes the origin of (x, y) -coordinate satisfy the equation:

$$\frac{y}{x} = \tan(\theta + \frac{\pi}{2}) = \frac{-\cos\theta}{\sin\theta} \Rightarrow \\ \Rightarrow x\cos\theta + y\sin\theta = 0$$

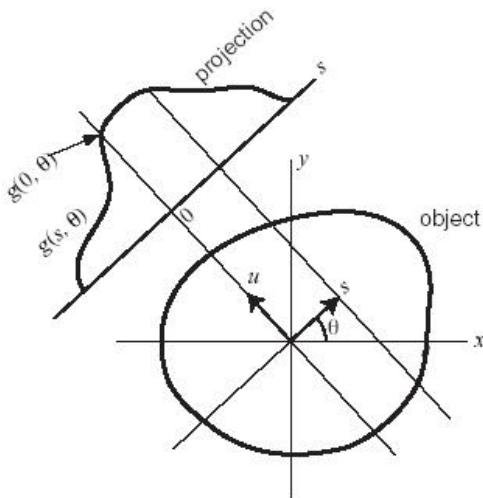


Figure2. The Radon Transform computation

The integration along the line whose normal vector is in θ direction and that passes the origin of (x, y) -coordinate means the integration of $f(x, y)$ only at the points satisfying the previous equation. With the help of the Dirac "function" δ , which is zero for every

argument except to 0 and its integral is one, $g(0, \theta)$ is expressed as:

$$g(0, \theta) = \iint f(x, y) \cdot \delta(x\cos\theta + y\sin\theta) dxdy \quad (24)$$

Similarly, the line with normal vector in θ direction and distance s from the origin is satisfying the following equation:

$$(x - s \cdot \cos\theta) \cdot \cos\theta + (y - s \cdot \sin\theta) \cdot \sin\theta = 0 \Rightarrow \\ x\cos\theta + y\sin\theta - s = 0$$

So the general equation of the Radon transformation is acquired: [10, 11, 15, 16, 18]

$$g(s, \theta) = \iint f(x, y) \cdot \delta(x\cos\theta + y\sin\theta - s) dxdy \quad (25)$$

The inverse of Radon transform is calculated by the following equation [14] :

$$f(x, y) = \int_{-\pi/2}^{\pi/2} \rho \cdot R_\theta(g(s, \theta)) d\theta \quad (26)$$

where R_θ is the Radon transformation, ρ is a filter and $s(x, y) = x\cos\theta + y\sin\theta$

Radon-Wigner transformation is a kind of projective transformation of linear integration. It is the Radon transformation of linear integration projection for the signal wigner transformation, as shown in the Figure3.

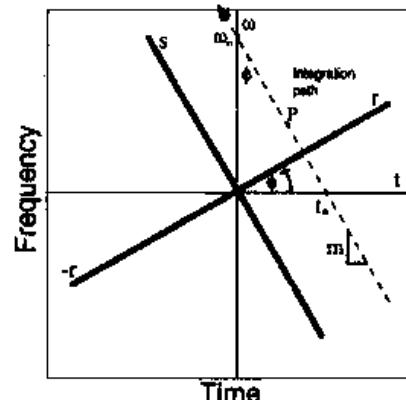


Figure3. Schematic illustrating the geometry for calculation of the Radon-Wigner spectrum

The Wigner distribution of the image $I(x, y)$ is defined as:

$$WD(x, y, \omega_x, \omega_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I(x + \xi/2, y + \zeta/2) I^*(x - \xi/2, y - \zeta/2) \\ \times e^{-j(\omega_x \xi + \omega_y \zeta)} d\xi d\zeta$$

its pseudo form is used in practical realizations

$$WD(x, y, \omega_x, \omega_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I(x + \xi/2, y + \zeta/2) I^*(x - \xi/2, y - \zeta/2) \\ \times w(\xi, \zeta) w^*(-\xi, -\zeta) e^{-j(\omega_x \xi + \omega_y \zeta)} d\xi d\zeta$$

where $w(\xi, \zeta)$ is a 2D window function. The 2D chirp signal, used here as a base for watermarking, has this form:

$$\begin{aligned}\Omega(x, y) &= 2A \cos(ax^2 + by^2 + c) \\ &= A(e^{j(ax^2 + by^2 + c)} + e^{-j(ax^2 + by^2 + c)})\end{aligned}$$

where A is the watermark amplitude or strength. Note that the Wigner distribution of this signal is highly concentrated.

$$\begin{aligned}WD(x, y, \omega_x, \omega_y) &= A^2 W(\omega_x - 2ax, \omega_y - 2by) + \\ &A^2 W(\omega_x + 2ax, \omega_y + 2by) + \\ &\text{cross-terms}\end{aligned}$$

where the Fourier transform of the window, $w(\omega_x - 2ax, \omega_y - 2by)$, is close to a delta function $\delta(\omega_x - 2ax, \omega_y - 2by)$ for a sufficiently wide window, since the cross-terms in the Wigner distribution will be eliminated by use of projections, they will be neglected in the sequel.

After a general linear geometrical transformation, signal can be written in this form

$$\Omega'(x, y) = 2A \cos(a_1 x^2 + a_2 y^2 + a_3 xy + a_4 x + a_5 y + a_6)$$

This transformation corresponds to a mapping of centered ellipse into the rotated one, whose center is displaced from the origin. From the point of view of the Wigner distribution concentration on the local frequency, we may say that it is invariant with respect to this transformation. Only the position of the local frequency of the distribution concentration will be changed.

$$\begin{aligned}WD(x, y, \omega_x, \omega_y) &= A^2 W(\omega_x - 2a_1 x - a_3 y - ax, \\ &\omega_y - 2a_2 y - a_3 x - a_5) + \\ &A^2 W(\omega_x + 2a_1 x + a_3 y + ax, \\ &\omega_y + 2a_2 y + a_3 x + a_5) + \\ &\text{cross-terms}\end{aligned}$$

This means that the described geometrical transformation does not influence the maximal value of the Wigner distribution which we intend to use for the watermark detection. The Wigner distribution of the watermark remains close to the delta pulse. Although, we have a very specific and recognizable function over the enter space, its energy could be much smaller than the energy of the Wigner distribution of an image because the value of A should be drastically smaller than the average image values. For this reason the watermark detection of us only the Wigner distribution is not reliable enough.

IV. THE IMBEDDING AND TEST OF DIGITAL WATERMARK

We consider how to construct a watermark to insert into the image. In the Srdjan Stankovic and Igor Djurovic's paper, the two-dimensional chirp signals are used as watermarks and in the algorithm two-dimensional Radon-Wigner transformation is applied to additionally concentrate the energy of the watermark signal and shows perfect robustness to the geometrical attacks. But the

computing of two-dimensional Radon-Wigner needs too much time and could be very difficult. This algorithm is very impractical and the ordinary computer could not finish this work. So we want to look for a new time-frequency distributions domain algorithm to solve this problem.

We imbed the watermark in the S transformation domain of image. In Stockwell's paper, the S transformation is introduced and can detect linear frequency-modulated signals. But 2D S transformation needs expensive computing. Obviously, it is necessary to apply one-dimensional S transformation on image and additionally concentrate the energy of the watermark signals. We select the linear frequency-modulated signals as watermark. The digital watermark is W with the sum of many linear frequency-modulated signals with different frequency:

$$W(n) = \cos[2\pi(f_1 + k_1 n)T] + L + \cos[2\pi(f_m + k_m n)T] \quad (28)$$

The length of W is n , and then choose D_0 and D_1 two areas with the same size of watermark in wavelet transformation middle frequency domain LH_0 and HL_0 of digital image frame C_{ij} . The method to imbed watermark is as followed:

$$D'_0(i, j) = (D_0(i, j) + W(n)), D'_1(i, j) = (D_1(i, j) + W(n)) \quad (29)$$

Then we synthesize wavelet to get watermark image. All the frames be done the same way as above-mentioned calculate ways. When withdrawing watermark, we carry on wavelet decomposition again and withdraw a 1-D signal from the known domains. We make S transformation on the 1-D signal and detect the linear frequency-modulated signals that are the watermarks. Then we take Radon transformation on the S transformation image and the more clear watermarks is shown.

In this paper, we use standard 256×256 gray image Lena as an original image. Applying Haar wavelet transformation in the algorithm, the image after imbedded the two and three linear frequency-modulated signals with different frequency as watermark is shown in Figure4. The detecting result is shown in Figure5 and Figure6. The picture frame decomposition adding three watermarks cuts pictures in the different position and the different size. After cutting an attack withdraw watermark. We cut 50% and 75% of the image random such as Figure7 Figure12. Then extracted watermark result is shown such as follows. According to the result of the experiment, it can be seen that the watermarking image can still be extracted well even the original image is shearing attacked by 75% with the S-Radon transformation. This algorithm is better than the S transformation and Wigner transformation. This proves the efficient of the method used above. In the testing process, this algorithm can be used in the reality.



Figure4. Watermarked image

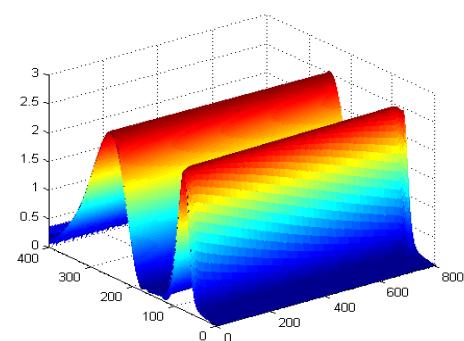


Figure8. The two Watermarks extracted with S transformation

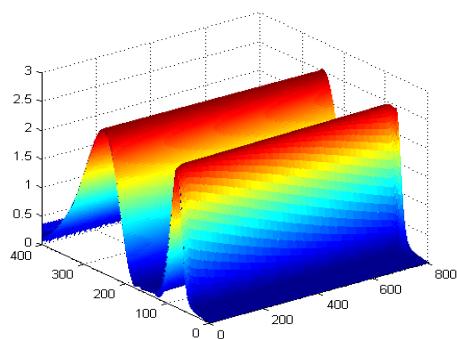


Figure5. The two Watermarks extracted with S transformation

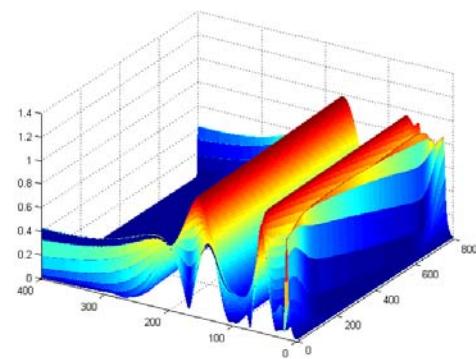


Figure9. The three Watermarks extracted with S transformation

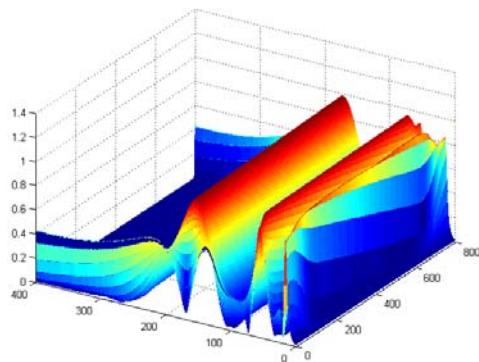


Figure6. The three Watermarks extracted with S transformation

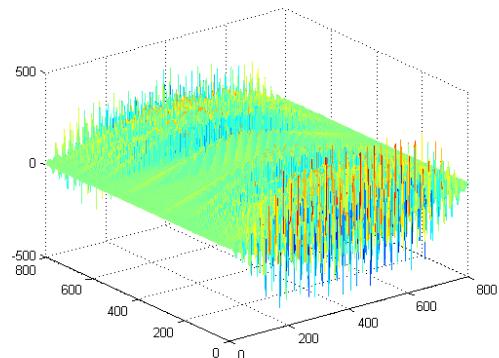


Figure10. The three Watermarks extracted with Wigner transformation



Figure7. Sheared by 50% in the middle

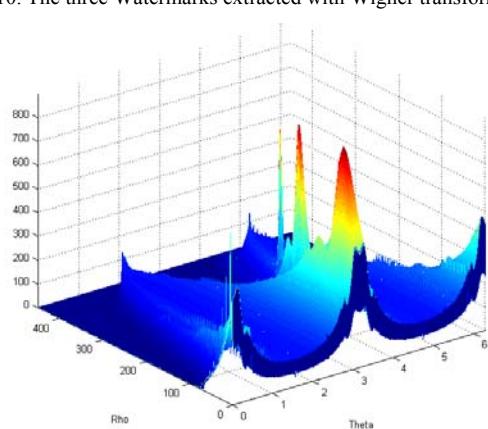


Figure11. The three Watermarks extracted with improved S-Radon transformation



Figure12. Sheared by 75% upside

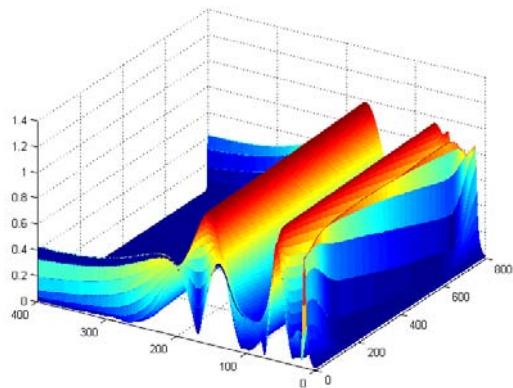


Figure13. The three Watermarks extracted with S transformation

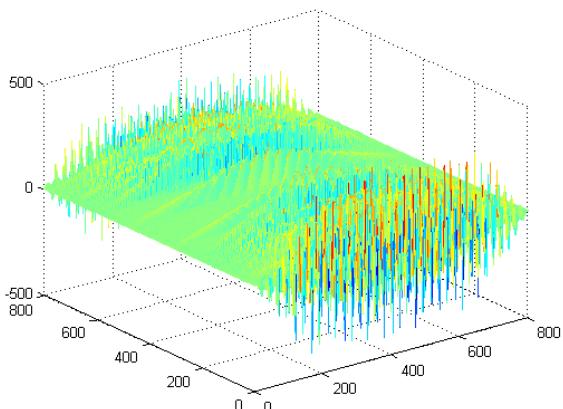


Figure14. The three Watermarks extracted with Wigner transformation

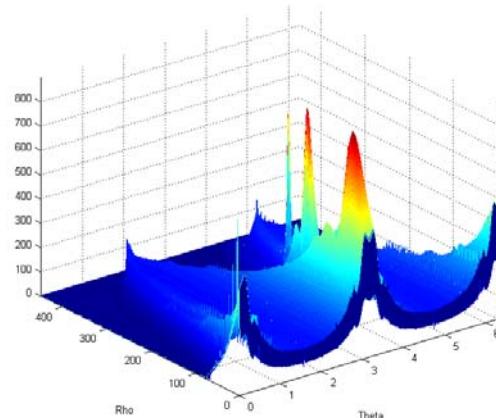


Figure15. The three Watermarks extracted with improved S-Radon transformation

V. CONCLUSIONS

In this paper, a robust watermarking method against shearing based on S-Radon transform is introduced. The proposed method makes use of the person's sense of vision characteristics and wavelet transformation to achieve the improved S-Radon transform on the image. The linear frequency-modulated signals are selected as watermarks and are added in middle frequency coefficients in the transformation matrix. Based on 1D S transform and Radon transformation, the watermark is extracted. The method improves the validity of watermarking and shows excellent advantage against shearing attack.

ACKNOWLEDGMENT

This paper is supported by the 2009 Natural Science Program of Heilongjiang Province Educational Committee (NO.11541016) and the Doctor Startup Funding of Northeast Agricultural University

REFERENCES

- [1] RUANAIDH J J K, PUN T. Rotation, scale and translation spread spectrum digital image watermarking [J]. *Signal Processing*, 1998, 66(3): 303~317.
- [2] SRDJAN STANKOVIC, IGOR DJUROVIC, IOANNIS PITAS. Watermarking in the space/spatial-frequency domain using two-dimensional radon-wigner distribution [J]. *IEEE Transactions on Image Processing*. 2001, 10(4): 650-658.
- [3] COX I J, KILIAN J, LEIGHTON F T. Secure Spread Spectrum Watermarking For Multimedia [J]. *IEEE Tran. of Image Processing*, 1997, 6(12), 1673~1687.
- [4] LINNARTZ J P, DEPOVERE G, and KALKER T. On the Design of a Watermarking system: Considerations and Rationales [A]. In: Information hiding Third International Workshop, IH'99, Dresden, Germany. (1999), 253~269.
- [5] R G VAN SCHYNDEL, A Z TIRKEL, CF OSBORNE. [A]. In: Proceeding of IEEE International Conference on Image Processing [C], Vancouver, Canada 1994, 2:86~94.
- [6] NIV Xia-mu, LU Zhe-ming, SVN Sheng-he. Digital watermarking of still image with gray -level digital watermarks [J]. *IEEE Trans .On Consumer Electronics*, 2000, 46(1), 137~239.

- [7] Huang Ji-wu, Tan Tie-niu. Summary of the Image In plate Watermark [J]. Electronic Paper, 2000, 26(5): 645~655.
- [8] STORCKWELL R G, MANSINHA L, LOWE R P, Localization of the Complex Spectrum: The S-Transform. IEEE Transactions on Signal Processing, 1996, 44(4): 998-1001.
- [9] Mathworks, <http://www.mathworks.com>, accessed 15 June 2005.
- [10] A. Asano, "Radon transformation and projection theorem", Topic 5, Lecture notes of subject Pattern information processing, 2002 Autumn Semester, <http://kuva.mis.hiroshima-u.ac.jp/~asano/Kougi/02a/PIP/>
- [11] A. Averbuch, R.R. Coifman, D.L. Donoho, M. Israeli, J. Wald'en, Fast Slant Stack: A notion of Radon Transform for Data in a Cartesian Grid which is Rapidly Computable, Algebraically Exact, Geometrically Faithful and Invertible., to appear in SIAM J. Scientific Computing, 2001
- [12] C. Meneses-Fabian, G. Rodríguez-Zurita, and J.F. Vázquez-Castillo "Optical tomography of phase objects with parallel projection differences and ESPI", Investigacion revista mexicana de fisica 49 (3) 251-257 Junio 2003.
- [13] V.V. Vlcek, "Computation of Inverse Radon Transform on Graphics Cards", International Journal of Signal Processing 1(1) 2004 1-12.
- [14] K. Sandberg, D. N. Mastronarde, G. Beylkina, "A fast reconstruction algorithm for electron microscope tomography", Journal of Structural Biology 144 (2003) 61-72, 3 September 2003.
- [15] P. Milanfar, "A Model of the Effect of Image Motion in the Radon Transform Domain", IEEE Transactions on Image processing, vol. 8, no. 9, September 1999
- [16] M. Barva and J. Kybic with J. Mari and C. Cachard, "Radial Radon Transform dedicated to Micro-object Localization from Radio Frequency Ultrasound Signal", In UFFC '04: Proceedings of the IEEE International Ultrasonics, Ferroelectrics and Conference. Piscataway: IEEE, 2004, p. 1836-1839. ISBN 0-7803-8412-1.
- [17] P.G. Challenor , P. Cipollini and D. Cromwell, "Use of the 3D Radon Transform to Examine the Properties of Oceanic Rossby Waves", Journal of Atmospheric and Oceanic Technology, Volume 18, 22 January 2001.



Deng Minghui was born 1976 in Harbin China. He received the B.S. degree in application of electronic technology from Harbin Engineering University in 1999 and the M.S. degree in mode distinguishing and brainpower system from Harbin Engineering University in 2002 and the Ph.D. degree in control theory and control engineering from Harbin Engineering University in 2005.

From 2005 to 2006, he worked in Heilongjiang University as a lecturer. Since 2007 he has been a vice professor in Northeast Agricultural University. In 2010 he made postdoctoral study in Harbin Institute of Technology.

Dr. Deng is working in image watermarking and information security.

Zeng Qingshuang was born 1965 in Harbin China. He received the B.S. degree in automatic control from Harbin Institute of Technology in 1987, the M.S. degree in control theory and applications from Harbin Institute of Technology in 1990 and the Ph.D. degree in navigation, guidance and control from Harbin Institute of Technology in 1997.

Dr. Zeng is working in Navigation and Control Theory as a professor in Harbin Institute of Technology.

Zhou Xiuli was born 1966 in Hegang China. He received the B.S. degree in Electrical Engineering from Northeast Agricultural University in 1986, the M.S. degree in control theory and control engineering from Northeast Agricultural University in 1992. Dr. Zeng is working in Information Processing and Control Theory as a professor in Northeast Agricultural University.

Multiple Model Predictive Control of Component Content in Rare Earth Extraction Process

Hui Yang

School of Electrical and Electronic Engineering, East China Jiaotong University, Nanchang, 330013, China
Email: yhshuo@263.net

Rongxiu Lu

School of Electrical and Electronic Engineering, East China Jiaotong University, Nanchang, 330013, China
Email: rxlu@4y.com.cn

Kunpeng Zhang

School of Electrical and Electronic Engineering, East China Jiaotong University, Nanchang, 330013, China
Email: ecjtu.zhangkunpeng@163.com

Xin Wang

School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University, 200240, China
Email: wangxin26@sjtu.edu.cn

Abstract—Aiming at the complicated characteristic of rare earth extraction process and combining the material balance model, a multiple models modeling and control method is proposed. Based on the data collected in an industrial field, an improved subtractive clustering algorithm is employed to obtain steady operation points for the process; the recursive least squares algorithm is adopted to identify submodel parameters and establish multiple linear models. According to the model switching index, an online optimal predictive model is obtained. And the efficiency of the model is verified by taking a certain rare earth company extraction as an example. In the end, generalized predictive controller of the corresponding sub-model is designed, so that component content is controlled in real time and accurately. Simulation results show the effectiveness of the method above.

Index Terms—rare earth extraction process, complicated characteristic, multiple models, generalized predictive control

I. INTRODUCTION

There are many controlled parameters in rare earth extraction process, among which purity of rare earth is the foremost. It's indispensable to measure the component content. However, the extraction process has complexities like strongly nonlinear, multivariable, strong decoupling, delay and time variance between component content and solvent flow-rate, material liquid flow-rate and hydrochloric acid flow-rate. Component content also varies with the disturbances of solvent saponification degree and feed-in compositions and so on^[1]. And it's hard to be measured online and get optimal control with conventional modeling and control methods.

Manuscript received Sep. 8, 2011; revised. Oct. 20, 2011; accepted Nov. 5, 2011.

The present chief methods for component content online measurement in rare earth extraction process include UV-VIS, FIA, LaF₃ ISE, Isotopic XRF and so on. Because of high cost of the equipments, low reliability and stability, their usage in industry are generally limited. The soft sensor technology has many advantages such as preciseness, reliability, economy dynamic fast response, easy to realize the preset control in the outcome product purity and so on. The soft sensor method provides a new way to the online measurement of component content in rare earth extraction process^[2]. In[3], a soft sensor method was proposed for the measurement of component content with hybrid models. However, the dynamic characteristics of rare earth extraction process can't be fully reflected for limitation of static analysis. With the development of neural network, it has been widely applied in the modeling of nonlinear system. In[4], a modeling method using neural network was proposed for rare earth extraction process, and with the limitation of data sampled, all working conditions could not be covered by training results with this method, which resulted in low predictive accuracy and poor training effect. To solve the problem above, in[2], an intelligent optimal control strategy was provided by combining the technologies based on soft sensor and CBR(case-based reasoning). But the strategies are still based on neural network.

The multiple model approach is an efficient and simple framework for identifying and modeling of complex nonlinear process^[5]. In[6], a soft sensor method with multiple model was proposed. It established a model set with 5 extraction stages to measure component content online, but it would lead to redundant model, which not only increased the number of models, caused large calculation, but also reduced the precision of the model without optimal control.

In industrial applications, model predictive control(MPC), an optimization model-based controller, has achieved great successes, and most of commercially available MPC products have utilized linear model^[7]. Nevertheless, industrial processes are nonlinear and operate over a broad range of operating conditions. On the other hand, an important advantage of multiple model approach is that existing analysis and synthesis tools for linear systems can easily be adapted to this class of models at the cost of very little modification. Therefore, many efforts were put in development and application of multiple model/controller solutions within the MPC field. It has been reported in literatures that the selection of proper number of models have shown to be an important issue^[7, 8, 9].

This work is mainly focused on the difficulty of modeling and control about component content in rare earth extraction process. Multiple linear models with only one extraction stage at monitoring point are established to reduce the number of models and the quantity of calculation, and generalized predictive controller has been developed by using multiple models running in series to cope with whole component content varies. Firstly, by using the improved subtractive clustering algorithm, the steady operation points are obtained. Secondly, recursive least squares algorithm is adopted to identify model parameters by using the data in an industrial field and the multiple linear models are built. Then model switching index is designed to choose the optimal model and the efficiency is verified. Finally, when the predictive value of component content doesn't satisfy the requirement, the outputs of controllers are adjusted so as to ensure quality of the outlet product. The good accuracy performance obtained with the designed soft sensors and controllers shows that the effectiveness of the proposed method in modeling and control for rare earth extraction process. In addition, the problem of large time-delay in the process is also solved.

II. DESCRIPTION OF RARE EARTH EXTRACTION PROCESS

Fractional extraction processes are generally adopted in industry for the separation of rare earth, because two kinds of high purity, high recovery rate products can be obtained at the same time for the separation of *A*, *B* component, where *A* is easy extracted component and *B* is hard extraction component, the two component *A* and *B* extraction procedure is shown in Fig.1.

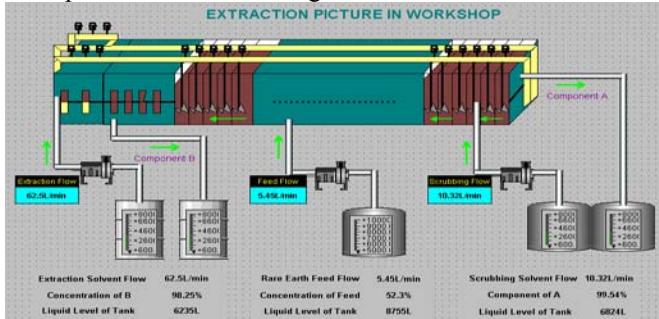


Fig. 1 Picture of rare earth countercurrent extraction process

As shown in Fig.1, the left side is extraction sections composed of *n* stage mix-clarifiers and the right side is

the scrubbing sections composed of *m* stage mix-clarifiers. In each stage, there is an agitator in each mixer and a flowmeter in tank respectively.

In Fig.1, the extraction solvent flow rate u_1 is added into the 1^{st} stage mix-clarifier, which is flowing from left to right through the agitator in mixer. Simultaneously the rare earth feed flow rare u_2 containing the element to be extracted is added into the n^{th} stage mix-clarifier, which is flowing from right to left. At the same time, the flow rate of scrubbing solvent u_3 is added into the extraction process at the $(n+m)^{\text{th}}$ stage, flowing from right to left and joining the rare earth feed at the n^{th} stage. Then, in extraction section, because the istristribution ratio is different between organic phase and aqueous phase among elements, more easily extraction component *A* and less hard extraction component *B* can be obtained and entered into the organic phase. So, in scrubbing section, through controlling the scrubbing condition, easily extraction component *A* can be obtained far more than hardly extraction component *B*, which means *A* and *B* can be separated well. Finally, repeating the exchanging and scrubbing in each stage, product *B* with the purity of Y_B can be obtained at the aqueous phase outlet in extraction section, while product *A* with the purity of Y_A at the organic phase outlet in scrubbing section.

The simplified schematic diagram of rare earth extraction process is shown in Fig.2, where u_1 is the flow rate of extraction solvent, u_2 is the flow rate of rare earth feed, u_3 is the flow rate of scrubbing solvent, $x_{i,F}$ is the concentration of rare earth feed, $z_{i,j}$, $x_{i,j}$ ($j = 1, \dots, n+m$) are the concentration of the i^{th} component at organic phase and aqueous phase respectively. Y_A is the purity of *A* and Y_B is the purity of *B*. Similarly, $Y_{A,k}$ is the organic phase content *A* at monitoring point in scrubbing section, and $Y_{B,k}$ is the aqueous phase content *B* at monitoring point in extraction section.

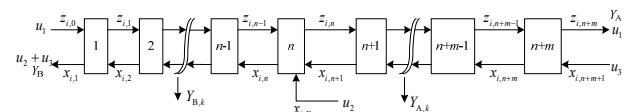


Fig.2 Schematic diagram of rare earth countercurrent extraction process

In order to improve the purity of the extraction product, dozens even hundreds of stages are built in the extraction process. So, it often takes a long-time delay(from several hours to twenty hours) for the control variables, such as rare earth feed, extraction solvent and scrubbing solvent, to adjust product purity. Therefore, process monitoring points have to be set near the outlet of the stage from 5th to 25th stage. Then, at each monitoring point, $Y_{A,k}$ and $Y_{B,k}$ should be measured and controlled to obtain the satisfactory products, i.e. Y_A and Y_B . Unfortunately, in practice $Y_{A,k}$ and $Y_{B,k}$ can not implement the real time measurement. To solve these problem above, a multiple model method is proposed.

III. MODELING OF RARE EARTH EXTRACTION PROCESS

A. An Improved Subtractive Clustering Algorithm

Rare earth extraction is a nonlinear process. So it's most important to choose the best operation points and establish the local linear model to approximate the output of nonlinear system above. Clustering is a kind of method which divides a sample set without any label into several subsets by the rules. It makes similar sample belong to the same group and different sample belong to different groups. In[10], a sort of subtractive clustering algorithm was introduced, which need not give number of clustering in advance, and can put up unsupervised learning with low calculation and fast clustering. Compared with Fuzzy C-Means(FCM) clustering algorithm, the algorithm in[10] can avoid bad clustering result and local optimum caused by the unsuitable initial parameters(cluster number and cluster center) in FCM. However, it can produce redundant cluster centers. By analyzing this problem in the algorithm, it has been found that data density satisfying $D_j^c \leq 0$ is one of the important reasons to lead such question. To solve this problem, an improved algorithm which can adjust data density to satisfy $D_j^c \leq \xi$ (ξ is a positive constant) is presented. Compared with the original method, the improved clustering algorithm above can avoid redundant cluster centers and reduce calculation. Based on the improved subtractive clustering algorithm, cluster centers or steady operations of rare earth extraction process can be obtained.

In multiple model modelling process, the main problem is to determine the number of submodel, ie. the cluster number n . Based on describing all working conditions, the cluster number n should be designed as small as possible. Smaller cluster numbers mean reduction of calculation and system stability when switching among submodels occurs frequently. So, the cluster number n has the following index function:

$$J_m = \sum_{i=1}^N \sum_{j=1}^n \mu_{ij}^2 \|X_i - X_j^c\|^2 \quad (1)$$

where N is the number of sampled data, n is the cluster numbers, $X_i \in R^q$ is the i^{th} sampled data input with $X_i = [u_{1i}, u_{2i}, u_{3i}, x_{i,F}]^T$, u_{1i}, u_{2i}, u_{3i} are the i^{th} flow rate of extraction solvent, rare earth feed and scrubbing solvent at the aqueous phase respectively, $x_{i,F}$ is the concentration of rare earth feed corresponding to the i^{th} flow rate. Clearly, rare earth extraction process is a four inputs system. In addition, $X_j^c \in R^q$ is the j^{th} cluster center with

$X_j^c = [u_{1j}^c, u_{2j}^c, u_{3j}^c, x_{i,F}^c]^T$, μ_{ij} is the membership of the i^{th} sampled data at the j^{th} clustering. Like the definition of membership function at fuzzy logic, μ_{ij} is defined as:

$$\mu_{ij} = \frac{\exp\left(-\frac{1}{2}\|X_i - X_j^c\|^2/\sigma^2\right)}{\sum_{k=1}^n \exp\left(-\frac{1}{2}\|X_i - X_k^c\|^2/\sigma^2\right)} \quad (2)$$

In this paper, the improved subtractive clustering algorithm above was used to classify sampled data and clustering effect was evaluated by index function J_m :

Step 1: Set initial parameter $\delta_a = \delta_{\min}$ and choose step value($\varepsilon > 0$);

Step 2: Calculate sampled data density:

$$D_i = \sum_{i=1}^n \exp\left(-\|X_i - X_j\|^2/(\delta_a/2)^2\right) \quad (3)$$

Step 3: Set the maximum density $D_1^c = \max D_i$ and choose the first cluster center $X_1^c = X_i |_{\max D_i}$;

Step 4: Choose $\delta_b = 1.5\delta_a$ and update data density:

$$D_i^c = D_i - D_1^c \exp\left(-\|X_i - X_1^c\|^2/(\delta_b/2)^2\right), \quad (i=1, \dots, N) \quad (4)$$

Step 5: Repeat step 3 and step 4 until $D_j^c \leq \xi$, and obtain the j^{th} cluster center $X_j^c (j=2, \dots, n)$;

with: n is cluster center numbers and satisfies $n < N$

Step 6: Calculate the k^{th} ($k > 1$) cluster index J_m and update $\delta_a = \delta_a + \varepsilon$, if $\delta_a \in [\delta_{\min}, \delta_{\max}]$, then repeat steps 2 to 5;

Step 7: Set $J_m = J_m^k (k=1, \dots, K)$; get cluster numbers n and cluster centers $X_l^c (l=1, \dots, n)$ to classify sampled data corresponding to J_m , and the corresponding data set Q_l is obtained by using the nearest-neighbour rule.

B. The Dynamic Model Establishing

Around the working point or the clustering center obtained above, considering the characteristics of rare earth extraction process, in every extractor it contains material balance and heat balance, etc. Here material balance can describe the main dynamic characteristics of all process. So the following concentration dynamic equilibrium relationship is constructed in the j^{th} stage at the aqueous phase of the i^{th} component:

$$\frac{dx_{i,1}}{dt} = \frac{1}{R_j} [(u_2 + u_3)x_{i,j+1}(t-\tau) + u_1 D_{i,j-1}x_{i,j-1}(t-\tau) - (u_2 + u_3)x_{i,j} - u_1 D_{i,j}x_{i,j}] \quad (5)$$

where u_1, u_2, u_3 are extraction solvent flow, rare earth feed flow at the aqueous phase and scrubbing flow at the aqueous phase respectively. When $j = 1, n+m$, $x_{i,0}(t-\tau) = x_{i,0}, x_{i,n+m+1}(t-\tau) = x_{i,n+m+1}$ are all known. When $j > n$, $u_2 = 0$. Here R_j is the holding volume, τ is the lagging time, $D_{i,j}$ is the distribution coefficient between the organic phase and the aqueous phase.

Because of the complication of the whole extraction process, the assumption, which whole extraction process is in the equilibrium state, is made. So the monitoring point is established at the l^{th} stage as in Fig.3 below.

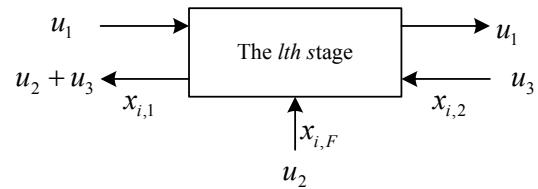


Fig.3 The diagram of component content monitoring point at the l^{th} stage

In order to detect component purity at the monitoring point, according to the concentration dynamic equilibrium relationship described in(5), the relationship of i^{th} component at the aqueous phase in the l^{th} stage is established.

$$\frac{dx_{i,1}}{dt} = \frac{1}{R_i} [u_3 x_{i,2}(t-\tau) + u_2 x_{i,F} + u_1 D_{i,0} x_{i,0}(t-\tau) - u_1 D_{i,1} x_{i,1} - (u_2 + u_3) x_{i,1}] \quad (6)$$

where $x_{i,0}(t-\tau) = x_{i,0}$, $x_{i,2}(t-\tau) = x_{i,2}$, are known.

From(6), it can be seen that it's a nonlinear equation. So, at the operation point $E = (x_{i,2}^e, x_{i,1}^e, x_{i,0}^e, u_1^e, u_2^e, u_3^e)$, it has:

$$\frac{dx_{i,1}^e}{dt} = f[x_{i,2}^e, x_{i,1}^e, x_{i,0}^e, u_1^e, u_2^e, u_3^e] = 0 \quad (7)$$

Then, transforming the operation point in (6) to origin and apply Taylor linear expansion method, the local linear model of i^{th} stage is obtained as:

$$\begin{aligned} \frac{dx_{i,1}}{dt} &= \frac{1}{R_i} [(u_2^e + u_3^e) x_{i,2}(t-\tau) + D_{i,1} u_1^e x_{i,0}(t-\tau) \\ &\quad - (D_{i,1} u_1^e + u_2^e + u_3^e) x_{i,1} + (x_{i,2}^e - x_{i,1}^e) u_2 \\ &\quad + (x_{i,2}^e - x_{i,1}^e) u_3 - (D_{i,1} x_{i,1}^e - D_{i,0} x_{i,0}^e) u_1] \end{aligned} \quad (8)$$

Discretize (8), it can be obtained as follows:

$$\mathbf{x}_i(k+1) = \mathbf{A}\mathbf{x}_i(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{D}\mathbf{v}(k) \quad (9)$$

where $\mathbf{u}(k) = [u_1(k), u_2(k), u_3(k)]^T$ is the input vector, $\mathbf{x}_i(k) = x_{i,1}(k)$ is the concentration state vector, and $\mathbf{v}(k) = x_{i,F}(k)$ is the disturbance vector. Here \mathbf{A} , \mathbf{B} and \mathbf{D} are model parameters matrices.

Then system output is

$$\mathbf{y}_i(k) = \mathbf{C}\mathbf{x}_i(k) \quad (10)$$

where \mathbf{C} is parameter matrix and $\mathbf{y}_i(k) = y_{i,l}(k)$.

Based on(9) and (10), the relationship between \mathbf{u} and \mathbf{y} is concluded as following:

$$\mathbf{y}(k+1) = \mathbf{A}\mathbf{y}(k) + \mathbf{G}\mathbf{u}(k) + \mathbf{H}\mathbf{v}(k) \quad (11)$$

where \mathbf{A} , \mathbf{G} and \mathbf{H} are all undetermined coefficient matrices.

Based on(11), it can get

$$\mathbf{y}(k+1) = \boldsymbol{\theta}^T \boldsymbol{\phi}(k) \quad (12)$$

where $\boldsymbol{\theta} = [\mathbf{A}, \mathbf{G}, \mathbf{H}]^T$, $\boldsymbol{\phi}(k) = [\mathbf{y}(k)^T, \mathbf{u}(k)^T, \mathbf{v}(k)^T]^T$.

With the system state observed, $\mathbf{y}(k)$ and $\boldsymbol{\phi}(k)$ in(12) are obtained and $\boldsymbol{\theta}$ is identified by using least-squares identification method.

Model described in(12) is only applied for the condition that is around the special operation point. When it has great changes in system running environment, identifying the parameters are so hard to track practical changes that the model is inaccurate. Through expanding the nonlinear system around multiple operation points and using local linear models established above to approach nonlinear system, dynamic model which is approximate to the practical system can be obtained. Therefore, rare earth extraction process is expanded around different operation points

and the system is described with n input-output models I_1, I_2, \dots, I_n . The discrete local linear model around each operation point is

$$I_l : \mathbf{y}^l(k+1) = \boldsymbol{\theta}_l^T \boldsymbol{\phi}(k), l = 1, \dots, n \quad (13)$$

where $\boldsymbol{\theta}_l = [\mathbf{A}^l, \mathbf{G}^l, \mathbf{H}^l]^T$ is parameter matrix around the l^{th} operation point, and $\boldsymbol{\phi}(k)$ is a matrix composed of input-output data.

According to (13), and combined with the rare earth countercurrent extraction process described in Fig.3, the relationship between component purity y at monitoring point and input u is

$$\begin{aligned} I_l : y(k+1) &= -a_1^l y(k) - a_2^l y(k-1) + g_1^l u_1(k) \\ &\quad + g_2^l u_2(k) + g_3^l u_3(k) + h_1^l x_F(k) \\ &= \boldsymbol{\theta}_l^T \boldsymbol{\phi}(k) \quad (l = 1, \dots, n) \end{aligned} \quad (14)$$

where $\boldsymbol{\theta}_l^T = [a_1^l, a_2^l, g_1^l, g_2^l, g_3^l, h_1^l]^T$ is a parameter vector of the l^{th} model and $\boldsymbol{\phi}(k) = [-y(k), -y(k-1), u_1(k), u_2(k), u_3(k), x_{i,F}(k)]^T$.

So, using improved subtractive clustering algorithm above, cluster center X_l^c is obtained as operation point of local linear model, which is steady operation point in practical system. The initial parameters $\boldsymbol{\theta}_l^o$ of local are identified with the cluster data set Ω_l and recursive least-square identification method.

C. Model Switching

Modeling switching is a kind of scheduling mechanism in multiple models modeling method. The switching index is chosen according to the factors like practical physiacl object, control accuracy, etc.

Considering the rare earth countercurrent extraction process introduced in this paper, model structure adopts the form of multiple models in(14). Each local linear model has the same structure but the different initial parameters. At every sampling time, according to the switching index, only one local model I_l is selected out to be the optimal model to approximate to the system. To the switching index, different forms have great influence on modeling accuracy and model switching times. In this paper, the switching index function is designed according to the accumulation of the identified error among local models, which has integral property, as

$$J_l(k) = \sum_{j=1}^k \beta(j)^{(k-j)} |y(k) - \hat{y}_l(k)| \quad (15)$$

where $y(k)$ is the real output, $y_l(k)$ is the l^{th} model output, and $0 < \beta < 1$ is the weighted factor.

The switching index in (15) evaluates the matching degree between each submodel and the system by comparing with current error and historical error. It means the less switching index value is, the more matching degree is. Therefore, the multiple models based on local linear model can be represented as

$$\hat{y}_l(k+1) = \sum_{i=1}^n \alpha(J_l) \hat{\theta}_l^T(k) \boldsymbol{\phi}(k) \quad (16)$$

with

$$\alpha(J_l) = \begin{cases} 1, & \text{if } l = \arg \min J_j, j = 1, \dots, n \\ 0, & \text{others} \end{cases}$$

When there are too many submodels, the above model established exists an obvious problem. The matching degree between each submodel and the real system in sequence will lead to huge calculation, which influences the real-time monitoring/controlling. So, the multiple models set must be optimized. Firstly, the established multiple models set is divided into two regions by usage frequency, some of which with higher usage frequency belongs to sensitive region, others are non-sensitive region. To the models in sensitive region, their outputs are computed each time. To the models in non-sensitive region, when the error of the models in sensitive region is more than the threshold designed before, their outputs will be computed. Compared with the past multiple model set, it need not calculate each submodel output in sequence, which causes the calculation reduced.

D. Verification of Model

In order to verify the effectiveness of the proposed online prediction method about rare earth extraction component content based on multiple linear models, the real industrial data from a company extracted product yttrium is used to test the model. Based on the model, the element component purity at monitoring point in sensitive stage is measured online. If its purity doesn't meet requirement, by adjusting the flow rate of extraction solvent, rare earth feed and scrubbing solvent in time, the purity at monitoring point is controlled in a reasonable region and the quality of outlet product is guaranteed.

In this paper, 150 groups of real industrial effective data are obtained at monitoring point. Firstly, through the improved subtractive clustering algorithm above, 100 groups of sampled data are devided to four groups and the cluster centers are regarded as operation points as shown in TABLE I. Secondly, in each group, the local linear model is identified off-line by using the corresponding data. Then four models of element component content are obtained, which are shown as model 1~model 4. At last, another 50 groups of ampled data are adopted to test the models. According to the introduced modeling method above, choosing(15) as switching index function, when $\beta = 0.55$, doing simulation with MATLAB, the last result is shown in Fig.4 and Fig.5, respectively.

Compared with the simulated results and defined the following errors: maximum positive error(MPE), maximum negative error(MNE), robust mean square error(RMS E), the errors in two stages of model fitting and model testing are analyzed. The result is shown in TABLE II.

From Fig.4 and TABLE II, 100 groups of data are used for the identification of initial models, and the model output error is larger because the model is optimized constantly. Then the other 50 groups of data are used for testing the last optimal model, result shows that the output error is less than before. From the switching curve in Fig.5, it can be seen that condition changes mainly at point 2 and point 3, so the multiple linear models are switched between model 2 and model 3.

TABLE I.
STEADY OPERATION POINTS

<i>Input Variables</i> (u_1, u_2, u_3)	<i>Feed Concentration</i> (x_F)
Point 1	(73.88, 6.82, 11.30)
Point 2	(26.1, 2.50, 3.88)
Point 3	(40.30, 3.50, 6.21)
Point 4	(55.47, 5.40, 8.49)

Model 1 ~ Model 4:

$$\begin{aligned} y(k+1) &= 0.0707 y(k) - 0.1067 y(k-1) + 0.0509 u_1(k) \\ &\quad - 0.1507 u_2(k) - 0.2620 u_3(k) + 1.1947 x_F(k) \\ y(k+1) &= 0.2649 y(k) - 0.2509 y(k-1) + 0.1102 u_1(k) \\ &\quad - 0.2307 u_2(k) - 0.5622 u_3(k) + 1.2366 x_F(k) \\ y(k+1) &= 0.4448 y(k) - 0.0124 y(k-1) + 0.1111 u_1(k) \\ &\quad - 0.2660 u_2(k) - 0.5533 u_3(k) + 0.7552 x_F(k) \\ y(k+1) &= 0.1372 y(k) + 0.0648 y(k-1) + 0.0676 u_1(k) \\ &\quad - 0.1592 u_2(k) - 0.3404 u_3(k) + 1.1038 x_F(k) \end{aligned}$$

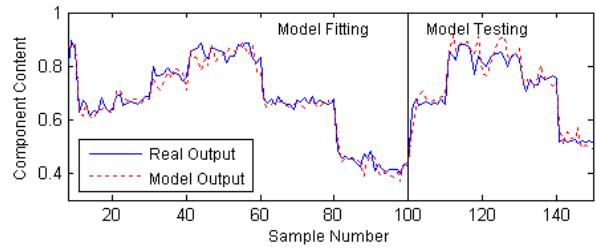


Fig.4 Curves between multiple linear models and real output

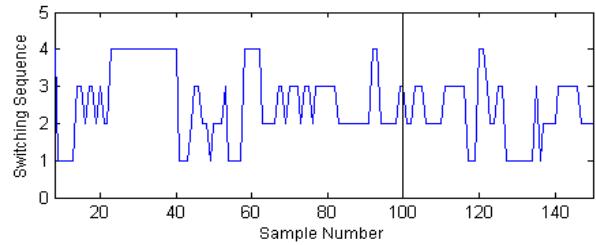


Fig.5 Switching curve of multiple linear models

TABLE II.
MODEL ERROR PERFORMANCE ANALYSIS

	MPE	MNE	RMSE
Model fitting	0.1011	0.0639	0.0281
Model testing	0.0397	0.0785	0.0299

IV. MULTIPLE MODELS PREDICTIVE CONTROLLER

According to the description in III, the online estimation and real-time monitoring of element component content are realized with the model obtained. When the component content at monitoring point doesn't meet objective value, based on the multiple linear models above and combined with the characteristics of Generalized Predictive Control(CPC)algorithm using CARIMA(Controlled Aut o-aggressive Integral Moving Average) model, a Multiple

Models GPC(MMGPC) is designed to adjust control variables to satisfy the requirement fastly. Depending on its historical dynamic information, a MMGPC can predict the control variable, which solves the problem of large time-delay in rare earth extraction process to some extent.

Model 1 ~ Model 4 are optimized using the sample data which are selected in an industry field. Then, taking controlling scrubbing solvent u_3 as an example, GPC C3 is designed as following.

C3:

$$\begin{cases} y(k) - 0.8519y(k-1) - 0.1633y(k-2) + 0.0152y(k-3) \\ \quad = -0.2887\Delta u_3(k-1) \\ y(k) - 1.6020y(k-1) + 0.6473y(k-2) - 0.0453y(k-3) \\ \quad = -0.4493\Delta u_3(k-1) \\ y(k) - 1.2033y(k-1) + 0.4501y(k-2) - 0.2468y(k-3) \\ \quad = -0.5339\Delta u_3(k-1) \\ y(k) - 1.1345y(k-1) + 0.0644y(k-2) + 0.0701y(k-3) \\ \quad = -0.3621\Delta u_3(k-1) \end{cases}$$

The other general predictive controller C1 is designed as the same method.

V. SIMULATIONS

In a certain rare earth company, product yttrium should be extracted with the purity reached more than 0.99. Judged by the experiences, the component content at monitoring point should be up to 0.8, otherwise, the outlet product can't meet the requirement. At some time, u_1 is 51.42 L, u_2 is 4.5 L, u_3 is 8.11 L and x_F is 0.481 mol/L. The output y obtained by the multiple linear models established above is 0.65 and according to the clustering algorithm submodel 3 is used to test component content online at this time. So using the corresponded controller to adjust the flow rate of scrubbing solvent u_3 , the component content at monitoring point is up to the objective value 0.8 fast. Finally, compared with PID algorithm and simulated with MATLAB, results are shown in Fig.6 and Fig.7, respectively.

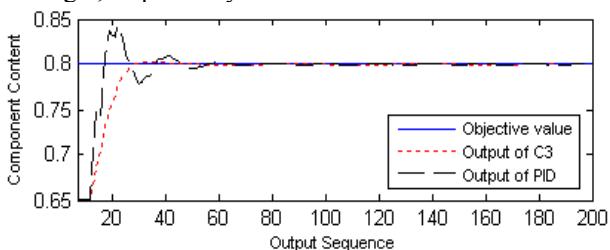


Fig.6 Comparison curves of tracking objective value

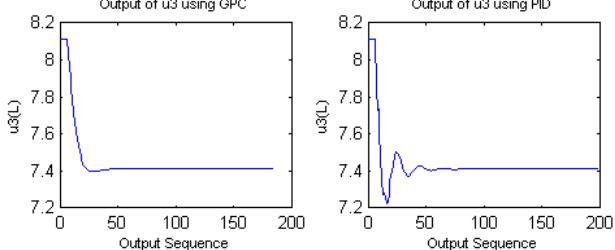


Fig.7 Varying curves of control variables as tuning u_3

From Fig.6, compared with PID algorithm, it shows fast response speed and stability using GPC. And from

Fig.7, the changes of u_3 using PID are unstable and oscillating, all of which are not suitable in practice.

VI. CONCLUSION

In this paper, considering the nonlinear dynamic model of rare earth extraction process, the multiple model modeling and control method is designed. Using the improved subtractive clustering algorithm, the extraction process is described by nonredundant multiple linear models with the same model structure and different parameters. Simultaneously, by using of switching index, the optimal model for measuring the component content online is constructed. Compared with the real industrial data, it shows that the modeling method has better generalization ability and higher prediction accuracy. Then, taking controlling the flow rate of scrubbing solvent as an example, GPC controller is designed in order to realize predictive control of component content in rare earth extraction process. Results show that the multiple model predictive control proposed is effective.

ACKNOWLEDGMENT

The authors acknowledged the National Natural Science Foundation of China(60864004, 51174091, 61164013) and the National 863 Plans Projects(2008AA04Z129) for having provided the financial support for this work.

REFERENCES

- [1] Tianyou Chai, Hui Yang, "Situation and developing trend of rare-earth countercurrent extraction process control", Journal of Rare Earth, vol.22, No.5, 2004, pp604-610.
- [2] Hui Yang, Chunyan Yang, Chonghui Song, Tianyou Chai, "Intelligent optimal control in rare-earth countercurrent extraction process vis soft-sensor", Lecture Notes in Computer Science, vol.3611, 2005, pp214-223.
- [3] Hui Yang, Xin Wang, "Component content soft-sensor based on hybrid models in countercurrent rare earth extraction process", Journal of Rare Earth, vol.23, Suppl, 2005, pp86-91.
- [4] Hui Yang, Tianyou Chai, "Component content soft-sensor based on Neural Network in countercurrent rare-earth extraction process", Acta Automatica Sinica, vol.32, No.4, 2006, pp489-495.
- [5] E.Domlan, B.Huang, F.Xu, A.Espejo, "A decoupled multiple model approach for soft sensors design", Control Engineering Practice, vol.19, N0.2, 2011, pp126-134.
- [6] Wenjun Jia, Tianyou Chai, "Soft-sensor of element component content based on multiple models for the rare earth cascade extraction process", Control Theory&Applications, vol.24, No.4, 2007, pp569-573. (in Chinese)
- [7] K.Konakom, P.Kittisupakom, I. Mujtaba, "Batch control improvement by model predictive control based on multiple reduced-models", Chemical Engineering Journal, vol. 145, 2008, pp129-134.
- [8] P.Overloop, S.Weijs, S.Dijkstra, "Multiple model predictive control on a drainage canal system", Control Engineering Practice, vol.16, 2008, pp531-540.
- [9] C.R.Pofirio, E.Almeida Neto, D.Odloak, "Multi-model predictive control of an industrial C3/C4 splitter", Control Engineering Practise, vol.11, 2009, pp765-779.

- [10] S.Chiu, "Fuzzy model identification based on cluster estimation", Journal of Intelligent and Fuzzy Systems", vol.2, No.3, 1994, pp267-278.



Hui Yang JiangXi Province, China. Birthdate: March, 1965. is Control Theory and Control Engineering Ph.D., graduated from Northeastern University. And research interests on complex system modeling, control and optimization,process industry integrated automation technology and application.

He is a professor and supervisor of doctoral students at school of Electrical and Electronic Engineering of East China Jiaotong University.



KunPeng Zhang HeNan Province, China. Birthdate: June, 1986. is a graduate student at the East China Jiaotong University. And research interests on complex system modeling, control and optimization.



RongXiu Lu GuangXi Province, China. Birthdate: December, 1976. is a Ph.D. student at the Nanchang University. And research interests on complex system modeling, control and optimization.

She is a lecturer in School of Electrical and Electronic Engineering of East China Jiaotong University.



Xin Wang LiaoNing Province, China. Birthdate: May, 1972. is Control Theory and Control Engineering Ph.D., graduated from Northeastern University. And research interests on intelligent decoupling control, multiple model adaptive control, complex system modeling, control and optimization.

He is a associate professor and master tutor at school of Electronic Information and Electrical Engineering of Shanghai Jiaotong University.

Multi-party Dialogue Games for Dialectical Argumentation

Jinping YUAN, Li YAO, Zhiyong HAO, Yanjuan WANG
 Science and Technology on Information Systems Engineering Laboratory,
 National University of Defense Technology, Changsha 410073, China
 Email: {yuanjp, liyao, zyhao, wjwang}@nudt.edu.cn

Abstract—This paper concerns a distributed argumentation system where different agents are equipped with argumentative knowledge base (henceforth referred as KB) within which conflict arguments are represented using attacking relations. This paper proposes the notion of “defensibility” of an argument in a distributed argumentation system and a multi-party dialogue game to compute the defensibility of an argument. In our multi-party dialogue game framework, we have proposed the notion of critical factor, legal move function and critical countermeasure, which act as the mechanism for avoiding idle attack and invalid attack in the course of dialogue games. Theoretically, the paper has also proved the soundness and completeness of multi-party dialogue games conducted by legal move and countermeasure function. It is anticipated that this research will contribute to argumentation research in MAS.

Index Terms—argumentation; distributed argumentation system; multi-party dialogue games; multi-agent system

I. INTRODUCTION

Belief conflicts are inevitable among different agents inhabiting in environments with imperfect information (such as incomplete, inconsistent or imprecise information). Finding effective ways for agents to reconciling conflicts is an active research area in multi-agent system (MAS). One mechanism of reducing conflicts is through argumentation where each agent presents arguments for or against the initial thesis and tries to convince each other during the argumentation or dialogical process^[1].

Much of recent work on argumentation dialogues^[2-4], are usually based on static and centralized argumentation framework^[5], which means that arguments and their attacking relationships are prescribed together in advance and unchangeable during the argumentation process. As our world is dynamic and distributed, this assumption doesn't normally hold and it is difficult to establish a complete argumentation framework in advance. In a dynamic situation, argumentation framework w.r.t. a certain proposition is gradually emerged during the argument process. It is therefore difficult to know whether an admissible set is preferred extension.

In this paper, we consider a distributed argumentation system, where arguments are stored in different agents' knowledge base as fragments of information and the

attacking relationship exists between arguments with conflict information (e.g. “Chris is a female” and “Chris is a male”). To resolve conflicts among agents, we adopt an argumentation approach to enable agents engaging in dialogue with each other. Typically, we propose a multi-party dialogue game to facilitate such dialogue interaction. We introduce the notions of defensibility for arguments, and a multi-party dialogue game for computing defensibility of an argument. The multi-party dialogue game is established on the notions of critical factor, eligible legal move function and critical countermeasure.

The remainder of this paper is organized as follows. Distributed argumentation system is briefly introduced in Section 2 where the notation of defensibility of an argument is proposed. We then introduce our proposed multi-party dialogue games in Section 3 where some key notations, e.g. critical factor, legal move function and critical countermeasure are introduced. Section 4 contains an algorithm we have proposed for computing critical factors and the defense set of a defensible proposition in a multi-party dialogue. Section 5 provides an example illustrating our proposed approach in action. We finally discuss some related work in the area and our planned further work.

II. DISTRIBUTED ARGUMENTATION SYSTEM

Over the last decade, a number of argumentation systems have been proposed to formally represent argumentation^[5-7]. The most influential one is Dung's argumentation framework^[5], which defines a set of atomic arguments together with a set of attack relationship. Generally speaking, these frameworks are based on the assumption that all arguments and attack relationship are prefixed. Following Dung's abstract argumentation framework, we propose a distributed argumentation system as follows.

Distributed argumentation system is established on several knowledge bases of participating agents (henceforth referred as participants). During the process of argumentation, participants are able to generate the initial arguments for a given topic, advance arguments for or against a position with the aim of convincing each other (including users) to adopt a certain proposition. The distributed argumentation system is defined as follows.

Definition 1. (Consistency) Let a set KB of program clauses, if there does not exist literal φ satisfies both

$KB \models \phi$ and $KB \models \square \phi$, KB is said to be consistent.

Here, it is notable that ϕ and $\square \phi$ are two conflicting literal, such as “Chris is a female” and “Chris is a male”.

Let a set $PAR = \{par_i | i \in \mathbb{I}\}$ of participants, the knowledge base KB_i of each participant $par_i \in PAR$ is a finite set of logical sentences. It is assumed that each agent's knowledge base is consistent, but conflicts may exist among different agents' knowledge bases, i.e. $\exists \phi. (KB_i \models \phi) \wedge (KB_j \models \square \phi), i \neq j$.

Definition 2. (Distributed argumentation system) A distributed argumentation system established on knowledge bases of all participants $PAR = \{par_i | i \in \mathbb{I}\}$ is a tuple $DAF = (A, R)$, in which:

① $A = \bigcup_{par_i \in PAR} A_i$, A_i be a set of arguments generated from the knowledge base of par_i ;

② $R = \bigcup_{i,j=1}^n R_{ij}$, R_{ij} is a set of binary (attack) relationship from A_i to A_j , i.e. $R_{ij} \subseteq A_i \times A_j$, $R_{ii} = \emptyset$.

Given two arguments $a \in A_i$ and $b \in A_j$, $(a, b) \in R$ (or aRb) means a attacks b ; $(a, b) \notin R$ represents a and b are conflict-free. R_+ba denotes the set of arguments attacked by argument a , i.e. $R_+ba = \{b \in A_j | aRb\}$; and $Rb(a)$ the set of arguments attacking argument a , i.e. $Rb(a) = \{b \in A_j | bRa\}$. A set S of arguments attacks an argument a if there is some argument b in S , such that $(b, a) \in R$. And an argument a attacks a set S of arguments if there is some argument b in S , such that $(a, b) \in R$. Let a set S of arguments, similarly, $R\#S = \bigcup_{a \in S} R\#a$ ($\omega \in \{+, -\}$) denotes the set of arguments which attacks or is attacked by S .

A set S of arguments is conflict-free iff for any arguments $a, b \in S$, it follows $(a, b) \notin R$. For a conflict-free set S of arguments, S is called admissible iff for any argument $a \in Rb(S)$, there exists an argument $b \in S$ such that $(b, a) \in R$.

With the assumption of consistency of each participant's knowledge base, it follows that, for each $par_i \in PAR$, A_i is conflict-free.

In a distributed argumentation setting, it is usually difficult to decide whether an admissible set is a preferred extension. Thus we introduce a notion of defensibility for an argument.

Definition 3. (Defensibility^[8]) An argument $a \in A$ of $DAF = (A, R)$ is a defensible w.r.t. an admissible set $S \subseteq A$, iff $\forall \alpha \in A. (xRb \rightarrow x \in R_+(\alpha))$.

If argument a is defensible w.r.t. admissible set $S \subseteq A$, then S is said to be a 's defense set.

Example 1. Let $\{a, e\}$, $\{b, d\}$ and $\{c\}$ are sets of arguments derived from KB_1 , KB_2 and KB_3 respectively, such that $R = \{(b, a), (d, a), (c, b), (c, d), (e, b)\}$. Then a is defensible on $DAF = (A, R)$ established on KB_1 , KB_2 and KB_3 , and whose defense set is $\{a, c, e\}$.

III. MULTI-PARTY DIALOGUE GAMES

Multi-party dialogue game is a suitable mechanism for the implementation of distributed argumentation. In a multi-party dialogue, each participant par_i puts forward their own proposition ϕ_i for a given topic t , then we get a collection $\Phi = \{\phi_i | par_i \in PAR\}$ of propositions. For each proposition $\phi \in \Phi$, there exists a multi-party dialogue starting by an initial move that contains ϕ , in which all participants make moves to attack or defend the initial one. The proponents win the multi-party dialogue if the opponents have no moves to change the status of the initial move, the status of the initial move is labeled as 1 when the opponents run out all valid moves to expand the multi-party dialogue, and labeled as 0 when the proponents has no valid moves to expand the dialogue tree. A formal definition of the status of an argument is given in Definition 9 below.

Traditionally, a move is tuple $m = (par_i, arg)$ ^[9], in which $par_i \in PAR$ is the player of the move, denoted by $Par(m)$; and $arg \in A_i$ the (counter-)argument expressed with move, denoted by $Arg(m)$. Meanwhile, for convenience, let $M_i = \{m | Arg(m) \in A_i\}$ be the set of moves held by participant par_i , and $M = \{m | Arg(m) \in A\}$ the set of moves in $DAF = (A, R)$, i.e. $M = \bigcup_{par_i \in PAR} M_i$. In addition, $Arg(m_i)R Arg(m_j)$ is also simplified by m_iRm_j which is called move m_i attacks m_j .

Definition 4. (Multi-party dialogue type) A multi-party dialogue type is a tuple $MDT = (PAR, DAF, \psi)$, where $PAR = PAR_p \cup PAR_o \cup PAR_n$ is the set of participants with three different positions to a certain proposition (i.e. proponents PAR_p , opponents PAR_o , and neutrals PAR_n); $DAF = (A, R)$ is a distributed argumentation system established on these participants' knowledge base; and $\psi : \mathcal{P}(M) \rightarrow \mathcal{P}(M)$ is legal move function, in which $\mathcal{P}(M)$ represents the power set of M .

For $\forall par \in PAR_p \cup PAR_o$, $C(par)$ means the complement of the set that par belongs to, e.g., for a move m if $Par(m) \in PAR_p$, then $C(Par(m)) = PAR_o \cup PAR_n$.

A multi-party dialogue is a sequence of moves $d : m_0, m_1, \dots, m_n, \dots$, in which players make moves to attack those moves made by the opposing party. It can be defined formally as follows.

Definition 5. (Multi-party dialogue) A multi-party dialogue is a sequence of moves $d : m_0, m_1, \dots, m_n, \dots$ satisfying: ① $Par(m_0) \in PAR_p$, and $Arg(m_0) \in \Phi$; and ② for any move m_n in d , such that $m_n \in \psi(\{m_0, m_1, \dots, m_{n-1}\})$, where ψ is legal move function.

The first condition says that multi-party dialogues always start with an argument in question, i.e. the proposition w.r.t. a given topic; and the second condition states all moves in dialogue games should be legal as defined in definition 6 below.

For convenience, $m \in_i d$ ($i = 0, 1, 2, \dots$) denotes move m appears in d at i th position (*i-appear* for short).

Definition 6. (Legal move function) A legal move function for multi-party dialogue games is defined as $\psi: \mathcal{P}(M) \rightarrow \mathcal{P}(M)$ that satisfies:

$$\psi(\bar{d}) = \begin{cases} M_\Phi, & \text{if } \bar{d} = \emptyset \\ \bigcup_{m \in_i d} (\mathbf{R}\underline{b}(m, M) \setminus \mathbf{R}\underline{b}(m, \bar{d})), & \text{otherwise} \end{cases}$$

where $\bar{d} = \{m \mid m \in_i d\}$ represents the set of all moves appear in d ; $M_\Phi = \{m_0 \mid \text{Arg}(m_0) \in \Phi\}$ represents the set of moves only containing the initial argument; and $\mathbf{R}\underline{b}(m, S) = \{m' \in S \mid m' R m \wedge \text{Par}(m') \in C(\text{Par}(m))\}$ the set of moves attacking m in set S , $S \subseteq \{M, \bar{d}\}$.

This definition suggests that for each move $m \in_i d$ ($i=1, 2, \dots$) (i.e. except for m_0), there exists a move attacked by m in multi-party dialogue, and it is not allowed to use one move attacking the same move more than once.

For a move m , $\mathbf{R}_+ b(m, S) = \{m' \in S \mid m R m'\}$ indicates the set of moves in S attacked by m . In multi-party dialogue d , it follows $\mathbf{R}_+ b(m_0, d) = \emptyset$.

Definition 7. (Extension) Let d and d' be two multi-party dialogues, d' be an expansion of d , iff for any $m \in_i d$, it also holds $m \in_i d'$;

Definition 8. (Legal extension) Let d' is an expansion of d , and there exists a set of moves $E = \{m \mid m \in_i d' \wedge m \notin_i d\} \neq \emptyset$. If $E \subseteq \psi(\bar{d})$, then d' be a legal expansion of d with E under ψ . E is also said a legal expansion factor of d under ψ .

Definition 9. (Move status function) Move status function $\text{label}: \bar{d} \rightarrow \{0, 1\}$, such that for any $m \in_i d$, it can be recursively defined as:

$$\text{label}(m \mid d) = \begin{cases} 0, & \text{if } \omega \geq 1 \\ 1, & \text{if } \mathbf{R}\underline{b}(m, \bar{d}) = \emptyset \text{ or } \omega = 0 \\ \text{undefined}, & \text{otherwise} \end{cases}$$

where $\omega = \sum_{m' \in \mathbf{R}\underline{b}(m, \bar{d})} \text{label}(m' \mid d)$, $\text{label}(m \mid d)$ denotes the status of m in d .

From above definition, each move in multi-party dialogue can be labeled either as 0 or 1. The definition also suggests a backward labeling procedure, through which we are able determine the status of the initial move m_0 .

Let $Win(d) = \{m \in_i d \mid \text{label}(m \mid d) = 1\}$ indicates the set of moves labeling as 1 in a multi-party dialogue d , and $Def(d) = \{m \in_i d \mid \text{label}(m \mid d) = 0\}$ the set of moves labeling 0.

Definition 6 provides a set of legal moves from the attacking relationship and repetition points of view. Further, we do not want to extend a multi-party dialogue with invalid moves (e.g. moves already been defeated by the opposing party) or useless moves (has no effect on the status of the initial move of dialogue game). In a dialogue game, each party hopes to achieve its goal with the least number of moves. Specifically, the proponents try to prove the rationality of initial argument by attacking all attacks from the opponents; while the opponents attempt to disprove the initial argument by attacking the

proponents' argument. Intuitively, some moves play critical role in a multi-party dialogue, since they have direct effect on the initial move. From this point of view, both the proponents and opponents could propose an optimal set of attackers aiming at critical moves. Critical factor depicts these moves playing a critical role in a multi-party dialogue.

Let $PRO(d) = \{m \in_i d \mid \text{Par}(m) \in PAR_p\}$ be the set of moves presented by the proponents in d , and $OPP(d) = \{m \in_i d \mid \text{Par}(m) \in PAR_o\}$ the set of moves presented by the opponents.

Definition 10. (Critical factor) Let multi-party dialogue d , and the initial move $m_0 \in_0 d$.

(1) If $\text{label}(m_0 \mid d) = 1$, $CF_{PRO}(d) \subseteq PRO(d) \cap Win(d)$ be a set of moves executed by the proponents in d , $CF_{PRO}(d)$ is called a pro critical factor iff the following conditions hold: (1) If $\sum_{m \in CF_{PRO}(d)} \text{label}(m \mid d) = 0$, then $\text{label}(m_0 \mid d) = 0$; and (2) $CF_{PRO}(d)$ is minimal w.r.t. set inclusion.

(2) If $\text{label}(m_0 \mid d) = 0$, $CF_{OPP}(d) \subseteq OPP(d) \cap Win(d)$ be a set of moves executed by the opponents in d , $CF_{OPP}(d)$ is called a con critical factor, iff the following conditions hold: (1) If $\sum_{m \in CF_{OPP}(d)} \text{label}(m \mid d) = 0$, then $\text{label}(m_0 \mid d) = 1$; and (2) $CF_{OPP}(d)$ is minimal w.r.t. set inclusion.

By Definition 10, if $\text{label}(m_0 \mid d) = 1$, the pro critical factor $CF_{PRO}(d)$ represents a minimal set of moves, such that if all moves in $CF_{PRO}(d)$ are defeated, then $\text{label}(m_0 \mid d) = 0$; and if $\text{label}(m_0 \mid d) = 0$, the con critical factor $CF_{OPP}(d)$ indicates a minimal set of moves such that if all moves in $CF_{OPP}(d)$ are defeated, then $\text{label}(m_0 \mid d) = 1$. Obviously, the pro (con) critical factor provides guidelines for the opponents (proponents) to attack proponents' (opponents') move.

Generally, there may be several critical factors in a multi-party dialogue, let $SCF_{PRO}(d)$ and $SCF_{OPP}(d)$ denote the set of pro and con critical factors respectively. Sometimes, pro critical factor and con critical factor are collectively called critical factor, denoted by $CF_{PO}(d)$ ($PO \in \{PRO, OPP\}$ from now on), which represents the minimal set of critical moves that makes $\text{label}(m_0 \mid d) = 1$ or $\text{label}(m_0 \mid d) = 0$. The element in critical factor is called critical element, it is obvious that for any critical element m , it follows $\text{label}(m \mid d) = 1$.

Example 2. Figure 1 depicts a multi-party dialogue. At the beginning, a proponent makes the initial move (m_0) , following Definition 9, it holds that $d_{(a)}$ contains a single move (as Fig.1 (a)) and $\text{label}(m_0 \mid d_{(a)}) = 1$. Then, the only pro critical factor is $\{m_0\}$, which means that only when m_0 is attacked will its status changes. Suppose the opponents makes moves m_1 and m_2 attacking m_0 (as Fig.1 (b)), it is obvious that $d_{(b)}: m_0, m_1, m_2$ and $\text{label}(m_0 \mid d_{(b)}) = 0$. For the proponents, they have to

counterattack both m_1 and m_2 in order to achieve $\text{label}(m_0) = 1$, as the con critical factor is $\{m_1, m_2\}$. Suppose the proponents make moves $\{m_3, m_4\}$ attacking critical factor $\{m_1, m_2\}$ (as Fig.1 (c)), then it is easy to know that $d_{(c)} : m_0, m_1, m_2, m_3, m_4$ and $\text{label}(m_0 | d_{(c)}) = 1$ and $SCF_{PRO}(d_{(c)}) = \{\{m_0\}, \{m_3\}, \{m_4\}\}$ is the set of pro critical factors, and so on.

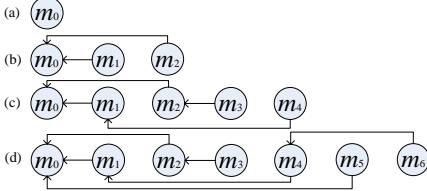


Figure 1. Multi-party dialogues

Proposition 1. Let a multi-party dialogue d , and the initial move $m_0 \in_0 d$.

- 1) If $\text{label}(m_0 | d) = 1$, then $\{m_0\}$ is a pro critical factor;
- 2) If $\text{label}(m_0 | d) = 0$, then $R\bar{b}(m_0, \bar{d}) \cap Win(d)$ is a con critical factor.

Definition 11. (Critical dependence) Let a multi-party dialogue d , and a critical element $m \in PO(d)$, such that $R\bar{b}(m, \bar{d}) \neq \emptyset$. The move m is critically dependent on $CD \subseteq PO(d) \cap Win(d)$, iff following conditions hold:
① If $\sum_{m' \in CD} \text{label}(m' | d) = 0$, then $\text{label}(m | d) = 0$;
② CD is minimal w.r.t. set inclusion.

The set of moves CD is also called a critical dependence of move m , denoted by $m \sim CD$. Specially, it is said that m has no critical dependence if $R\bar{b}(m, \bar{d}) = \emptyset$.

Theorem 1. Let $m \in_i d$ is a critical element in critical factor $CF_{PO}(d)$ of d , and $R\bar{b}(m, \bar{d}) \neq \emptyset$. If m is critically dependent on CD , then $(CF_{PO}(d) \setminus \{m\}) \cup CD$ is also a critical factor of d .

Proof. For convenience, and without loss of generality, let $PO = PRO$, i.e. $CF_{PRO}(d)$ be a pro critical factor in d .

For critical element $m \in_i d$ in $CF_{PRO}(d)$, let $CF_{PRO}^+(d) = (CF_{PRO}(d) \setminus \{m\}) \cup CD$ where CD is a critical dependence of m . By Definition 11, it holds that if $\sum_{m' \in CF_{PRO}^+(d)} \text{label}(m' | d) = 0$, then $\text{label}(m_0 | d) = 0$. Because m is a critical element in $CF_{PRO}(d)$ and critically dependent on CD (i.e. $m \sim CD$).

Then it remains to show that $CF_{PRO}^+(d)$ is minimal. Suppose $CF_{PRO}^+(d)$ is not minimal, thus there exist $CF_{PRO}^*(d) \supsetneq CF_{PRO}^+(d)$, such that if $\sum_{m' \in CF_{PRO}^+(d)} \text{label}(m' | d) = 0$, then $\text{label}(m_0 | d) = 0$. In other words, there exists at least one move m^* which holds $m^* \in CF_{PRO}^+(d)$ but $m^* \notin CF_{PRO}^*(d)$. There are two cases:

- 1) m^* be a critical element in $CF_{PRO}(d)$ (i.e. $m^* \in CF_{PRO}(d)$), then it follows that there exists a set

$CF_{PRO}^-(d) = CF_{PRO}(d) \setminus \{m^*\}$ of moves, such that if $\sum_{m' \in CF_{PRO}^-(d)} \text{label}(m' | d) = 0$, then $\text{label}(m_0 | d) = 0$. This violates the minimality of critical factor $CF_{PRO}(d)$.

2) m^* be an element in CD (i.e. $m^* \in CD$), then it holds that there exists a set $CD^- = CD \setminus \{m^*\}$ satisfying if $\sum_{m' \in CD^-} \text{label}(m' | d) = 0$, then $\text{label}(m | d) = 0$. It violates the minimality of critical dependence CD . Therefore $CF_{PRO}^+(d)$ is minimal w.r.t. set inclusion.

Thus, $CF_{PRO}^+(d) = (CF_{PRO}(d) \setminus \{m\}) \cup CD$ is a critical factor of d . \square

Definition 11 states the substitutability of critical element, in other words, for any critical factor $CF_{PRO}(d)$, a new critical factor can be established from $CF_{PRO}(d)$ by substituting a critical element with one of its critical dependences. And Theorem 1 suggests an approach to computing the set of critical factors (it remains to details later).

Proposition 2. Let a multi-party dialogue d , if $\text{label}(m_0 | d) = 1$, then m_0 critically depends on all pro critical factors.

Theorem 2. For each critical element $m \in_i d$, if $R\bar{b}(m, \bar{d}) \neq \emptyset$, then $m \cong b(m', \bar{d}) \cap Win(d)$, where $m' \in R\bar{b}(m, \bar{d})$.

Proof. For convenience, and without loss of generality, we assume that $m \in CF_{PRO}(d)$, then it follows that $m \in PRO(d) \cap Win(d)$, i.e. $\text{label}(m | d) = 1$. By Definition 9, it follows $\sum_{m' \in R\bar{b}(m, \bar{d})} \text{label}(m' | d) = 0$, because $R\bar{b}(m, \bar{d}) \neq \emptyset$, i.e. $R\bar{b}(m, d) \subseteq OPP(d) \cap Def(d)$. Hence, for any $m' \in R\bar{b}(m, \bar{d})$, it easy to see that $R\bar{b}(m', \bar{d}) \neq \emptyset$ and $\sum_{m'' \in R\bar{b}(m', \bar{d})} \text{label}(m'' | d) \geq 1$ (otherwise $\text{label}(m' | d) = 1$, it follows $\text{label}(m | d) = 0$ from Definition 9. Contradiction.).

Let $CD = R\bar{b}(m', \bar{d}) \cap Win(d)$, it is obvious that if $\sum_{m'' \in CD} \text{label}(m'' | d) = 0$, then $\text{label}(m' | d) = 1$. Therefore $\text{label}(m | d) = 0$, in other words, critical element $m \in_i d$ and the set of moves CD satisfy condition ① in Definition 11.

Thus it remains to show that CD is minimal w.r.t. set inclusion. Suppose CD is not minimal set, i.e. there exists $CD^- \supsetneq CD$, such that if $\sum_{m'' \in CD^-} \text{label}(m'' | d) = 0$, it holds $\text{label}(m | d) = 0$. It follows $CD^- \supseteq R\bar{b}(m', \bar{d}) \cap Win(d)$. Therefore $CD^- = CD$, i.e. $CD = R\bar{b}(m', \bar{d}) \cap Win(d)$ is minimal w.r.t. set inclusion. \square

From Theorem 1 and Theorem 2, it is easy to see that:

Corollary 1. Let $m \in_i d$ be critical element. For any $m' \in R\bar{b}(m, \bar{d})$, there exists at least one critical factor $CF_{PO}(d)$ of d , such that $R\bar{b}(m', \bar{d}) \cap Win(d) \subseteq CF_{PO}(d)$.

Corollary 2. Let a set CF be a critical factor with single critical element m , then for any $m' \in R\bar{b}(m, \bar{d})$,

$Rb(m', \bar{d}) \cap Win(d)$ is also a critical factor.

Corollary 1 and Corollary 2 shows that for each critical dependence CD of critical element, there exists critical factor $CF_{PO}(d)$, such that $CD \subseteq CF_{PO}(d)$. Theoretically, following these ideas, we can acquire all critical factors (it remains to discuss in section 4).

By Definition 10, both pro and con critical factors are those moves that take direct effects on the initial move of multi-party dialogues. They provide guidelines for participants making targeted and valid moves. From another angle of view, critical factors suggest a novel approach to avoiding invalid and useless moves in dialogue games.

Definition 12. Let a multi-party dialogue d , the legal move function $\psi_E : \mathcal{P}(M) \rightarrow \mathcal{P}(M)$ is eligible iff

$$\psi_E(\bar{d}) = \begin{cases} M_\Phi, & \text{if } \bar{d} = \emptyset \\ \psi(\bigcup_{CF_{PRO} \in SCF_{PRO}} CF_{PRO}) \setminus Def(d), & \text{if } \bar{d} \neq \emptyset \text{ and } label(m_0 | d) = 1 \\ \psi(\bigcup_{CF_{OPP} \in SCF_{OPP}} CF_{OPP}) \setminus Def(d), & \text{if } \bar{d} \neq \emptyset \text{ and } label(m_0 | d) = 0 \end{cases}$$

It can be seen that eligible legal move function ψ_E finds attack move based on critical factors. That is to say each move suggested by ψ_E is aiming to a critical element of multi-party dialogue.

Definition 13. (Critical countermeasure) Let CF_{PO} be a critical factor of d , a set $CM \subseteq \psi_E(\bar{d})$ of moves is said to be critical countermeasure to CF_{PO} , denoted by $CM \underline{R} CF_{PO}$, iff following conditions hold: ① for any move m in CF_{PO} , CM has move m' attacking m ; ② CM is minimal w.r.t. set inclusion.

By Definition 13, a critical countermeasure is a minimal set of moves which can be used to attack a critical factor $CF_{PO}(d)$ such that $\sum_{m \in CF_{PO}(d)} label(m | d) = 0$.

For example, as in Fig 1 (c), $SCF_{PRO}(d_{(c)}) = \{\{m_0\}, \{m_3\}, \{m_4\}\}$ is a set of critical factors, and in Fig.1(d), it is clear that $\{m_5\}$, $\{m_6\}$ are critical countermeasures aiming to $\{m_0\}$, $\{m_4\}$ respectively.

Definition 14. (Countermeasure function) Let a set $SCF_{PO}(d)$ of critical factors of d , countermeasure function $\Theta : \mathcal{P}(M) \rightarrow \mathcal{P}(\mathcal{P}(M))$ defined as:

$$\Theta(\bar{d}) = \begin{cases} \{CM \in \mathcal{P}(M_\Phi) \mid CM \models 1\}, & \text{if } \bar{d} = \emptyset \\ \{CM \mid \exists CF \in SCF_{PO}(d), \text{ s.t. } CM \underline{R} CF\}, & \text{if } \bar{d} \neq \emptyset \end{cases}$$

Countermeasure function suggests a set of critical countermeasures from which the proponents or opponents can select to reach their goal. Because all moves provided by eligible legal move function aim at critical factor, there exit no idle and invalid moves in each critical countermeasure. And each critical countermeasure suggested by countermeasure function is also minimal.

Following the critical countermeasure and M extension (see 错误! 未找到引用源。), we are now

considering whether both the proponents and opponents can expand multi-party dialogues with critical countermeasure in the course of dialogue games.

Definition 15. (Critical extension) Given an eligible legal move function ψ_E and countermeasure function Θ . d' is said a critical extension of d under (ψ_E, Θ) , if there exists $E_C \in \Theta(\bar{d})$ such that d' is a legal extension of d with E_C .

It is also said that E_C is a *critical extension factor* from d to d' . Obviously, any critical extension is legal extension under ψ_E .

Definition 16. (Termination and outcome) (1) A multi-party dialogue d terminates if there exists no critical extension factor for d under (ψ_E, Θ) .

(2) Let multi-party dialogue d terminates. For any $m \in d$, if $label(m | d) = 1$, then argument $Arg(m)$ included in m is said to be vindicable in d ; otherwise argument $Arg(m)$ is defeated if $label(m | d) = 0$.

(3) A terminating multi-party dialogue d is said to be successfully terminated by the proponents if $label(m_0 | d) = 1$; otherwise it is successfully terminated by the opponents.

Next, the soundness and completeness of multi-party dialogue game conducted by (ψ_E, Θ) are discussed.

Definition 17. (Soundness and completeness) Given a finite distributed argumentation system $DAF = (A, R)$, and a multi-party dialogue d w.r.t. argument $\phi \in A$.

(1) A multi-party dialogue game is sound if the set $DS = \{Arg(m) \mid m \in \bigcup_{CF \in SCF_{PRO}(d)} CF\}$ of a multi-party dialogue d , which is successfully terminated by the proponents with this game, is a defense set of ϕ .

(2) A multi-party dialogue game is said to be complete if for each defensible argument ϕ in $DAF = (A, R)$, there exists multi-party dialogue d w.r.t. ϕ which is successfully terminated by the proponents, and the set $DS = \{Arg(m) \mid m \in \bigcup_{CF \in SCF_{PRO}(d)} CF\}$ is a defense set of ϕ .

Theorem 3. Let ψ_E , Θ are eligible legal move function and countermeasure function respectively, then:

(1) Multi-party dialogue game conducted by (ψ_E, Θ) is sound;

(2) Multi-party dialogue game conducted by (ψ_E, Θ) is complete.

Proof. (1) Soundness part.

Let multi-party dialogue d conducted by (ψ_E, Θ) is successfully terminated by the proponent. Then, we want to show that $DS = \{Arg(m) \mid m \in \bigcup_{CF \in SCF_{PRO}(d)} CF\}$ is a defense set for ϕ . In other words, we need prove that DS is admissible and $\phi \in DS$ (see Definition 3).

Since the proponents successfully terminate d , hence $label(m_0 | d) = 1$, i.e. $\phi \in DS$ (see 0);

It remains to show that DS counterattacks every attack against it.

For each argument $a \in A$ attacking an argument $b \in DS$, then let move $m^a = (par, a)$, $m^b = (par, b)$, it follows that $m^a R m^b$. The move m^b is critical element because $b \in DS$. Also by Theorem 2, it holds $m^b \geq b(m^a, d) \cap Win(d)$ since $m^a \in R b(m^b, d) \neq \emptyset$. By Corollary 1, it follows that there exists critical factor $CF_{PRO}(d) \in SCF_{PRO}(d)$, such that $R b(m^a, \bar{d}) \cap Win(d) \subseteq CF_{PRO}(d)$. Then it is obvious that $R b(m^a, \bar{d}) \cap Win(d) \subseteq \bigcup_{CF \in SCF_{PRO}(d)} CF$, i.e. there exists $m \in \bigcup_{CF \in SCF_{PRO}(d)} CF$ satisfying $m R m^a$. In other words, DS counterattacks every attack against it.

DS is conflict-free. Otherwise, suppose that there are moves $m, m' \in \bigcup_{CF \in SCF_{PRO}(d)} CF$ satisfying $m R m'$. It is easy to see that $label(m | d) = 1$, because move m is critical element in mdg . By Definition 9, hence $label(m' | d) = 0$. However, we are informed that $label(m' | d) = 1$, because m' is also critical element in d . Contradiction! So, DS is a conflict-free set.

Thus, DS is defense set of ϕ .

(2)Completeness part.

Let $\phi \in A$ is defensible in $DAF = (A, R)$, i.e. there exists admissible set $S \subseteq A$ such that $\phi \in S$. Let multi-party dialogue d conducted by (ψ_E, Θ) starts by the initial move m_0 containing ϕ . Then there exists a dialogue d such that $PRO(d) \subseteq M_S$:

If d is successfully terminated by the proponent, then we are done, it holds the conclusion.

It is easy to see that d is not successfully terminated by the opponent either (otherwise it contradicts that admissible set S contains ϕ). It follows that d is not successfully terminated, i.e. there exists legal expansion factor $E \in \psi_E(\bar{d})$ for d under ψ_E . Specially, there are two cases:

Case 1: $label(m_0 | d) = 1$. For the set $SCF_{PRO}(d)$ of pro critical factors, there exists legal expansion factor $E \in \psi_E(\bar{d})$ under ψ_E for the opponent.

Case 2: $label(m_0 | d) = 0$. For the set $SCF_{OPP}(d)$ of con critical factors, there exists legal expansion factor $E \in \psi_E(\bar{d}) \subseteq M_S$ under ψ_E for the proponent.

Since the distributed argumentation system DAF is finite, then it is easy to see that the multi-party dialogue w.r.t. ϕ is successfully terminated by the proponents finally (Otherwise it violates the defensibility of argument ϕ). Meanwhile, $DS = \{\text{Arg}(m) | m \in \bigcup_{CF \in SCF_{PRO}(d)} CF\} \subseteq S$ is defense set of argument ϕ by the soundness theorem. \square

IV. CRITICAL FACTOR AND THE DEFENSIBILITY OF ARGUMENT

From previous sections, it can be seen that critical factor, a significant notion in multi-party dialogue games,

plays as both the guideline for the proponents and opponents to make moves and the component of the defense set of initial move (for a multi-party dialogue terminated by the proponents successfully). From Definition 11 and Theorem 1, the substitutability of critical element suggests an approach to computing the set of critical factors.

For any multi-party dialogue d , it follows that if $label(m_0 | d) = 1$, then $\{m_0\}$ is one of the pro critical factors, i.e. $\{m_0\} \in SCF_{PRO}(d)$ (by 0 (1)); and it analogously holds that if $label(m_0 | d) = 0$, then the set $R b(m_0, \bar{d}) \cap Win(d)$ is one of the con critical factors, i.e. $R b(m_0, \bar{d}) \cap Win(d) \in SCF_{OPP}(d)$ (by 0 (2)). Intuitively, the set of pro (or con) critical factors can be established from $\{m_0\}$ (or $R b(m_0, \bar{d}) \cap Win(d)$) with substitutability of critical element in critical factor depicted in Theorem 1.

Therefore, starting by $\{m_0\}$ (or $R b(m_0, \bar{d}) \cap Win(d)$), algorithm 1 exposites the computation of all pro (or con) critical factors of multi-party dialogue with DFS (depth-first search).

Algorithm 1. COMPUTE_SETOFCRITICALFACTORS

Input: multi-party dialogue d ;

Output: a set of critical factors $SCF_{PO}(d)$;

```

01 if  $label(m_0 | d) = 1$ 
02   let  $SCF_{PRO}(d) \leftarrow \{m_0\}$ 
      /*initialize the set of pro critical factors*/
03    $CF_{PRO}(d) \leftarrow \{m_0\}$ 
04   CriticalDependence( $m_0, CF_{PRO}(d), SCF_{PRO}(d)$ );
05 else if  $label(m_0 | d) = 0$ 
06   let  $SCF_{OPP} \leftarrow \{R b(m_0, d) \cap Win(d)\}$  ;
      /*initialize the set of con critical factors*/
07    $CF_{OPP}(d) \leftarrow R b(m_0, d) \cap Win(d)$ 
08   for all  $m \in CF_{OPP}(d)$ 
09     CriticalDependence( $m, CF_{OPP}(d), SCF_{OPP}(d)$ );
Function CriticalDependence( $m, CF_{PO}(d), SCF_{PO}(d)$ )
10 if  $R b(m, d) \neq \emptyset$ 
11    $SCD = \{R b(m', d) \cap Win(d) | m' \in R b(m, d)\}$  ;
      /* critical dependences of  $m$  */
12   for all  $CD \in SCD$ 
13      $newCF_{PO}(d) \leftarrow (CF_{PO}(d) \setminus \{m\}) \cup CD$  ;
      /*  $newCF_{PO}(d)$  is critical factor established by replacing
          $m$  in  $CF_{PO}$  with its critical dependence */
14      $SCF_{PO}(d) \leftarrow SCF_{PO}(d) \cup \{newCF_{PO}(d)\}$  ;
15   for all  $m^* \in CD$ 
16     CriticalDependence( $m^*, newCF_{PO}(d), SCF_{PO}(d)$ );
17 else if  $R b(m, d) = \emptyset$  /*  $m$  has no critical dependence*/
18   return null;

```

Proposition 3. (Termination) For any finite multi-party dialogue d , algorithm 1 is guaranteed to terminate returning a set of pro or con critical factors of d .

For any terminated multi-party dialogue d , $label(m_0 | d) = 1$ represents that the proponents successfully terminate d and $\phi = \text{Arg}(m_0)$ is defensible; contrarily $label(m | d) = 0$ means that the opponents

successfully terminate d and thus $\phi = \text{Arg}(m_0)$ is defeated.

For a multi-party dialogue d terminated by the proponents successfully, then we usually consider the defense set of the initial argument. This can be solved by Definition 17 and Theorem 3, which state that all arguments contained in pro critical factors of d constitute the defense set of $\phi = \text{Arg}(m_0)$.

V. RELATED WORK AND COMPARISON

With the development of abstract argument systems in the last decade, more and more researchers are devoted to the computation of the credulous and skeptical semantics of an abstract argumentation framework. One method adopted widely is through the use of dialogue games (or argument games), in which agents present arguments for or against a certain proposition (i.e. argument in question) by attacking those arguments proposed by the opponents. The TPI-disputes (Two Party Immediate Response Disputes) constructs disputes between the proponent and opponent, in which the players try their best to attack each other's most recent argument alternately^[3]. In addition, Prakken has discussed the notion of relevance of arguments in the dialogues^[10]. He argued that each move executed in the dialogue should have effect on the status of the initial move. In [11], PADUA (Protocol for Argumentation Dialogue Using Association Rules) is elaborated for classification of a new example through persuasion dialogue between two agents. Based on PADUA, the same authors propose PISA (Pooling Information from Several Agents) which allows more than two protagonists arguing from experience according a classification question^[4]. To our best knowledge, PISA is the only existing multi-party dialogue game.

The work reviewed above however put their focus on solving semantic of a static and centralized argumentation system where arguments and attack relationship are known and stored in one location. Using Defeasible Logic Programming(DeLP), Thimm et al. proposed distributed argumentation in multi-agent systems, called ArgMAS^[12]. In their framework, agent is capable of generating arguments for a given query and counterarguments against the arguments of other agents. All arguments are monitored by a role called moderator, which is similar to a judge overlooking the defender and accuser in a legal case. But they did not consider how to guide the agents making valid moves during the argumentation process. This paper provides a multi-party dialogue game, established on the notions of critical factor and critical countermeasures, for distributed argumentation. Critical factor is a minimal set of moves which play critical role in a multi-party dialogue game. It provides a guideline for participants making the most helpful moves to change the status of the initial move. Critical countermeasure is a minimal set of moves attacking a critical factor so as to change the status of the initial move. Obviously, critical countermeasure is analogical to the notion of relevance proposed by Prakken in [10].

VI. CONCLUSIONS

Belief conflicts are inevitable among different agents under imperfect information environment. Multi-party dialogue game is a useful means for heterogeneous agents to resolve conflicts in a distributed setting where agents put forward their arguments to, convince each other the truth of a certain proposition.

In this paper, we concern argumentation in a distributed scenario where both the proponents and opponents consist of several participating agents. Each agent is equipped with ability to generate initial argument for a given topic, and to propose counterarguments to attack arguments made by the opponents. We have proposed the notions of critical factor and critical countermeasure, and legal move and countermeasure functions for eliminating repeated, idle and invalid moves in a multi-party dialogue game. With these notions, the mechanism of multi-party dialogue game is discussed for computing the defense set of a defensible argument. We have also proved the soundness and completeness of multi-party dialogue game for a defensible argument in a distributed argumentation system. In addition, we have specified an algorithm for computing the set of pro (or con) critical factors of multi-party dialogue by DFS.

Our immediate further work is to implement our proposed multi-party dialogue game and then study the consequence. We will then explore the strategies for individual agents making high quality argument contributions along the line of [13]. We are also planning to explore different applications of our distributed argumentation system, e.g. in cloud computing and arguing agent competition^[14].

ACKNOWLEDGMENT

The authors acknowledge the support of National Natural Science Foundation of China (No. 70971134 and 9071615).

REFERENCES

- [1] Oliva E, McBurney P, Omicini A: Co-argumentation artifact for agent societies, I. Rahwan, S. Parsons, C. Reed, editor, Argumentation in Multi-Agent Systems, Berlin: Springer-Verlag Berlin, 2008: 31-46.
- [2] Karunatillake N C, Jennings N R, Rahwan I, et al. Dialogue games that agents play within a society. Artificial Intelligence, 2009,173 (9-10): 935-981.
- [3] Vreeswijk G A W, Prakken H: Credulous and skeptical argument games for preferred semantics, M. OjedaAciego, I.P. DeGuzman, G. Brewka, L.M. Pereira, editor, Logics in Artificial Intelligence, Berlin: Springer-Verlag Berlin, 2000: 239-253.
- [4] Wardeh M, Bench-Capon T, Coenen F: Multi-Party Argument from Experience, P. McBurney, I. Rahwan, S. Parsons, N. Maudet, editor, Argumentation in Multi-Agent Systems, 2009: 216-235.
- [5] Dung P M. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence 1995,77 321-357.
- [6] Gordon T F, Karacapilidis N. The Zeno argumentation framework [C]. In Proceedings of the Sixth International Conference on AI and New York, NY, USA: ACM Press,1997. 10-18.

- [7] Bench-Capon T J M. Value-based argumentation frameworks [C]. In *Proceedings of Nonmonotonic Reasoning*.2002. 444-453.
- [8] YAO L, YUAN J, QI X. Dialectic Analysis Model for Multilateral Argument Game [C]. In *Proceeding of Chinese Artificial Intelligence*.Beijing: 2009. 104-109.
- [9] Jakobovits H, Vermeir D. Dialectic Semantics for Argumentation Frameworks [C]. In *Proceedings of the 7th International Conference on Artificial Intelligence and Law*.ACM Press,1999. 53-62.
- [10] Prakken H. Coherence and flexibility in dialogue games for argumentation. *Journal of Logic and Computation*, 2005,15 1009-1040.
- [11] Wardeh M, Bench-Capon T, Coenen F P. Arguments from Experience: The PADUA Protocol [C]. In *COMMA'2008*.IOS press,2008. 405-416.
- [12] Thimm M: Realizing Argumentation in Multi-agent Systems Using Defeasible Logic Programming, P. McBurney, I. Rahwan, S. Parsons, N. Maudet, editor, *Argumentation in Multi-Agent Systems*, 2009: 175-194.
- [13] Yuan, T., Svansson, V., Moore, D. and Grierson, A. A Computer Game for Abstract Argumentation. In *Proceedings of IJCAI'2007 Workshop on Computational Models of Natural Argument*, pp.62-68, Hyderabad, India, 2007..
- [14] Yuan, T., Schulze, J., Devereux, J. and Reed, C. Towards an Arguing Agents Competition: Building on Argumento. In *Proceedings of IJCAI'2008 Workshop on Computational Models of Natural Argument*, Patras, Greece, 2008.

An Approach of Trustworthiness Evaluation of Software Behavior Based on Multidimensional Fuzzy Attributes

Zhen Li and Junfeng Tian

College of Mathematics and Computer, Hebei University, Baoding, 071002, China

Email: lizhen_hbu@yahoo.com.cn, jftian@hbu.edu.cn

Abstract—In order to increase the accuracy of trustworthiness evaluation of software, the paper improves the traditional construction process of the expected behavior trace of software and proposes an approach of trustworthiness evaluation of software behavior based on multidimensional fuzzy attributes. First, training samples of the same monitoring point are clustered based on multidimensional fuzzy attributes to construct a more accurate expected behavior trace of software. Second, an improved weight distribution method of multidimensional fuzzy attributes is presented based on correlation coefficient and standard deviation integrated approach (CCSD) for weight distribution of attributes in multiple attribute decision making. The improved weight distribution method is suitable for one-class samples from monitoring point and it considers both the dispersion of fuzzy attribute's value and the influence among these fuzzy attributes. Finally, experiments and analyses show that: ① the expected behavior trace of software constructed by training after clustering is more accurate than without clustering; and ② our improved weight distribution method of multidimensional fuzzy attributes has better effect of trustworthiness evaluation than CCSD and other methods of weight distribution for one-class samples.

Index Terms—software behavior, trustworthiness evaluation, fuzzy attribute, clustering, weight distribution

I. INTRODUCTION

With continuous deepening of the application of software in the sensitive fields such as finance, military affairs and economy, the requirement of software trustworthiness becomes more urgent. How to ensure high confidence of software during software development and running has become an important research direction of software theory and technology [1]. If the software behavior is always accordant with the expected behavior, we call the software is trustworthy [2]. For trustworthy software, the behavior and results can be expected and the behavior states can be monitored when it runs.

The expected behavior trace of software is usually

Supported by the National Natural Science Foundation of China (Grant No. 60873203, 61170254), the Outstanding Youth Foundation of Hebei Province (Grant No. F2010000317), the Natural Science Foundation of Hebei Province (Grant No. F2010000319, F2011201039) and the Science Foundation of Hebei University (Grant No. 2008Q09).

composed of a sequence of monitoring points and events or actions causing monitoring points' transition. The common monitoring points are system call, function module, component, etc. The fine-grained monitoring points result in a high degree of trustworthiness, but with low software running efficiency. Therefore, the monitoring points should be set by comprehensively considering the needs of the degree of trustworthiness and software running efficiency in actual application. For each monitoring point, there are a group of attributes describing the expected running situation when the software runs to it. It is very important to study the multidimensional attributes of monitoring points for the trustworthiness evaluation of software behavior.

II. RELATED WORK

Because system call is the interface provided by the operating system to access system resources, system call has become the important monitoring point of software behavior. A lot of software behavior automaton models based on system call have appeared, such as finite state automaton (FSA) model and pushdown automaton (PDA) model built by static analysis of source code [3], Vt-Path model built by dynamic learning [4], HPDA model combining static analysis with dynamic learning [5], context-sensitive Dyck model [6], HFA model with static-dynamic hybrid approach [7] and the model with dataflow analysis [8]. These models can describe the running trace of software and some of them introduce arguments policies and context attributes for system call. Jones [9] and Pu et al. [10] involve time interval attribute of system calls as the criteria for determining abnormal from normal signatures. In our previous work [11], we evaluate the trustworthiness of software by several attributes of monitoring points, such as function, arguments policies, context, timestamp, memory occupancy rate and CPU occupancy rate. Li et al. [12] monitor the software behavior from the aspects of availability, reliability and security and consider many attributes such as CPU occupancy rate, IP transmission efficiency, memory occupancy rate, bandwidth utilization, average throughput, and the number of illegal connections.

The above attributes of software monitoring point can be divided into two categories: deterministic attributes

and fuzzy attributes. For deterministic attributes, once any of them deviates from the normal value, the monitoring point is determined to be untrustworthy directly. These attributes include function, arguments policies, context, etc. Fuzzy attributes cannot be expressed as accurate numbers. They are fuzzy and granted the prescribed error bounds. These attributes include CPU occupancy rate, IP transmission efficiency, memory occupancy rate, time interval, etc. The above references construct the expected behavior of software monitoring points by running the software many times in the training phase, and take all samples of the same monitoring point as one training set. The sample values of a fuzzy attribute for the same monitoring point can vary obviously because the path from last monitoring point to the current monitoring point can vary. Therefore, the expected behavior trace trained by the above approaches is inaccurate.

According to the feature of deterministic attributes, weight distribution is meaningless for deterministic attributes. If all the deterministic attributes of the monitoring point are trustworthy, the trustworthiness of monitoring point depends on its fuzzy attributes and the weight distribution method of fuzzy attributes determines the accuracy of trustworthiness evaluation directly for monitoring point and then for software. The common methods can be grouped roughly into three categories: subjective [13-15], objective [16,17] and integrated [18,19]. Because it is difficult to determine the weights of fuzzy attributes for monitoring point from the subjective experience, we mainly consider the objective methods. Li et al. [12] determine the weight of attributes by information entropy theory which makes the results be objective and adaptive relatively. Wang et al. [20] propose a feature weight learning algorithm which gives each feature a feature weight by minimizing the feature evaluation index through gradient descent technique. Wang et al. [21] propose a correlation coefficient and standard deviation integrated approach (CCSD) for determining the weights of attributes in multiple attribute decision making. The training samples for weight distribution used in these approaches either belong to different classes or are alternative schemes for multiple attribute decision making. While, the training samples for one monitoring point belong to the same class, the above methods cannot be used to the weight distribution of one-class samples. The methods of weight distribution for one-class samples are less currently. Information entropy [22,23] can be a solution and its basic idea is to determine the weight of each attribute according to the attribute's information entropy for one-class samples. The smaller attribute's information entropy means that the sample data is more regular, and the model constructed is better, so the weight of the attribute is larger, and vice versa. However, information entropy only considers the dispersion of attribute value, and doesn't involve the influence among these attributes. The methods of weight distribution for one-class samples need to improve greatly.

On the basis of the above problems, this paper presents an approach of trustworthiness evaluation of software

behavior based on multidimensional fuzzy attributes. First, the expected behavior of software monitoring points is constructed for each class of samples according to the clustering results based on multidimensional fuzzy attributes. Second, an improved weight distribution method of multidimensional fuzzy attributes is proposed based on CCSD method for determining the weights of attributes in multiple attribute decision making. Both the dispersion of attribute value and the influence among these attributes are considered, which ensures the better effect of trustworthiness evaluation.

III. CLUSTERING BASED ON MULTIDIMENSIONAL FUZZY ATTRIBUTES

In the training phase, the sample values of a fuzzy attribute for the same monitoring point can vary obviously because the path from last monitoring point to the current monitoring point can vary. Therefore, the expected behavior trace constructed is inaccurate if all samples of the same monitoring point are taken as one training set. We solve the problem by clustering these samples based on multidimensional fuzzy attributes.

For n samples X_1, X_2, \dots, X_n of monitoring point mp , each sample has m fuzzy attributes A_1, A_2, \dots, A_m , denoted by sample matrix $X = (x_{ij})_{n \times m}$ ($1 \leq i \leq n, 1 \leq j \leq m$). These samples are clustered by a clustering algorithm based on entropy [24].

For monitoring point mp , the clustering process of n samples is as follows:

① Number each sample as a class C_k , where $1 \leq k \leq m$ and C_k is composed of N_k samples.

② Choose any sample, add it to another class, and compute the entropy of C_k after adding the sample to it according to (1) (h is the smoothing coefficient). The sample is allocated to the class which has the smallest added value of entropy.

$$H(C_k) = -\log\left[\frac{1}{(2\pi)^m N_k^2 h^{2m}} \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} \exp\left(-\frac{(x_i - x_j)(x_i - x_j)^T}{2h^2}\right)\right] \quad (1)$$

③ Number the classes again after allocating samples.

④ Repeat ② and ③ until there is no class which has only one sample.

⑤ Compute the between-class entropy of any two classes according to (2) and merge two classes with smallest between-class entropy. The number of classes and samples for each class must be also modified.

$$H(C_i, C_j) = -\log\left[\frac{1}{(2\pi)^m N_i^2 N_j^2 h^{2m}} \sum_{i=1}^n \sum_{j=1}^n M(x_i, x_j) \exp\left(-\frac{(x_i - x_j)(x_i - x_j)^T}{2h^2}\right)\right] \quad (2)$$

$M(x_i, x_j)$ in (2) is as follows:

$$M(x_i, x_j) = \begin{cases} 1, & x_i \in C_i, x_j \in C_j \text{ or } x_i \in C_j, x_j \in C_i \\ 0, & \text{else} \end{cases} \quad (3)$$

Repeat ⑤ until there is only one class. If the changes of the smallest between-class entropy at a certain time are

more significant than before, the classes at this time, represented as C_1, C_2, \dots, C_M (M is the number of classes), are the result of clustering.

IV. TRUSTWORTHINESS EVALUATION OF SOFTWARE BEHAVIOR

A. Trustworthy Degree of Training Sample

For the training samples $Y_1^q, Y_2^q, \dots, Y_n^q$ of class C_q ($1 \leq q \leq M$), let $y_{1j}^q, y_{2j}^q, \dots, y_{nj}^q$ be n' sample values of the fuzzy attribute A_j . After removing various effects of the environment, fuzzy attributes are approximately normally distributed [25] and the normal value of each fuzzy attribute fluctuates around the average value. Therefore, we can determine the trustworthy degree of fuzzy attribute according to the degree of deviation from the average value. The average value μ_j^q of fuzzy attribute A_j is as follows:

$$\mu_j^q = (\sum_{i=1}^{n'} y_{ij}^q) / n'. \quad (4)$$

The trustworthy degree d_{ij}^q of fuzzy attribute A_j for training sample Y_i^q ($1 \leq i \leq n'$) is as follows:

$$d_{ij}^q = \begin{cases} \frac{y_{ij}^q}{\mu_j^q}, & y_{ij}^q \in [y \min_j^q, \mu_j^q] \\ 1 + \frac{\mu_j^q}{y \max_j^q} - \frac{y_{ij}^q}{y \max_j^q}, & y_{ij}^q \in [\mu_j^q, y \max_j^q] \end{cases}, \quad (5)$$

where $y \min_j^q = \min_{1 \leq i \leq n'} \{y_{ij}^q\}$, $y \max_j^q = \max_{1 \leq i \leq n'} \{y_{ij}^q\}$, $j = 1, 2, \dots, m$.

From (5), we can see $d_{ij}^q \in [0, 1]$. The trustworthy degree d_i^q of training sample Y_i^q is $d_i^q = \sum_{j=1}^m w_j^q d_{ij}^q$, where w_j^q is the weight of fuzzy attribute A_j for class C_q .

B. Trustworthiness Evaluation

In the test phase, for the test sample Z of software monitoring point mp , let z_1, z_2, \dots, z_m be the values of m fuzzy attributes. The trustworthy degree d_j^q of fuzzy attribute A_j for class C_q is as follows:

$$d_j^q = \begin{cases} \frac{z_j}{\mu_j^q}, & z_j \in [y \min_j^q, \mu_j^q] \\ 1 + \frac{\mu_j^q}{y \max_j^q} - \frac{z_j}{y \max_j^q}, & z_j \in [\mu_j^q, y \max_j^q] \\ 0, & \text{else} \end{cases}. \quad (6)$$

The trustworthy degree d^q of test sample Z for class C_q is: $d^q = \sum_{j=1}^m w_j^q d_j^q$. For software monitoring point mp , the trustworthy degree d of test sample Z is: $d = \max_{1 \leq q \leq M} \{d^q\}$.

For given threshold of trustworthy degree τ , if $d \in [\tau, 1]$, the software monitoring point mp is trustworthy; if $d < \tau$, mp is untrustworthy. The degree of untrustworthy is increased as the deviation degree from 1 increases. When any software monitoring point is reported to be untrustworthy, the software stops running.

C. An Improved Weight Distribution Method of Multidimensional Fuzzy Attributes (Improved CCSD)

For the clustered samples, the weights of fuzzy attributes are very important to the accuracy of trustworthiness evaluation for monitoring point. CCSD method [21] for determining the weights of attributes in multiple attribute decision making considers attributes with big standard deviations should be given more important weights than those attributes with small standard deviations. However, this property is not suitable for the weight distribution of fuzzy attributes of monitoring points. We solve the problem and give an improved CCSD method for monitoring points.

For the training samples $Y_1^q, Y_2^q, \dots, Y_n^q$ of class C_q , the standard deviation σ_j^q of fuzzy attribute A_j is as follows:

$$\sigma_j^q = \sqrt{[\sum_{i=1}^{n'} (y_{ij}^q - \mu_j^q)^2] / (n' - 1)}. \quad (7)$$

The smaller standard deviation means that the sample data is more centralized, and the ability to describe normal behavior is stronger, so the weight of fuzzy attribute A_j is larger, and vice versa.

When fuzzy attribute A_j is dropped out, the trustworthy degree d_i^{qj} of sample Y_i^q is: $d_i^{qj} = \sum_{k=1, k \neq j}^m w_k^q d_{ik}^q$. The correlation coefficient R_j^q between d_{ij}^q and d_i^q is as follows:

$$R_j^q = \frac{\sum_{i=1}^{n'} (d_{ij}^q - \bar{d}_j^q)(d_i^{qj} - \bar{d}^{qj})}{\sqrt{\sum_{i=1}^{n'} (d_{ij}^q - \bar{d}_j^q)^2 \cdot \sum_{i=1}^{n'} (d_i^{qj} - \bar{d}^{qj})^2}}, \quad (8)$$

where $\bar{d}_j^q = \frac{1}{n'} \sum_{i=1}^{n'} d_{ij}^q$, $\bar{d}^{qj} = \frac{1}{n'} \sum_{i=1}^{n'} d_i^{qj}$.

If R_j^q is high enough and close to one, then fuzzy attribute A_j has little effect on the trustworthiness of the sample and it can therefore be assigned a very small weight. If R_j^q is very low, say close to minus one, then fuzzy attribute A_j has significant impact on the

trustworthiness of the sample and it should be given a very important weight.

We integrate the standard deviation and correlation coefficient and define the weights of fuzzy attributes as follows:

$$w_j^q = \frac{\sqrt{1 - R_j^q} / \sigma_j^q}{\sum_{k=1}^m \sqrt{1 - R_k^q} / \sigma_k^q}. \quad (9)$$

Equation (9) is a system of nonlinear equations. It contains m equations, which can uniquely determine m weight variables. To solve the equation, we convert it into the following nonlinear optimization model for solution:

$$\text{Minimize } J = \sum_{j=1}^m (w_j^q - \frac{\sqrt{1 - R_j^q} / \sigma_j^q}{\sum_{k=1}^m \sqrt{1 - R_k^q} / \sigma_k^q})^2, \text{ subject to}$$

$\sum_{j=1}^m w_j^q = 1$, $w_j^q \geq 0$, which can be solved by LINGO software and at optimality the objective function value $J^* = 0$.

V. EXPERIMENTS AND ANALYSIS

A. Test of Clustering Based on Multidimensional Fuzzy Attributes

We have performed experiments on a PC with Intel (R) Core (TM)2 Duo E7500 2.93 GHz and 2 GB of main memory running Linux kernel 2.4.20. The grain of monitoring point is set to system call. Each system call is intercepted by loadable kernel module (LKM) and modified to capture the attributes' values of the system call. Because the trustworthiness of deterministic attributes determines the trustworthiness of monitoring point directly, we mainly discuss the effect of fuzzy attributes on the trustworthiness of monitoring point.

Software vi6.1 is the editor in Red Hat 9 Linux. In clean environment, we capture 15 samples along three different paths in a monitoring point of software vi6.1 where three paths converge. Three fuzzy attributes time interval, memory variation and CPU variation are involved and they are the absolute value of difference of time, memory occupancy rate and CPU occupancy rate respectively between the current monitoring point and last monitoring point, as shown in Table I. Table II shows the fuzzy attributes' values of two normal traces and two abnormal traces in the monitoring point for test.

The experiment clusters the training samples based on multidimensional fuzzy attributes. Let smoothing coefficient h be 0.5. The result of clustering is: $C_1: \{1, 2, 3, 4, 5, 6\}$, $C_2: \{7, 8, 9, 10\}$, $C_3: \{11, 12, 13, 14, 15\}$. Suppose the weight of each fuzzy attribute is equal and let the threshold of trustworthy degree τ be 0.85, the trustworthy degrees of test samples without clustering and after clustering for each class are shown in Table III. After clustering, the trustworthy degree of test sample is the maximum of the trustworthy degrees for all classes, that is, the trustworthy degree of Normal 1, Normal 2,

TABLE I.
MULTIDIMENSIONAL FUZZY ATTRIBUTES' VALUES FOR TRAINING SAMPLES

Training sample	Time interval(μs)	Memory variation(%)	CPU variation(%)
1	11	9.2	10.2
2	12	9.4	9.4
3	11	8.3	9.8
4	11	10.1	8.5
5	12	10.7	10.0
6	11	10.2	9.6
7	25	9.8	8.7
8	27	8.3	9.3
9	23	8.8	9.6
10	25	10.2	10.0
11	50	16.5	20.1
12	53	16.7	21.4
13	52	18.0	21.8
14	51	16.4	20.5
15	53	17.8	20.9

TABLE II.
MULTIDIMENSIONAL FUZZY ATTRIBUTES' VALUES FOR TEST SAMPLES

Test sample	Time interval(μs)	Memory variation(%)	CPU variation(%)
Normal 1	12	10.4	9.9
Normal 2	52	17.9	21.6
Abnormal 1	40	19.3	9.5
Abnormal 2	69	25.6	24.9

TABLE III.
THE TRUSTWORTHY DEGREES OF TEST SAMPLES

Test sample	Trustworthy degree		
	Without clustering	Class C_1	Class C_2
Normal 1	0.68	0.95	0.77
Normal 2	0.62	0	0
Abnormal 1	0.70	0	0.48
Abnormal 2	0.32	0	0
			0.67

Abnormal 1 and Abnormal 2 is 0.95, 0.97, 0.70 and 0.67 respectively. Without clustering, two normal traces are determined to be untrustworthy by mistake. While, after clustering, the trustworthiness of both two normal traces and two abnormal traces are determined correctly. Therefore, the expected behavior trace of software is constructed more accurately by training after clustering.

B. Comparison of Weight Distribution Methods

Our experiment uses machine learning databases (MLDBs) from UCI Repository [26] including lots of different databases. We select Iris database which has three classes (setosa, versicolor and virginica). Each class has 50 samples and each sample has four attributes (sepal length, sepal width, petal length and petal width) which are all continuous real variables. There is a great difference between the attributes' values of class setosa and other two classes; While, there is little difference between the attributes' values of class versicolor and class virginica.

Let's take class setosa and class versicolor for example. For each class, we take the first forty-five samples as training samples and the last five samples as test samples. Fig. 1 and Fig. 2 show the weights of each fuzzy attributes computed with four weight distribution methods (equal weight, CCSD [21], information entropy [23] and improved CCSD) for class setosa and class

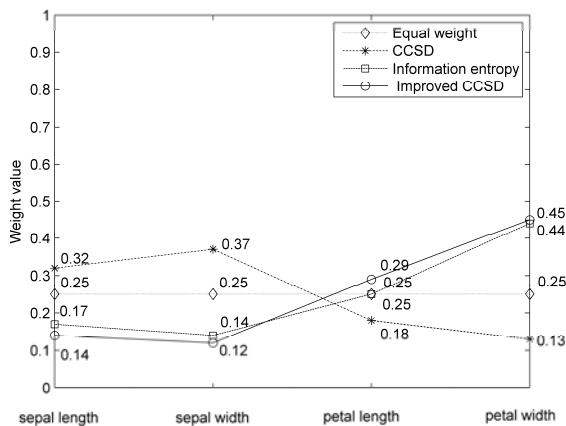


Figure 1. The weights of attributes for class setosa.

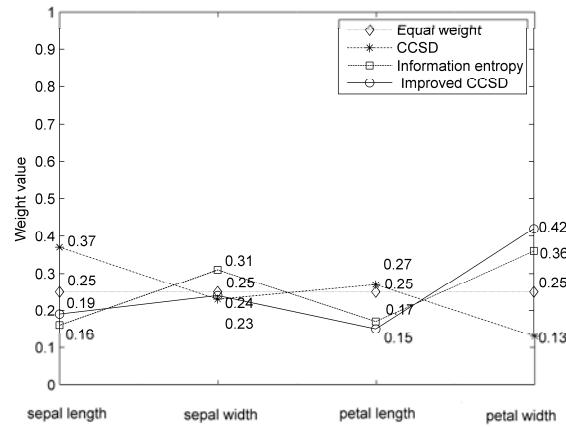


Figure 2. The weights of attributes for class versicolor.

versicolor respectively. For class setosa, there is little difference among the correlation coefficients of four attributes in the improved CCSD, so the weights of attributes mainly depend on the standard deviation. The results of improved CCSD and information entropy are without big difference. For class versicolor, there is a great difference among the correlation coefficients of four attributes in the improved CCSD, so the correlation coefficient has a great influence on the weights of attributes. The results of improved CCSD and information entropy have larger difference.

Let the threshold of trustworthy degree τ be 0.85. In the first experiment, we take the training samples of class setosa as training samples and take the test samples of class setosa, all samples of class versicolor and versicolor as test samples. Because there is a great difference between the attributes' values of class setosa and other

two classes, all test samples of class setosa are determined to be trustworthy and all samples of class versicolor and virginica are determined to be untrustworthy correctly for the four weight distribution methods. In the second experiment, we take the training samples of class versicolor as training samples and take the test samples of class versicolor, all samples of class setosa and virginica as test samples. The test results of four weight distribution methods are shown in Table IV. For the test samples of class versicolor, the test results of four weight distribution methods are same. For samples of class setosa, the evaluation results of four weight distribution methods are same and have no misjudgment samples; For the average of trustworthy degree of four weight distribution methods, improved CCSD has the smallest average of trustworthy degree which means the largest degree of deviation, so it can determine the untrustworthy sample most easily. For samples of class virginica, the evaluation results of improved CCSD have the smallest misjudgment samples and the smallest average of trustworthy degree; the second is information entropy. CCSD for determining the weights of attributes in multiple attribute decision making is the worst one. Therefore, our improved CCSD has better effects for trustworthiness evaluation.

VI. CONCLUSION

The paper proposes an approach of trustworthiness evaluation of software behavior based on multidimensional fuzzy attributes. It clusters the samples of the same monitoring point based on multidimensional fuzzy attributes to construct a more accurate expected trace of software. For a better trustworthiness evaluation effect, the paper proposes an improved weight distribution method of multidimensional fuzzy attributes based on CCSD method for weight distribution of attributes in multiple attribute decision making. The improved weight distribution method is suitable not only for the multidimensional fuzzy attributes of monitoring points, but also for any weight distribution according to one-class samples, so it is of widespread usage. Our future work is to consider the selection of multidimensional fuzzy attributes for monitoring points to achieve a better effect for trustworthiness evaluation of software.

REFERENCES

- [1] H. W. Chen, J. Wang, and W. Dong, "High Confidence Software Engineering Technologies", *Acta Electronica Sinica*, vol. 31, no. 12A, 2003, pp. 1933-1938. (in Chinese)

TABLE IV.
TEST RESULTS OF THE SECOND EXPERIMENT

The class of test sample	The average of trustworthy degree				Number of trustworthy samples determined /Number of Total samples			
	Equal weight	CCSD	Information entropy	Improved CCSD	Equal weight	CCSD	Information entropy	Improved CCSD
versicolor	0.94	0.94	0.94	0.94	4/5	4/5	4/5	4/5
setosa	0.54	0.61	0.51	0.48	0/50	0/50	0/50	0/50
virginica	0.79	0.82	0.77	0.74	15/50	22/50	9/50	4/50

- [2] K. Liu, Z. G. Shan, J. Wang, J. F. He, Z. T. Zhang, and Y. W. Qin, "Overview on Major Research Plan of Trustworthy Software", *Bulletin of National Natural Science Foundation of China*, vol. 22, no. 3, 2008, pp.145-151. (in Chinese)
- [3] D. Wagner and D. Dean, "Intrusion Detection via Static Analysis", *Proc. IEEE Symp. on Security and Privacy*, IEEE Computer Society, May 2001, pp.156-169, doi: 10.1109/SECPRI.2001.924296.
- [4] H. H. Feng, O. M. Kolesnikov, P. Fogla, L. Wenke, and G. Weibob, "Anomaly Detection Using Call Stack Information", *Proc. IEEE Symp. on Security and Privacy*, 2003, pp.62-75, doi: 10.1109/SECPRI.2003.1199328 .
- [5] Z. Liu, S. M. Bridges, and R. B. Vaughn, "Combining Static Analysis and Dynamic Learning to Build Accurate Intrusion Detection Models", *Proc. of the 3rd IEEE Int'l Workshop on Information Assurance*, College Park, March 2005, pp.164-177, doi: 10.1109/IWIA.2005.6.
- [6] J. Giffin, S. Jha, and B. Miller, "Efficient context-sensitive intrusion detection", *Proc. of the 11th Network and Distributed System Security Symposium*, San Diego, USA, 2004, pp.1-15.
- [7] W. Li, Y. X. Dai, Y. F. Lian, and P. H. Feng, "Context Sensitive Host-Based IDS Using Hybrid Automaton", *Journal of Software*, vol.20, no.1, 2009, pp.138-151. (in Chinese)
- [8] B. Sandeep, C. Abhishek, and R. Sekar, "Dataflow Anomaly Detection", *Proc. of the 2006 IEEE Symposium on Security and Privacy*, 2006, pp. 48-62, doi: 10.1109/SP.2006.12.
- [9] A. Jones and S. Li, "Temporal Signatures for Intrusion Detection", *Proc. of 17th Annual Computer Security Applications Conference*, Dec. 2001, pp. 252 - 261, doi: 10.1109/ACSAC.2001.991541.
- [10] S. Pu and B. Lang, "An Intrusion Detection Method Based on System Call Temporal Serial Analysis", *Proc. of the 3rd International Conference on Intelligent Computing*, 2007, pp. 656-666, doi: 10.1007/978-3-540-74171-8_65.
- [11] J. F. Tian, Z. Li, and Y. L. Liu, "A Design Approach of Trustworthy Software and Its Trustworthiness Evaluation", *Journal of Computer Research and Development*, vol.48, no.8, 2011.(in Chinese)
- [12] X. Y. Li, X. L. Gui, Q. Mao, and D. Q. Leng, "Adaptive Dynamic Trust Measurement and Prediction Model Based on Behavior Monitoring", *Chinese Journal of Computers*, vol.32, no.4, 2009, pp. 664-674. (in Chinese)
- [13] F. H. Barron, "Selecting a Best Multiattribute Alternative with Partial Information about Attribute Weights", *Acta Psychologica*, vol. 80, 1992, pp.91-103, doi: 10.1016/0001-6918(92)90042-C.
- [14] M. Deng, W. Xu, and J. B. Yang, "Estimating the Attribute Weights through Evidential Reasoning and Mathematical Programming", *International Journal of Information Technology & Decision Making*, vol. 3, 2004, pp. 419-428, doi: 10.1142/S0219622004001124.
- [15] B.S. Ahn, and KxS. Park, "Comparing Methods for Multiattribute Decision Making with Ordinal Weights", *Computers & Operations Research*, vol. 35, no. 5, 2008, pp. 1660-1670, doi: 10.1016/j.cor.2006.09.026.
- [16] H. Deng, C. H. Yeh, and R. J. Willis, "Inter-company Comparison Using Modified TOPSIS with Objective Weights", *Computers & Operations Research*, vol. 27, 2000, pp. 963-973, doi: 10.1016/S0305-0548(99)00069-6.
- [17] G.W. Wei, "Maximizing Deviation Method for Multiple Attribute Decision Making in Intuitionistic Fuzzy Setting", *Knowledge-Based Systems*, vol. 21, no.8, 2008, pp. 833-836, doi: 10.1016/j.knosys.2008.03.038.
- [18] J. Ma, Z. P. Fan, and L. H. Huang, "A Subjective and Objective Integrated Approach to Determine Attribute Weights", *European Journal of Operational Research*, vol. 112, 1999, pp. 397-404, doi: 10.1016/S0377-2217(98)00141-6.
- [19] Y. M. Wang, and C. Parkan, "A General Multiple Attribute Decision-making Approach for Integrating Subjective Preferences and Objective Information", *Fuzzy Sets and Systems*, vol. 157, 2006, pp. 1333-1345, doi: 10.1016/j.fss.2005.11.017.
- [20] L. J. Wang, S. Y. Guan X. L. Wang, and X. Z. Wang, "Fuzzy C Mean Algorithm Based on Feature Weights", *Chinese Journal of Computers*, vol.29, no.10, 2006, pp.1797-1803. (in Chinese)
- [21] Y. M. Wang, and L. Ying, "Integration of Correlations with Standard Deviations for Determining Attribute Weights in Multiple Attribute Decision Making", *Mathematical and Computer Modeling*, vol. 51, 2010, pp. 1-12, doi: 10.1016/j.mcm.2009.07.016.
- [22] H. Zhang, H. L. Bian, L. F. Wu, Y. S. Zhang, M. W. Cui, and Q. K. Zeng, "Study on Program Behavior Control Based on LSM", *Journal of Software*, vol.16, no.6, 2005, pp.1151-1158. (in Chinese)
- [23] J. F. Tian and Y. Zhu, "Trusted Shell Based Constitution Model of Trusted Software", *China Communications*, vol. 8, no. 4, 2011, pp.11-22.
- [24] H. C. Wang and H. Peng, "A Clustering Algorithm Based on Entropy", *Computer Science*, vol. 34, no. 11, 2007, pp. 178-180. (in Chinese)
- [25] R. J. Larsen and M. L. Marx, *An Introduction to Mathematical Statistics and Its Applications*, Pearson Prentice Hall, 2006.
- [26] UCI Repository of machine learning databases, <http://archive.ics.uci.edu/ml/index.html>, 1998.



Zhen Li, born in Baoding, China, 1981, received M.S. degree of computer application from Hebei University, Baoding, China in 2006.

She is a lecturer of the College of Mathematics and Computer of Hebei University. In the past five years, she has published over 10 technical papers in refereed journals and conference proceedings. Her current research interests include computer security, trust computing and distributed computing.



Junfeng Tian, born in Baoding, China, 1965, received Ph.D degree of computer science from University of Science and Technology of China, Hefei, China in 2004.

He is a professor of Computer Science at Hebei University. In the past five years, he has published over 30 technical papers in refereed journals and conference proceedings. His research interests include network security, trust computing, distributed computing and wireless sensor network.

More information about Prof. Tian can be found at <http://int.hbu.edu.cn/int/szll/tjf.html>.

Personalized Web Search Using Clickthrough Data and Web Page Rating

XuePing Peng

Beijing Institute of Technology, Beijing, 100081, China

Email: pengxp@bit.edu.cn

ZhenDong Niu*, Sheng Huang and Yumin Zhao

Beijing Institute of Technology, Beijing, 100081, China

Email: {zniu, 20812036, oblivion}@bit.edu.cn

Abstract—Personalization of Web search is to carry out retrieval for each user incorporating his/her interests. We propose a novel technique to construct personalized information retrieval model from the users' clickthrough data and Web page ratings. This model builds on the user-based collaborative filtering technology and the top-N resource recommending algorithm, which consists of three parts: user profile, user-based collaborative filtering, and the personalized search model. Firstly, we conduct user's preference score to construct the user profile from clicked sequence score and Web page rating. Then it attains similar users with a given user by user-based collaborative filtering algorithm and calculates the recommendable Web page scoring value. Finally, personalized information retrieval is modeled by three case applies (rating information for the user himself; at least rating information by similar users; not make use of any rating information). Experimental results indicate that our technique significantly improves the search performance.

Index Terms—Personalization, Web page rating, information retrieval, clickthrough data

I. INTRODUCTION

As the amount of information on the Web rapidly increases, it creates many new challenges for Web search. Millions of searches are conducted every day on search engines such as Yahoo!, Google and Bing, etc. Despite the popularity, search engines have their deficiencies: given a query, they usually return a huge list of results, the pages ranked at top may not meet users' needs and the same result regardless of who submitted the query [1]. One reason for this problem is the keyword-based query interface, which is difficult for users to describe exactly what they need. Besides, typical search engines often do not exploit user information. Even two users submit the same query, their information need may be different [21-22].

Personalized Web search is to carry out retrieval for

each user incorporating his/her own information need. To solve this problem, researchers have developed systems that adapt their behavior to the goals, tasks, interests, and other characteristics of their users. Based on models that capture important characteristics of users, these personalized systems maintain their users' preferences and take them into account to customize the content generated or its presentation to the different individuals [2]. Some Web search systems use relevance feedback to refine user needs or ask users to register their demographic information beforehand in order to provide better service[23]. Since these systems require users to engage in additional activities beyond search to specify/modify their preferences manually, approaches that are able to implicitly capture users' information needs should be developed.

This paper focuses on utilizing clickthrough data and Web page ratings to improve Web search. Clickthrough data can be extracted from a large amount of search logs accumulated by web search engines. These logs typically contain user-submitted search queries, the URL of Web pages which are clicked by users in the corresponding search result page[24]. The data objects contained in the clickthrough data are of different types: user, query and Web page. By performing analysis on the clickthrough data, we attempt to discover the latent factors among these multi-type objects[1]. However, most of these references extract only clickthrough data for analysis, and ignore the specific characteristics of Web pages. Page rating is one important characteristic, which can be calculated from explicit relevance rates of users who browsed the Web page. By analyzing associations among clickthrough data multi-type objects and computing Web page rating, we construct a personalized search model, and then re-rank search results by the model.

In this paper, by analyzing the clickthrough data and calculating Web page rating, we propose a novel, effective and efficient personalized Web search model. In this model, we give solutions to the following two problems: (1) How to create user profiles, and (2) How to return the different results when the same query is submitted by different users?

Manuscript received Sep. 8, 2011; revised Oct. 20, 2011; accepted Oct. 28, 2011.

Project number: 1110012040112, 91101,3070012231001

*Corresponding author's email: zniu@bit.edu.cn

The remainder of this paper is organized as follows. Section 2 provides related work. Section 3 gives a brief introduction to personalized Web search model. Section 4 presents the experimental results and Section 5 offers some concluding remarks and directions for future research.

II. RELATED WORK

A. Personalized Web Search

Different users may prefer different results for the same query. Personalized search [25-29] aims to provide the most relevant search results to individual users based on their interests. Personalized search comprises two major components: (1) User profiles, and (2) The actual search algorithm [19].

Approaches focused on the User Profile. Sugiyama et al. [31] analyzed surfing behavior and generated user profiles as features (terms) of the visited pages. Upon issuing a new query, the search results were ranked based on the similarity between each URL and the user profile. Machine Learning [32] was used on the past click history of the user in order to determine topic preference vectors and then apply Topic-Sensitive PageRank [33]. User profiling based on browsing history has the advantage of being rather easy to obtain and process. This is probably why it is also employed by several industrial search engines.

Approaches focused on the Personalization Algorithm. Effectively building the personalization aspect directly into PageRank [34] has received much attention recently. Haveliwala [33] computed a topic-oriented PageRank, in which 16 PageRank vectors biased on each of the main topics of the Open Directory were initially calculated off-line, and then combined at run-time based on the similarity between the user query and each of the 16 topics. More recently, Nie et al. modified the idea by distributing the PageRank of a page across the topics it contains in order to generate topic oriented rankings. Jeh and Widom proposed an algorithm that avoids the massive resources needed for storing one Personalized PageRank Vector (PPV) per user by precomputing PPVs only for a small set of pages and then applying linear combination. As the computation of PPVs for larger sets of pages was still quite expensive, several solutions have been investigated, the most important ones being those of Fogaras and Racz, and Sarlos et al., the latter using rounding and count-min sketching in order to fastly obtain accurate enough approximations of the personalized scores.

Only by opening to the outside world, can it bring in adequate flow of negative entropy, make the dissipation occur between telecom industry and the environment, and ultimately evolve to the dissipative structure. The open feature of telecom industry system is the prerequisite and essential condition of self-organized industrial system. As a nation's economical and foundational industrial, telecom industry is in a complex environment and is interdependent with the external environment. On the one hand, the environment provides a variety of factors required by the system to survive and develop, such as

materials, technology, information, capital elements. On the other hand, telecom companies export to the environment products and services and dynamically improve the environment through efforts, to create a more favorable environment for the development.

B. Collaborative Filtering

We may distinguish two broad categories of collaborative recommendation systems, namely content-based and collaborative filtering. A content-based system selects items based on the correlation between the content of the items (e.g. keywords describing the items, such as album genre, artists, etc., for music tracks) and the users' preferences [35]. However, it is limited to dictionary-bound relations between the keywords used by users and the descriptions of items and therefore does not explore implicit associations between users.

Collaborative filtering systems are divided into two categories, i.e. memory-based and model-based. In the memory-based systems we calculate the similarity between all users, based on their ratings of items using some heuristic measure such as the cosine similarity or the Pearson correlation score. Then we predict a missing rate by aggregating the ratings of the k nearest neighbours of the user we want to recommend to. The problem with memory-based systems is that we have to decide on a rather arbitrary basis over parameters such as the number of neighbours. What is more, in the case of social networks there is no straightforward way to introduce similarities between users based on friendships and social tagging, other than some way of ad hoc interpolation of similarity weights from those different sources.

The model-based filtering systems assume that the users build up clusters based on their similar behaviour in rating of items. A model is learned based on patterns recognised in the rating behaviours of users using clustering, Bayesian networks and other machine learning techniques. The problem with model-based methods is that it is necessary to fine-tune several parameters of the model as well as the fact that the models produced might not generalise well in radically different context. What is more, as in the case of memory-based systems extra effort and training needs to be done in order to introduce knowledge from social networks [18].

C. Clickthrough Data

User click-through data can be extracted from a large amount of search logs accumulated by web search engines. These logs typically contain user-submitted search queries, followed by the URL of Web pages which are clicked by users in the corresponding search result page. Although these clicks don't reflect the exact relevancy, they provide valuable indications to the users' intention by associating a set of query terms with a set of web pages. If a user clicks on a web page, it is likely that the web page is relevant to the query, or at least related to some extent. Many valuable applications have been proposed along this direction, such as query suggestion [3][4][5], query expansion [6], query clustering [7-8][14-15], web page summarization [12], web search results

optimizing[9-10][11][13] and conducting other interesting work [16-17].

III. PROPOSED APPROACH

A. User Profile

1) Sequence Score

Definition: A retrieval transaction is user's browsed sequence for search results, and is noted as *tran*.

A *tran* can record the accessed information of the search results after the user put the query strings in search engine. For example, if a user query "user model", the search result is a list in some order (eg: "page1, page2, page3, page4"), And the user's accessed sequence for the results is "page3, page2, page4", our model will capture a tuple "<SessionId, "user model", (page3, page2, page4)>" by analyzing the user's clickthrough data. The score of each item of the user's accessed sequence give the following evaluation equation,

$$\text{Score}(\text{sequence})_{\text{page}} = \frac{1}{N} \cdot \sum_{i=1}^N \text{Score}(\text{page})_i \quad (1)$$

$$\text{Score}(\text{page})_i = \frac{m-j+1}{m} \quad (2)$$

Where $\text{Score}(\text{page})_i$ is the web page score in the "SessionId" i , N is the number of the different sessions in which a user browsed the same web page. m is length of user's accessed sequence in the session, and j is page's position in the user's accessed sequence.

2) Web Page Rating

The score of Web page rating,

$$\text{Score}(\text{rate})_{\text{page}} = \frac{1}{n} \cdot \sum_{i=1}^n \text{rate}_i \cdot f(n) \quad (3)$$

Where n is the number of users who give relevance evaluation to the Web page, and rate_i is score value of relevance rate given by user i . $f(n)$ is an increasing function of parameter n ; the greater is the value of n , the more popular is the Web page.

3) Preference Score

User k preferences score for Web page p ,

$$\text{pref}(k, p) = \delta \cdot \text{Score}(\text{rate})_{\text{page}} + (1-\delta) \text{Score}(\text{sequence})_{\text{page}} \quad (4)$$

Where $\text{Score}(\text{rate})_{\text{page}}$ is the score of Web page rating and $\text{Score}(\text{sequence})_{\text{page}}$ is the score of user's browsed sequence. And δ is an impact factor, whose range is in $[0, 1]$. The user profile can be created as following Table I.

TABLE I

USER PROFILE

UserId	PageId	Rating	Time
58743	29086	0.7321	22:50:21 24-11-2010
89301	8329	0.6859	22:50:33 24-11-2010
6741	73429	1.2942	22:50:45 24-11-2010
...

B. User-based Collaborative Filtering

User-based collaborative filtering predicts a test user's interest in a test item based on rating information from similar user profiles [1][5][14]. Each user profile (row vector) is sorted by its dis-similarity towards the test user's profile. Ratings by more similar users contribute more to predicting the test item rating. The set of similar users can be identified by employing a threshold or selecting top- N . In the top- N case, a set of top- N similar users $S_u(u_k)$ towards user k can be generated according to,

$$S_u(u_k) = \{u_a \mid \text{rank } S_u(u_k, u_a) \leq N, x_{a,m} \neq \Phi\} \quad (5)$$

Where $|S_u(u_k)|=N$. $s_u(u_k, u_a)$ is the similarity between users k and a . Cosine similarity and Pearson's correlation are popular similarity measures in collaborative filtering, see e.g. [1][5]. The similarity could also be learnt from training data [9]. This paper adopts the Pearson's correlation similarity measure, comparing two user profiles by the Pearson's correlation of the similarity between the corresponding row vectors [21].

Consequently, the predicted rating $\text{pref}_{\text{rec}}(k, p)$ of test item p by test user k is computed as following,

$$\text{pref}_{\text{rec}}(k, p) = \sum_{j=1}^N \text{pref}(j, p) \bullet (s_u(u_k, u_j) + 1.0) \quad (6)$$

Where $\text{pref}_{\text{rec}}(k, p)$ is the recommendable Web page scoring value, N is the number of the top- N most similar users, and $\text{pref}(j, p)$ is preference value of user j for resource p . And $s_u(u_k, u_j)$ represents the similarity between user u and user j .

C. Personalized Search Model

According to selecting top-scoring documents from (6) and documents of interest to users including users accessed to and system predicted, we proposed a personalized Web search retrieval model which different users entering the same query keywords, the search results list is different. The model is described as following,

$$\text{Score}(q, p) = \begin{cases} (1-\alpha)\text{Sim}(q, p) + \alpha \cdot \text{pref}(k, p), & p \in \text{list}_{\text{pref}} \\ (1-\beta)\text{Sim}(q, p) + \beta \cdot \text{pref}_{\text{rec}}(k, p), & p \in \text{list}_{\text{rec}} \\ \text{Sim}(q, p), & p \notin \text{list}_{\text{pref}} \text{ and } p \notin \text{list}_{\text{rec}} \end{cases} \quad (7)$$

Where $\text{Score}(q, p)$ is the score of Web page p for query q , $\text{Sim}(q, p)$ is the similarity between Web page p and query q ; $\text{pref}(k, p)$ is the directly preference of given user k for Web page p , $\text{pref}_{\text{rec}}(k, p)$ is the predicted rating of given user k for Web page p ; α and β are impact factors of the $\text{pref}(k, p)$ and $\text{pref}_{\text{rec}}(k, p)$; $\text{list}_{\text{pref}}$ is the collection of user k explicitly interested documents and list_{rec} are the collection of user k implicitly predicted documents, and $\text{list}_{\text{pref}} \cap \text{list}_{\text{rec}} = \emptyset$.

If there is rating information for the user himself for the particular page, then the first case applies.

If there is at least rating information by similar users for the particular page, then the second case applies.

Otherwise, the third case applies, which does not make use of any rating information.

IV. EXPERIMENTAL SETUP

A. Data Sets

Clickthrough data can be recorded with little overhead and without compromising the functionality and usefulness of the search engine. In particular, compared to explicit user feedback, it does not add any overhead for the user. The query q and the returned ranking r can easily be recorded whenever the resulting ranking is displayed to the user. For recording the clicks, a simple proxy system can keep a logfile [10]. In this paper, we collect clickthrough data by using a proxy server of Web server side. The data include user login information, query string, the Web page id, Session Id, clicked sequence of search results, and the visiting time. The Table II describes the data information.

The Web page rating can be explicitly recorded after user browsed the page, and meanwhile provided relevance score to it. The range of relevance score is from 0 to 1, and includes 0 and 1. If a user thought a page browsed was not relevant, he/she could give relevance score of the page to 0. On the contrary, he/she could give relevance score of the page to a number greater than 0.

Our experiments are performed on the China Education Television (CETV) Learning Mall Resource Set, which contains 312,477 pieces of resource, and uploaded, by 5,664 resource producers. We have 165,379 users for our system, and get 130,452 rating records.

In our experiments system, we trace user's searching and browsing activity, and to update user's interest, then we provide personalized Web search to users according to their preferences.

B. Evaluation Metrics

We evaluate the ranking algorithms over a range of accepted information retrieval metrics, namely *Precision at K* ($P(K)$) and *Mean Average Precision* (MAP). Each metric focuses on a different aspect of system performance, as we describe below [9].

Precision at K: As the most intuitive metric, $P(K)$ reports the fraction of documents ranked in the top K results that are labeled as relevant. In our setting, we require a relevant document to be labeled "Good" or higher. The position of relevant documents within the top K is irrelevant, and hence this metric measures overall user satisfaction with the top K results.

MAP: Average precision for each query is defined as the mean of the precision at K values computed after each relevant document was retrieved. The final MAP value is

defined as the mean of average precisions of all queries in the test set. This metric is the most commonly used single-value summary of a run over a set of queries.

C. Ranking Methods Compared

BM25F: As a strong web search baseline we used the BM25F scoring, which was used in one of the best performing systems in the TREC 2004 Web track [12, 13]. BM25F and its variants have been extensively described and evaluated in IR literature, and hence serve as a strong, reproducible baseline. The BM25F variant we used for our experiments computes separate match scores for each "field" for a result document (e.g., body text, title, and anchor text), and incorporates query-independent link based information (e.g., PageRank, ClickDistance, and URL depth). The scoring function and field-specific tuning is described in detail in [12]. Note that BM25F does not directly consider explicit or implicit feedback for tuning.

BM25FP: The ranking produced by incorporating clickthrough statistics and Web page rating to reorder web search results ranked by BM25F above.

Lucene: Apache Lucene [30] is a high-performance and full-featured text search engine library written entirely in Java. It is a technology suitable for nearly any application that requires full-text search. Lucene is scalable and offers high-performance indexing, and has become one of the most used search engine libraries in both academia and industry. Lucene ranking function, the core of any search engine applied to determine how relevant a document is to a given query, is built on a combination of the Vector Space Model (VSM) and the Boolean model of Information Retrieval. The main idea behind Lucene approach is the more times a query term appears in a document relative to the number of times the term appears in the whole collection, the more relevant that document will be to the query. Lucene uses also the Boolean model to first narrow down the documents that need to be scored based on the use of Boolean logic in the query specification.

LuceneP: The ranking produced by reordering the Lucene results using clickthrough statistics and Web page rating.

D. Users Evaluation

We use user's browsing sequence and page turning activity to test the accuracy of the search results list in our model, which is also called users evaluation.

The higher is the similarity of the search results sequence and user's browsing sequence; the higher is the

TABLE II
INFORMATION FORMAT OF CLICKTHROUGH DATA

ID	Query	PageId	Rank	UserId	SessionId	Time
1	User model	47806	3	58743	8232328228986249	9:21:43 24-11-2010
2	User model	38570	4	58743	8232328228986249	9:22:15 24-11-2010
3	User model	29086	6	58743	8232328228986249	9:22:15 24-11-2010
4	Web search	8329	2	89301	1923744500763862	9:24:27 24-11-2010
5	Personalized search	73429	1	6741	2785098742726650	9:30:36 24-11-2010
...

retrieval precision.

The sequence of the search results list is $\text{vector1} = \langle (1, \frac{m-1+1}{m}), (2, \frac{m-2+1}{m}), \dots, (i, \frac{m-i+1}{m}), \dots, (m, \frac{m-m+1}{m}) \rangle$, where m represents the resource number in the recommended results list, i represents the i -st resource in the list.

User's browsing sequence is $\text{vector2} = \langle (1, 0), (2, \frac{n-1+1}{n}), (3, \frac{n-2+1}{n}), \dots, (i, \frac{n-j+1}{n}), \dots, (m, \frac{n-k+1}{n}) \rangle$, where n represents the user-browsed resource number, j represents the j -st resource in the browsing list, k represents the k -st resource in the browsing list.

For example, the search results list is $\text{item1}, \text{item2}, \dots, \text{itemi}, \dots, \text{item9}, \text{item10}$, and m is 10 here.

And the user-browsed resource list is $\text{item2}, \text{item4}, \text{item6}, \text{item8}, \text{item7}$, and n is 5. Then we get $\text{vector1} = \langle (1, 1), (2, 0.9), (3, 0.8), (4, 0.7), (5, 0.6), (6, 0.5), (7, 0.4), (8, 0.3), (9, 0.2), (10, 0.1) \rangle$; $\text{vector2} = \langle (1, 0), (2, 1), (3, 0), (4, 0.8), (5, 0), (6, 0.6), (7, 0.2), (8, 0.4), (9, 0), (10, 0) \rangle$. We now use the laws of cosines to calculate the similarity between the two vectors. The greater is the similarity value, the higher is the retrieval precision.

We compare the search results between our information retrieval model (BM25FP) and the base information retrieval model (BM25F).

Fig.1 shows the sequence similarity between BM25FP and BM25F. The sequence similarity of the BM25FP is much better than the BM25F from the Fig.2 because the BM25F is only related to the similarity of query and document without considering user's preference.

E. Impact of Parameters

Recall the two parameters in (7): α balance the scores between the query similarity and user's preference score, and β balances the scores between the query similarity and the predicted rating, we first test the sensitivity of α , setting β to zero. This scheme counts directly on the user preference score, but does not use user-based collaborative filtering prediction. Fig. 2 shows web search MAP against varying α from zero (a pure information retrieval model) to one (a user preference score approach). The value of the optimal α demonstrates that interpolation between pure information retrieval model and user preference score approaches improves the Web search performance. More specifically, the best results are obtained with α around 0.4.

Fig. 3 shows the sensitivity of β after fixing α to 0.4. The graph plots the MAP when parameter β is varied from zero (a pure information retrieval model approach) to one (the predicted rating approach). We observe that β reaches its optimal in 0.2.

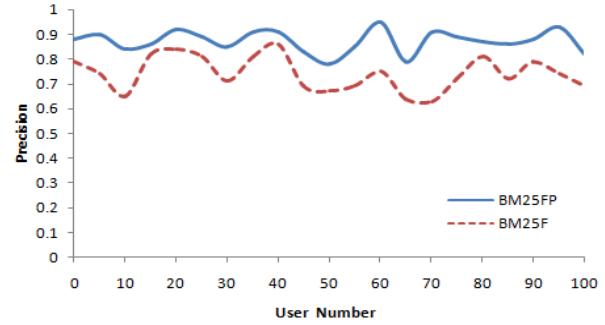


Figure 1. Sequence similarity between two models

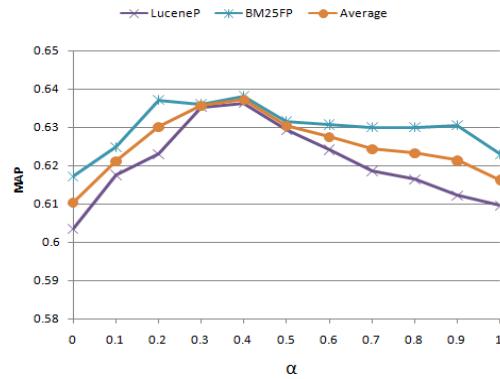


Figure 2. Impact of the parameter α

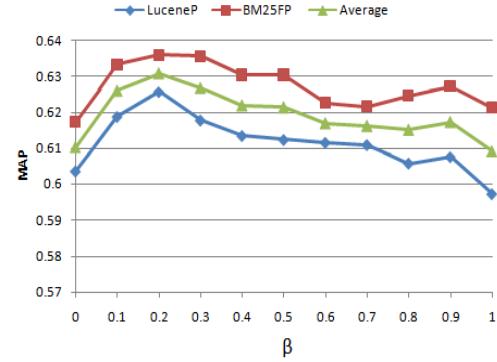


Figure 3. Impact of the parameter β ($\alpha=0.4$)

Additional experiments (not reported here) verified that there is little dependency between the choice of α and the optimal value of β . The optimal parameters can be identified by using the cross validation from the training data.

F. Personalized Search Performance

We continue with a comparison to results obtained with other methods, setting α to 0.4 and β to 0.2. We first compare our results (BM25FP) to the standard BM25F. We report results for test the precision at 5, 10 and 20. The first two rows of Table III summarize the results, showing the performance of the BM25FP is better than the BM25F. Next, we first compare the LuceneP to the standard Lucene with the same condition. The last two rows of Tab.III summarize the results, showing the performance of the latter is better than the former too.

TABLE III
PRECISION COMPARISONS

	P@5	P@10	P@20	MAP
BM25F	0.80	0.76	0.72	0.6035
BM25FP	0.88	0.85	0.80	0.7207
Lucene	0.88	0.72	0.70	0.6172
LuceneP	0.88	0.82	0.78	0.7178

V. CONCLUSION AND FUTURE WORK

This article has proposed a personalized Web search model based on the method which calculates users' preferences according to the user's search behaviors and resource properties. This model has fully used the information in these two areas, does not need the user to make the appraisal when he or she glances over information, the system will analyze and quantize user's behaviors automatically. According to the user model which formerly established, this article simultaneously proposed the resources filtering and recommendation algorithm, which was based on Top-N resource recommending method.

ACKNOWLEDGMENT

This work was supported by Program for New Century Excellent Talents in University, China (NCET-06-0161, 1110012040112), Fok Ying Tong Education Foundation China(91101), BIT Major Foundational Research Project Supported (3070012231001), Beijing Institute of Technology graduate student scientific and technological innovation project.

REFERENCES

- [1] J. T. Sun, H. J. Zeng, H. Liu, Y. Lu, and Z. Chen, "Cubesvd: a novel approach to personalized web search," Proceedings of the 14th international conference on World Wide Web, Chiba, Japan, May 10-14, 2005.
- [2] Y. E. Ioannidis and G. Koutrika, "Personalized systems: models and methods from an IR and DB perspective.", Proceedings of the 31st International Conference on Very Large Data Bases, Trondheim, Norway, August 30 - September 2, 2005.
- [3] C. Huang, L. Chien, and Y. Oyang, "Relevant term suggestion in interactive web search based on contextual information in query session logs," JASIST 54(7): 638-649,2003.
- [4] N. J. Belkin, "Helping people find what they don't know," Communications of the ACM, No.8, 2000, pp58-61.
- [5] H. Ma, H. Yang, I. King, and M. R. Lyu, "Learning latent semantic relations from clickthrough data for query suggestion," Proceedings of the 17th ACM Conference on Information and Knowledge Management, Napa Valley, California, USA, October 26-30, 2008.
- [6] H. Cui, J.R. Wen, J.Y. Nie, and W.Y. Ma, "Query expansion by mining user logs," IEEE Transaction on Knowledge and Data Engineering, Vol. 15, No. 4, July/August 2003.
- [7] D. Beeferman and A. Berger, "Agglomerative clustering of a search engine query log," Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining, , Boston, MA, USA, August 20-23, 2000.
- [8] J.R. Wen, J.Y. Nie, and H.J. Zhang, "Clustering user queries of a search engine," Proceedings of the Tenth International World Wide Web Conference, Hong Kong, China, May 1-5, 2001.
- [9] E. Agichtein, E. Brill, and S. Dumai, "Improving web search ranking by incorporating user behavior information," Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, Washington, USA, August 6-11, 2006.
- [10] T. Joachims, "Optimizing search engines using clickthrough data," Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton, Alberta, Canada, July 23-26, 2002.
- [11] T. Joachims and F. Radlinski, "Search engines that learn from implicit feedback," Computer, No.8, 2007, pp34-40.
- [12] J. T. Sun, D. Shen, H. J. Zeng, Q. Yang, Y. Lu, and Z. Chen, "Web-page summarization using clickthrough data," Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Salvador, Brazil, August 15-19, 2005.
- [13] X. Wang and C. Zhai, "Learn from web search logs to organize search results," Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Amsterdam, The Netherlands, July 23-27, 2007.
- [14] J. R. Wen, J. Y. Nie, and H. Zhang, "Query clustering using user logs," ACM Trans. Inf. Syst., No.1, 2002, pp59-81.
- [15] D. Beeferman and A. Berger, "Agglomerative clustering of a search engine query log," Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining, Boston, MA, USA, August 20-23, 2000.
- [16] M. Pasca and B. V. Durme, "What you seek is what you get: Extraction of class attributes from query logs," Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007.
- [17] D. Shen, M. Qin, W. Chen, Q. Yang, and Z. Chen, "Mining web query hierarchies from clickthrough data," Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence, Vancouver, British Columbia, Canada, July 22-26, 2007.
- [18] I. Konstas, V. Stathopoulos, and J. M. Jose, "On social networks and collaborative recommendation," Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Boston, MA, USA, July 19-23, 2009.
- [19] P.A. Chirita, C.S. Firin, and W. Nejdl, "Personalized query expansion for the web," Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Amsterdam, The Netherlands, July 23-27, 2007.
- [20] J. Wang, A. P. de Vries and M. J. T. Reinders, "Unifying user-based and item-based collaborative filtering approaches by similarity fusion," Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, Washington, USA, August 6-11, 2006.
- [21] R. B. Almeida and V. A. F. Almeida, "A community-aware search engine." Proceedings of the 13th international conference on World Wide Web, WWW 2004, New York, NY, USA, May 17-20, 2004.
- [22] F. Liu, C. Yu, and W. Meng, "Personalized web search by mapping user queries to categories," Proceedings of the International Conference on Information and Knowledge Management, McLean, VA, USA, November 4-9, 2002.
- [23] L. Fitzpatrick and M. Dent, "Automatic feedback using past queries: social searching?," Proceedings of the 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Philadelphia, PA, USA, July 27-31, 1997.

- [24] G.R. Xue, H.J. Zeng, Z. Chen, Y. Yu, W.Y. Ma, W.S. Xi, and W.G. Fan, "Optimizing web search using web clickthrough data," Proceedings of the International Conference on Information and Knowledge Management, Washington, DC, USA, November 8-13, 2004.
 - [25] Z. Dou, R. Song, J. Wen, "A large-scale evaluation and analysis of personalized search strategies," Proceedings of the 16th International Conference on World Wide Web, Banff, Alberta, Canada, May 8-12, 2007.
 - [26] F. Qiu, and J. Cho, "Automatic identification of user interest for personalized search," Proceedings of the 15th international conference on World Wide Web, Edinburgh, Scotland, UK, May 23-26, 2006.
 - [27] B. Tan, X. Shen, C. Zhai, "Mining long-term search history to improve search accuracy," Proceedings of the Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Philadelphia, PA, USA, August 20-23, 2006.
 - [28] J. Teevan, S.T. Dumais, and E. Horvitz, "Personalizing search via automated analysis of interests and activities," Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Salvador, Brazil, August 15-19, 2005.
 - [29] J. Teevan, E. Adar, R. Jones, and M. Potts, "Information re-retrieval: Repeat queries in yahoo's logs," Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Amsterdam, The Netherlands, July 23-27, 2007.
 - [30] O. Gospodnetic, and E. Hatcher, Lucene in action, Manning Publications Co.2005 .
 - [31] K. Sugiyama, K. Hatano, and M. Yoshikawa, "Adaptive web search based on user profile constructed without any effort from users," Proceedings of the 13th international conference on World Wide Web, New York, NY, USA, May 17-20, 2004.
 - [32] F. Qiu and J. Cho, "Automatic identification of user interest for personalized search," Proceedings of the 15th international conference on World Wide Web, WWW 2006, Edinburgh, Scotland, UK, May 23-26, 2006.
 - [34] T. Haveliwala, "Topic-sensitive pagerank," Proceedings of the Eleventh International World Wide Web Conference, WWW2002, Honolulu, Hawaii, USA, 7-11 May 2002.
 - [35] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: Bringing order to the web," Technical report, Stanford University, 1998.
- Xueping Peng** Anhui Province, China. Birthdate: October, 1980. is a Computer Software and Theory Ph.D. student in the School of Computer Science and Technology at Beijing Institute of Technology. And research interests on information retrieval, personalized recommendation, web data mining.
- Zhendong Niu** Anhui Province, China. Birthdate: October, 1968. is Computer Software and Theory Ph.D., graduated from Dept. Computer Science and Technology at Beijing Institute of Technology. And research interests on computer software architecture, knowledge management, intelligent education software system, digital library, neural information system. He is a professor in the School of Computer Science and Technology at Beijing Institute of Technology, China.
- Sheng Huang** Henan Province, China. Birthdate: June, 1986. is a Computer Software and Theory Ph.D. student in the School of Computer Science and Technology at Beijing Institute of Technology. And research interests on text mining and opinion mining.
- Yumin Zhao** Henan Province, China. Birthdate: June, 1980. is a Computer Software and Theory Ph.D. student in the School of Computer Science and Technology at Beijing Institute of Technology. And research interests on digital library, data mining, computer network.

Study on Learner Modeling in Adaptive Learning System

Bing Jia

Jilin University, Changchun, 130012, China

Email: bing_jia@qq.com

Yongjian Yang and Jun Zhang

Jilin University, Changchun, 130012, China

Email: {yyj, zhangjun10}@jlu.edu.cn

Abstract—In light of the present situation research on learner model in Adaptive Learning System at home and abroad, This paper summarizes the main problems of learner model, such as the unscientific attention dimension, poor calculated representation method and single and subjective method to obtain Characteristic value. So this paper has put forward a new learner model to achieve self-organization and push of learning resources on the basis of learning goals and learner's personal conditions. It includes three characteristic items such as knowledge level, cognitive ability and preferences. We respectively introduce the representation and acquisition of characteristic value of learner model in detail. After that, we propose a new Adaptive Learning System based on this learner model as a case. Experimental results show this learner model is effective and practical in the application.

Index Terms—Adaptive Learning System, Advanced Learner Model, characteristic, representation, acquisition

I. INTRODUCTION

Learner Model(LM for short) is the representation of the learners' cognitive states[1]. It reflects the learners' learning process, knowledge proficiency, misconceptions and the gap between the desired goals[2]. In adaptive learning system, learner model is a single input system[3]. Currently there are three representative researches on it. We analysis and summarize these models from different prospects which are focusing on dimension, representation and acquiring method[4] in Table I. We discover three common problems exist in the learner model.

Beginning with the prospective of focusing on dimension, it only concerns about the learners' interests, psychological characteristics etc, however ignores such features as the learners' original knowledge level and cognitive ability which are relevant to the learning resources' recommending.

Then, from the prospective of representation method, it generally uses the qualitative description method which is inconvenient in calculating and reasoning.

In the end, from the prospective of acquiring method, it all uses single subjective methods such as questionnaire

TABLE I.
COMPARISON OF THE CLASSICAL MODEL

Model View	Cross's	Kaye's	DingXingfu's
Dimensions	Environmental Individual	Education Life Style	Cognitive Psychological
Representation	Qualitative	Qualitative	Qualitative
Acquisition	Questionnaire	Questionnaire Test	Measurement Statistics

and scale, which made the acquisitive characteristics inaccurate.

In this article, in order to provide effective support to adaptive learning, we mainly put our focus on the research of the three above problems, aiming at researching the basic components and the relevant representation and acquiring method of the learner model.

In the view of the above mentioned analysis, in this study, we propose a new LM called Advanced Learner Model (ALM for short) that can achieve self-organization and push of learning resources according to learning goals and learner's personal situation. Construct ALM by knowledge level module, cognitive ability module and preferences module. In this study, analysis the theoretical basis of LM an in section 2, introduce the representation method of ALM in section 3, give the acquisition of characteristic value of learner model of ALM in section 4, propose a new Adaptive Learning System based on this learner model as a case to explore the related technologies to achieve the model and the application in section 5 and summarize the study and forecast the future in section 6.

II. THE RELATED THEORETICAL BASIS

LM refers to produce one reliable way of expression to demonstrate what the student understood and could do, what he does not understand and could not do, what he wants to do as well as he should do[5]. The study of LM is based on ALS, also is the key link in ALS.

A. The Adaptive Learning System

The basic idea of Adaptive Learning System is a student-centered, and a targeted study guide accord to the

student mastery and acceptance of knowledge. It has the following characteristics.

Adaptability: The system automatically provide the learners with the most suitable knowledge to learn, through interact with the learners and recognize the knowledge level and characteristics of the learners.

Autonomy: learners have the right to choose whether to participate in the system provided learning activities or not. And also can choose the way or strategy to participate in the learning process.

Resources construction: system constructs the appropriate learning resources and makes it adapt to the learners' all kinds of needs, which is based on the learners' acquired knowledge.

B. Theory Basis of Learner Model

ALM is an abstract description and representation of learner's characteristics information[6][7]. So how to determine the characteristics of ALM? First, Gagne[8] believes that any kind of learning new knowledge and skills are based on the knowledge already learned and subordinated to. Second, learner's cognitive capacity reflects his study ability (Speed, the way to accept knowledge, etc). Clearly, the premise of effective learning activities is to understand learner's knowledge level and cognitive ability. In addition, the learner's learning is driven by motives, we should pay attention to understand the learner's interests and preferences information to promote the learning activities of learners. In this paper, we construct ALM from three characteristics of learner, such as knowledge level, cognitive ability and preferences.

III. REPRESENTATION OF ALM

In order to provide adaptive learning services for learners in ALS, LM must have functions as follows:

First, estimates learner's knowledge level and cognitive ability through test, then uses them as the basis for pushing learning content.

Second, provides personalized guidance for the learner in accordance with learner's needs, interests and habits to stimulate interest in learning[9].

Finally, renewals LM by mining and extraction to the information in learner's learning process and test process.

So, we propose a new LM called Advanced Learner Model (ALM for short) consist of knowledge level module, cognitive ability module and preferences module. We respectively introduce the representation

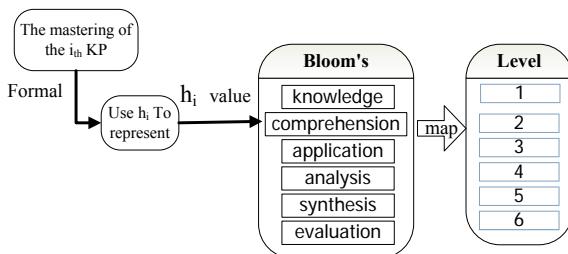


Figure 1. The representation of KP mastering

methods of the three modules in detail in the following sections.

A. Knowledge Level

Knowledge Level refers to a set of learner's existing knowledge under a certain goal. It has two parameters, one is knowledge point(KP), the other is mastering. We use k_i to represent the i_{th} (KP), and we use h_i to represent learner's mastering of the i_{th} KP. For the mastering, when learner have no idea of the KP, we use "0" to represent the situation. For others, we divide the mastering into six levels based on Bloom's taxonomy of cognitive objectives, such as knowledge, comprehension, application, analysis, synthesis, evaluation, mapping with 1~6, as shown in Fig.1.

To sum up, we use "Knowledge-how" to represent the set. Definition 1 : Knowledge-how = $\{(k_1, h_1), \dots, (k_n, h_n)\}$, where k_i represents the i_{th} knowledge point(KP), n represents the total of KPs, h_i represents learner's mastering of the i_{th} KP, $h_i \in H$, $H=\{0,1,2,3,4,5,6\}$, where "0" represents the situation that the learner have no idea of the KP, 1~6 respectively represents educational objectives proposed by Bloom[10](knowledge, comprehension, application, analysis, synthesis , evaluation).

B. Cognitive Ability

Cognitive Ability(CA) is individual's ability possessed while reconstructing and employing knowledge[11]. We can describe the level of learners' ability by depicting the cognitive status of learners. It also has two parameters, one is the type of CA, The other is the level value of CA. The type of CA should be divided into eight capability based on Gardner's multiple intelligences theory of cognitive abilities such as Inductive capacity, memory, observation, abstraction ability, analytical ability, calculation ability, imagination, and logical reasoning ability. And in turn be expressed with a_1 to a_8 ,as Fig.2 shows. The level value can be represented by l_i . As the ability to change is continuous, l_i is defined as the value greater than 0 and less than 1.

We use Ability-level to represent the set of learner's CA, so, Definition 2 : Ability-level = $\{(a_1, l_1), \dots, (a_n, l_n)\}$, where a_i is the i_{th} CA, l_i is the level of the i_{th} CA, n is the total of CAs.

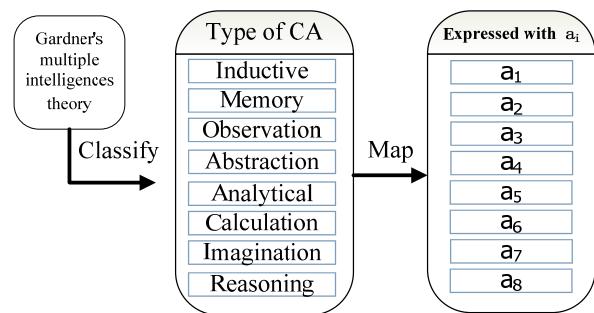


Figure 2. The representation of cognitive type

C. Preferences

Preferences are defined as demonstrated interest, hobby and other information in learning. Through the analysis, we preferred preferences information into the background material preferences, learning strategy preferences, learning time preferences, system features preferences and presentation preferences. In turn is expressed as P_b , P_s , P_t , P_f , P_p . They are five concept sets, A concept is a pair of (c, σ) , where c is a concept, and σ is a coefficient denotes the preference degree of the concept, as Fig.3 shows. For the value, take the presentation of resources for example, it can be taken the value of a set of $\langle \text{text}, \sigma_1 \rangle \langle \text{flash}, \sigma_2 \rangle$, as Fig.4 shows.

Use a five-tuple “Preferences-set” to represent it, so Definition 3: Preferences-set= $\langle P_b, P_p, P_s, P_t, P_f \rangle$, where P_b represents background material preferences(sport, entertainment, education, etc.), P_p represents resource presentation preferences (text, audio, animation, video), P_s represents learning strategies preferences(teaching type, inquisition type, cooperation type), P_t represents study time preferences(time period), P_f represents system function preferences(using the system frequency-related features). P_b , P_p , P_s , P_t , P_f are five concept sets, A concept is a pair of (c, σ) , where c is a concept, and σ is a coefficient denotes the preference degree of the concept.

IV. ACQUISITION OF CHARACTERISTIC VALUE OF ALM

We acquire the characteristic values of the learner model through analyzing the learners' characteristics. In traditional teaching process, through observation, survey and communication, teachers would obtain the learners' characteristics which are required in the teaching process. However, under the network environment, the learners and system are invisible to each other, which made the face to face communication impossible. So it could only be done through the analysis and process of the captured

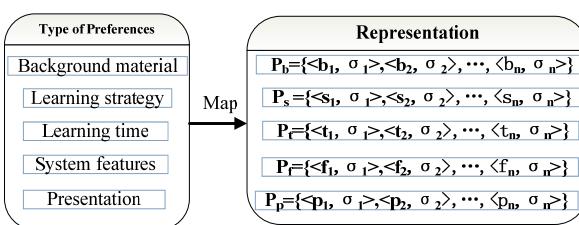


Figure 3. The representation of preferences type

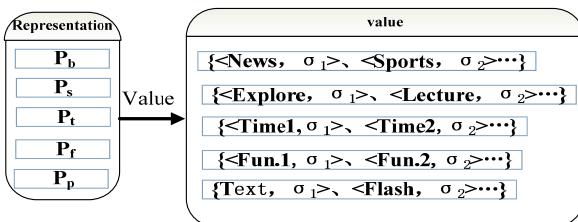


Figure 4. The representation of preferences value

information by computers. Thus, the conception of learners' characteristic analysis is with the proper use of computer technology to acquire the learner model's characteristic values, under the guidance of some reasoning methods[12].

A. Estimation method of Knowledge Level

1) Access to the set of KPS

For the set of knowledge points, we give the acquiring flow chart as shown in the Fig.5. The basic idea is to find its precursor according to the set of goal knowledge points, and estimate whether the value of the learners' master degree of these precursors is “puzzle”, if not, end the process and record this knowledge point. if so, carry on finding this knowledge point's precursor.

2) Estimate learner's mastering of a KP

There are two effective methods of reasoning such as Bayes theorem and fuzzy logic. This paper mainly introduces the application of fuzzy logic in estimating the degree of learner's mastering of a certain KP[13]. fuzzy logic is a logic based on the concept of fuzzy set, in which membership is expressed in varying probabilities or degrees of truth—that is, as a continuum of values ranging from 0 (does not occur) to 1 (definitely occurs). In the process of this study, we used the research methods to push down, the first step to determine the representation of the mastery of by fuzzy set, the second step gives changes rules of the membership in fuzzy set, and finally gives the expression of the parameter q in the rules. So you can reverse the process to achieve the mastery of KP.

According to Definition 1, we use these seven ranks (1~6) to express the degree of learner's mastering of a certain KP. The membership of each rank is expressed as $\mu_k(i)$, $i=1\sim 6$. Express the degree of learner's mastering of a certain KP by fuzzy set K, so $K=\{i | (i, \mu_k(i)), i \in H\}$, where, $0 \leq \mu_k(i) \leq 1$, $\sum \mu_k(i) \leq 1$. To a KP, the highest level of membership is the level that the learner has achieved currently about this KP.

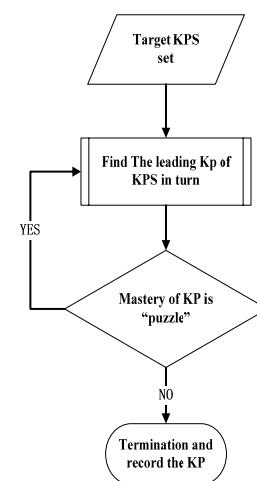


Figure 5. Access to the set of KPS

For example, assume that the membership of the six ranks are $\{0.1, 0.2, 0.4, 0.2, 0.1, 0\}$, therefore, the degree of learner's mastering of a certain KP are expressed as $K=\{(1, 0.1), (2, 0.2), (3, 0.4), (4, 0.2), (5, 0.1), (6, 0)\}$. We can see that the level of membership for the rank 3 is the highest. Then can estimate the degree of learner's mastering of a certain KP is the third rank.

We definite the changing rules of membership of fuzzy set "K" as Table II, assume "m" is the rank of test questions, $i=1 \sim 6$.

The basic principle of changing rule: when learner does a certain rank of test question correctly, adopt up rule to the rank and the following, and adopt down rule to the above. When learner do test question wrongly, the membership does not change. Take table III for example, here supposes parameter $q=0.5$, we can infer that the degree of mastering about the KP is level 2 from table III. The learner does the third question in rank 2 correctly, but he does the fourth question in same rank wrongly, we may surmise that he possibly guesses the third question right or this question is possibly neglectful. When does the fifth question in the same rank, he is also wrong. It seems more likely that he may not reach rank 3, the system declines the rank of the question from rank 3 to rank 2. As a result of this algorithm can exclude some interference information, such as guess right and negligence. It can classify the knowledge level of learners better.

From the "down rule" in Table III, using " Δ " to represent the variation of $\mu_k(i)$, $\Delta=(\mu_k(i+1)-\mu_k(i))q$, so parameter "q" decides the scope of membership changing , It decides the declining speed of learner's mastering of KP. The determination of parameter "q" should consider the following two aspects: the difficulty of the question and the degree of learner's familiarity with the question. The difficulty of the question is expressed as "d", where $0 < d < 1$.the degree of learner's familiarity with the question will be obtained by the speed of the learner to answer the question correctly. We stipulate two time sections: normal time and longest time.

If the learner answer the question correctly in normal time range, indicate that he is familiar with the knowledge. If he answers the question correctly between the normal time and the longest time, indicate that he is not very familiar with it. If beyond the longest time, indicate that he is not familiar with it. Then, use the function to express the familiar degrees with the questions ,we can obtain the definition 4.

Definition 4 : assume that α is the normal time, γ is the longest time, $\beta=(\alpha+\gamma)/2$.The function of the degree of learner's familiarity with the question $F(t)$ is defined as:

TABLE III.
CHANGING RULES OF MEMBERSHIP

If	Rule	Type
$i < m$	$\mu_k(i) = \mu_k(i) - \mu_k(i)q + \mu_k(i-1)q$	up rule
$i = m$	$\mu_k(i) = \mu_k(i) + \mu_k(i-1)q + \mu_k(i+1)q$	up rule
$i > m$	$\mu_k(i) = \mu_k(i) - \mu_k(i)q + \mu_k(i+1)q$	down rule

TABLE II.
THE EXAMPLE OF CHANGING RULES OF MEMBERSHIP

Initial $\mu=\{1, 0, 0, 0, 0, 0\}$	$q=0.5$
1 $\mu=\{0.5, 0.5, 0, 0, 0, 0\}$	Do rank 2 test correctly
2 $\mu=\{0.25, 0.75, 0, 0, 0, 0\}$	Do rank 2 test correctly
3 $\mu=\{0.125, 0.5, 0.375, 0, 0, 0\}$	Do rank 3 test correctly
4 $\mu=\{0.125, 0.5, 0.375, 0, 0, 0\}$	Do rank 3 test wrongly
5 $\mu=\{0.0625, 0.75, 0.1625, 0, 0, 0\}$	Do rank 3 test wrongly
6 $\mu=\{0.0625, 0.75, 0.1625, 0, 0, 0\}$	Do rank 2 test correctly

(where α, γ of each question can be different).

From the "down rule" in Table III, using " Δ " to represent the variation of $\mu_k(i)$, $\Delta=(\mu_k(i+1)-\mu_k(i))q$, so parameter "q" decides the scope of membership changing , It decides the declining speed of learner's mastering of KP. The determination of parameter "q" should consider the following two aspects: the difficulty of the question and the degree of learner's familiarity with the question. The difficulty of the question is expressed as "d", where $0 < d < 1$.the degree of learner's familiarity with the question will be obtained by the speed of the learner to answer the question correctly. We stipulate two time sections: normal time and longest time. If the learner answer the question correctly in normal time range, indicate that he is familiar with the knowledge. If he answers the question correctly between the normal time and the longest time, indicate that he is not very familiar with it. If beyond the longest time, indicate that he is not familiar with it. Then, use the function to express the familiar degrees with the questions, we can obtain the definition 4.

Definition 4: assume that α is the normal time, γ is the longest time, $\beta=(\alpha+\gamma)/2$.The function of the degree of learner's familiarity with the question $F(t)$ is defined as: (where α, γ of each question can be different).

$$F(t) = 1 - S(t; \alpha, \beta, \gamma) = \begin{cases} 1 & \text{当 } t \leq \alpha \\ 1 - 2\left(\frac{t-\hat{\alpha}}{\gamma-\alpha}\right)^2 & \text{当 } \alpha \leq t \leq \beta \\ 2\left(\frac{t-\gamma}{\gamma-\alpha}\right)^2 & \text{当 } \beta \leq t \leq \gamma \\ 0 & \text{当 } \gamma \leq t \leq \gamma \end{cases}$$

So that $q = c^d * F(t)$, c is a constant larger than 1, $0 < d < 1$. When the learner does the question correctly in the normal time range, $F(t)=1$, q become large. When the learner does the question correctly exceed the normal range, $F(t)$ becomes small, q become small too. It proves that the learner's mastering degree of KP is changing slower When exceeds the longest time, $F(t)=0$, $q=0$. Similarly, q increases with increasing d . When the learner does a difficulty bigger question correctly, It proves he can learn this content well, also proves his mastering degree of KP is rising faster.

So, the steps to determine "Knowledge-how"

a) Obtains a set of learner's goal knowledge:
Target-knowledge= $\{t_1, t_2, \dots, t_m\}$.

b) Discovers the pre-KPs corresponding with the m KPs in the "Target-knowledge" set according to domain model. The set of pre-KPs: Pre-knowledge={p₁,p₂,...,p_n}.

c) Judges the learner's mastering of pre-KPs one by one (marked with 0~6). so the set of learner's pre-KPs level:Pre-Knowledge-how={(p₁,h₁),..., (p_i,h_i),..., (p_n,h_n)}, h_i∈H.

d) Records KPs whose mastering degree between 1 and 6 in "Knowledge-how". Set p_j as target-KPs and turn to step b when h_j=0.

e) Achieves the disorder set "Knowledge-how".

B. Estimation method of Cognitive Ability

We can obtain the learner's CA by testing. We design a form of every test is : Test(i) = (A(i),T(i),η , λ , Φ:Q), A series of test questions constituted a test paper. TEST= {t_i= (A₁,T₁, η₁, λ₁, Φ₁ : Q₁), ..., t_n= (A_n,T_n, η_n, λ_n, Φ_n:Q_n)}. For example, in order to obtain the level of induction ability and memory ability. The series of test questions are as follows:

$$\text{Test}(1)=(A(1), T_1, \eta_1, \lambda_1, \Phi_1 : Q_1).$$

$$\text{Test}(2)=(A(1), T_2, \eta_2, \lambda_2, \Phi_2 : Q_2).$$

$$\text{Test}(3)=(A(2), T_3, \eta_3, \lambda_3, \Phi_3 : Q_3).$$

$$\text{Test}(4)=(A(2), T_4, \eta_4, \lambda_4, \Phi_4 : Q_4).$$

Then obtain the answers of the test questions. The set of answers : Answer =Answer ={answer (t₁, φ'₁), ..., answer (t_i, φ'_i), ..., answer (t_n, φ'_n)}. So, we can definite the level of the i_{th} CA as follow: Ability-how (A_i)=

$$\sum_{j=1}^n \left[\left(\frac{\phi_j \cap \phi'_j}{\phi_j} \right) \times \eta_j \times \lambda_j \right] / \sum_{j=1}^n (\eta_j \times \lambda_j)$$

Where (φ' _j ∩ φ _j)/φ _j is the accuracy rate of the learner's answers to the j_{th} question, using R_j to express the accuracy rate, so 0≤ R_j≤ 1. n is the total of the questions which the student replied to test the i_{th} kind of ability. Let l_i =Ability-how (A_i), and we can obtain the set of learner's CA "Ability-level".

C. Estimation method of Preferences

For the estimation of preferences, it divided into two steps which are initial estimation and dynamic estimation. We have to pretest the learners' interests before them using the system for learning and also initialize the learners' interests. With the proceed of learners' learning activities, we can discover the learners' interests then correct and maintain those values through the data mining about the data of learners' searching concepts, browsing websites' types and topics of discussions. For the initializing values, we can gather users' interest information through users' registration forms or scale, and use direct or indirect matching methods processing those original data to acquire the initial values. For dynamic values, it concern about the learners' learning procedure information and historical testing information. Then process the acquired values through the data mining model to obtain the needed data.

In addition, when using the characteristic value, we can build the vectors about the feature item of learning objective and the relevant weight to present the learning resources according to the 5D vector space consists of learners' interests' five factor group. Based on the angle of the feature vectors in the vector space to determine the similarity between the learner's interest information and the learning resource, and return the most similar learning resource to the learner.

V. A CASE STUDY BASED ON ALM

In order to verify the performance of ALM, this paper presents a new adaptive learning system based on ALM (short for NalsALM), to complete the case study for ALM.

A. The Framework Design of NalsALM

In this study, NalsALM is mainly constituted by two models (learner model, domain model), three processor(inference engine(IE), resources recommendation engine(RRE), information collection and extraction engine(ICEE)) and four database (knowledge base, test base, resources base, learner information base). In this system, domain model describes the relationship between the structure of knowledge points. ALM is core. All knowledge push plan are based on domain model and ALM. Learner need to choose learning goals if he uses the system for the first time. Based on the learner's goal, ALM chooses knowledge points as goal knowledge points set in the knowledge base, using "Target-knowledge" to represent

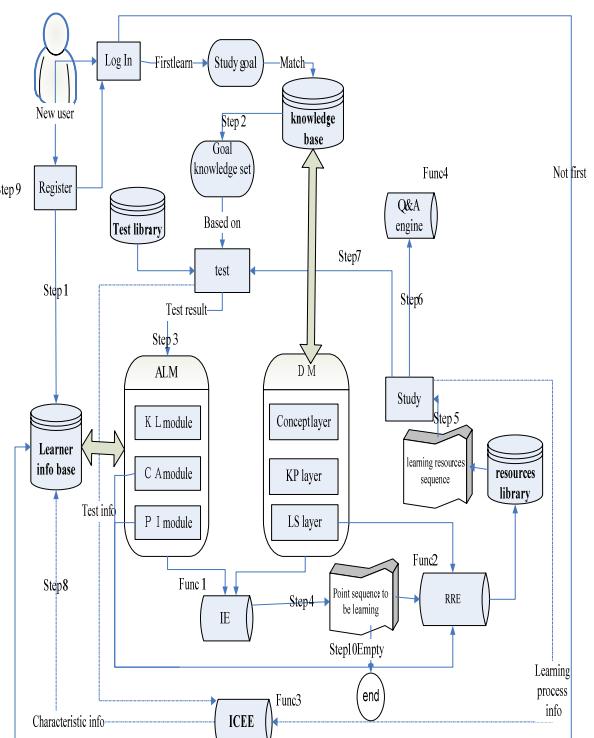


Figure 6. The representation of preferences value

the set, then designs tests according to “Target-knowledge”. ALM will deduce learner’s knowledge level and cognitive ability from test data. After that, IE achieves the “Knowledge-sequence” set that the learner will learn according to the deducing result of ALM, domain model and “Target-knowledge”. RRE recommends learning resources to meet learning needs of the learner best based on “Knowledge-sequence”. ICEE collects and arranges the information of learning process and testing process to reappear for learner’s characteristic information. System’s main flow is shown in Fig.6.

The descriptions of main flow as follows:

Step 1: The new user need to register. The learner’s personal information, academic information and other information are recorded in the learner info base.

Step 2: After user log in, need to choose learning goals if it is the first time the learner use the system. The learning objectives can be a chapter, a section or a knowledge point. Relations between them are included. That is, a chapter formed by a number of sections, and a section also formed by a number of knowledge points. Based on the learner’s goal, choose knowledge points as goal knowledge points set in the knowledge base under the goal, Express this set with target-knowledge.

Step 3: Design the test questions according to this set, and the test data need to feed back to ALM.

Step 4 : Educe the sequence of the knowledge points that to be learning for the learner from the relevant data between ALM and DM .

Step 5: The system shows the selected learning resources to users sequentially.

Step 6: Learners can get the answers from the Q&A engine when they are in trouble.

Step7: Learners can test at any time when they are learning.

Step8: Record learners' learning process information in learner information base

Step 9: If the learner has used the system, it directly shows the resources that the learner should learn according to his learning record.

Step10: If the sequence is empty, then end. Learners need to choose a further target.

B. The Major Works of NalsALM

1) The construction of Domain Model

The domain model may regard as is a professional field knowledge library, it is containing the fact knowledge and a Meta-knowledge. How is the system auto-adapted height decided to a great extent by organizes the knowledge, the expression knowledge and the utilization knowledge. DM is a graphical representation of domain knowledge, and is an abstract of each constituent of domain knowledge[10] . As shown in Fig.7 the relations of learning objects are from the course ontology. System realization is divided into three tiers, namely concept tier, knowledge-point (KP) tier and resources tier. The concept tier in Fig.7 is the abstract description of course ontology, and the knowledge-point in Fig.7 is an atomic unit, associated with learning objects, and cannot be separated. The resource tier presents the specific learning objects.

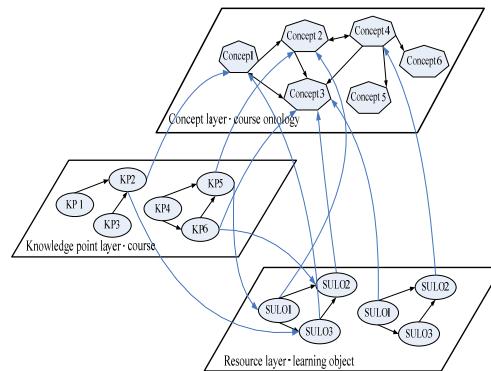


Figure. 7 DM based on ontology

The relationship between knowledge-points depends on the relationship of course concept, In other words, the relationship between learning objects depends on course ontology, so we need to determine the course ontology and its semantic relationship first.

2) Learning resources push mechanism

We can obtain “Knowledge-how”, “Ability-level”, and “Preference” by Estimation methods in section 4. Here, we will introduce how to achieve to push the learning resources effectively and personalized based on these three information.

a) Access to “Knowledge-sequence”

IE Arranges the disorder KP in learner’s knowledge level set “Knowledge-how” educed by ALM according to the relations of KP in DM. knowledge base has stored the requirements of mastering degree of KPs. It Judges KP one by one, and deletes the KP which meet the requirement from the set. Finally obtains a set of pre-KPs as “Pre-sequence” to be learning. So ,

$$\text{Pre-sequence} = \{(k_i, r_i), \dots, (k_j, r_j), \dots, (k_m, r_m)\}, 1 \leq i < j < \dots < m \leq n$$
where k_j represents the j^{th} KP, r_j represents the rank of the j^{th} KP, and $r_j = h_j + x$ where h_j is the learner’s mastering of the j^{th} KP in “Knowledge-how”, x is given by the system in accordance with the teaching objectives. Then obtains a set of target KPs sequence as “Target-sequence” according to “Target-knowledge”, so

$$\text{Target-sequence} = \{(t_1, 1), (t_2, 1), \dots, (t_l, 1), (t_n, 1)\}$$
, the target KPs need to learn from rank 1, step-by-step . So the set of KPs to be learning “Knowledge-sequence”= Pre-sequence \cup Target-sequence.

b) Recommendation of learning resources

RRE selects learning resources which meet the rank needs of each KP of “Knowledge-sequence” in resources base as a collection “R” of learning resources, Then match the “Ability-level” of the learner to the related CA attribute of each resource, Obtain a collection “R₁” of learning resources whose similarity are within a certain range(Ideally, it is necessary to make the recommended learning resources can not exceed any level of CA of the learners, but also help to improve certain CA of learners) Finally, according to the learner’s “preferences-set” selects the learning resources that best meet the learning needs of learners in R₁.

Appropriate for a certain knowledge point learning resources, learners need to be tested after learning that the resource.

If pass, add the mastery of knowledge with 1 pace. Next, determine whether the rank meet the Level requirements of the KP. If it meets, delete the KP. If not, improve the level of KP and continue to learn until meet the demand.

If not pass, re-select other resources in this level to learn for learner until he pass the test.

For example, student A has reached the rank 2(comprehension) to K_i in the current by test, but the knowledge base required level is 4 (analysis). So for KP K_i , the system should give the learning resource in rank 3(application) to A. If A pass the test, we can infer that A has reached the rank 3, but still not reach rank 4. Next, the system will provide the learning resource to A in rank 4(synthesis).If A don't pass the test, the system will still provide other learning resource to A in rank 3(application).

c) Technologies and applications

The system is implemented in MVC mode. It mainly uses JAVABEAN to enclose business logical layer. The presentation of interface adopts JSP and XML technology. The database is SQL server. Besides above-mentioned technologies, the system also consults JavaScript, DOM and others. The used software includes Jena, JBuilder, Dreamweaver, SQL server etc[14][15]. This work has been applied in junior high school for assisted learning. Through experiments it proves to be an efficient method for adaptive learning.

VI. CONCLUSION

The construction of learner model is a complex process in Adaptive Learning System. In this paper, based on Gagne's learning hierarchy theory , We proposed a new LM called Advanced Learner Model (ALM for short) consist of knowledge level, cognitive ability and preferences. We give the representation of ALM separately. Then give acquisition of characteristic value of ALM in detail. Such as, We use the idea of fuzzy logic to estimate the learner's knowledge level, because its maneuverability is quite strong. Use linear computation method to estimate the learner's cognitive capacity, it is simple and practical. Use vector space theory to represent preferences. After that, we propose a new Adaptive Learning System based on this learner model as a case. Experimental results show this learner model is effective and practical in the application.

In addition, this model has the following two points to study further: 1) In estimating learner's mastering of a certain knowledge point, the method need to be researched to determine the degree of difficulty. 2) The more appropriate data mining algorithm need to be selected to realize the renewal and maintenance of learner model.

ACKNOWLEDGMENT

We would like to thank professor Yongjian Yang for his insightful comments in reviewing drafts of this paper. We also would like to thank Hi-Tech Foundation of Jilin Province Development and Reform Commission (2010642) for the sponsor to our study.

REFERENCES

- [1] X.P.Wang, Q.S.Hu and B. Li. "Personalized Information Retrieval for the Internet Intelligent," Journal of Computer Research and Development, vol.36, September 2009, pp.1039-1046.
- [2] Z.Li and G.Q.Zhao. "Research on the Student Modeling of the ICAI System," Computer Engineering and Science, March 2002, pp.101-104.
- [3] C.P.Zhou and X.Q.Zhou. "The research of adaptive student model on Bayesian networks," Journal of Shandong Institute of Light Industry(Natural Science Edition), April 2007, pp.3-6.
- [4] B. Jia, S.C. Zhong,et al.The Construction and Evolution of Learner Model in Adaptive Learning System. In: Hj. Kamaruzaman Jusoff, eds . Proc. of The 2009 International Conference on Computer Technology and Development (ICCTD 2009),USA:IEEE,pp.148-152, November 2009.
- [5] X. P. Lai, Y. P. Wu and W. D. Cheng, "The Design and Implementation of Multimedia Intelligent CAI in Mining Science," Beijing: Computer Systems Applications, No. 2, pp.34-35, February 1997.
- [6] P. D. Chen and K. D. Li, "Design of student model centered web based adaptive learning system, " Proceedings of SITE 2002, Nashville, Tennessee, March 2002.
- [7] B. Jia , S.C. Zhong,et al. "The Study and Design of Adaptive Learning System Based on Fuzzy Set Theory", Transactions on Edutainment.Changchun, pp.1-11,August 2010.
- [8] R. M. Gagne, "Educational technology,as technique," Educational Technology, pp.5-13, August 1968.
- [9] H. Feng , L. Wang and H. Yang, "The Research of Multi-purposes Based Learner Model", Beijing: Computer Engineering and Applications, vol. 39, No. 36, pp.98-100, December 2003.
- [10] B. S. Bloom, M. D. Engelhart , E. J. Furst, W. H. Hill and D. R. Krathwohl, "Taxonomy of Educational Objectives," handbook I , Cognitive domain, London, Longman,1989.
- [11] S. Q. Tang and Z. Zhong, "Evaluation of Cognition of Students," Nanning: Journal of Academy of Sciences, Vol.18, No. 4, pp.281-284, April 2002.
- [12] Z.H.Jiang, Z.J.Wu. Network environment, multi-dimensional characteristics of the learner model to build. E-Education Research, April 2005, pp.71-73.
- [13] Z. Zhuang, "Estimation method of student's cognitive level in self-adaptive distance tutoring system," Beijing : Computer Engineering and Applications, vol. 43, No. 3, pp.220-223, January 2007.
- [14] B. Jia , S.C. Zhong,et al. "The Applied Research of Mobile Technology in Family Education about the Health of the Elderly". The 2nd International Conference on Advanced Computer Theory and Engineering (ICACTE 2009), Egypt,pp.729-735,September 2009.
- [15] Y. J. Zhong, J. Liu, S. C. Zhong, Y. M. Zhang and X. C. Cheng, "Programming of informatized instructional design platform for physics," Proceedings of Technologies for E-learning and Digital Entertainment, pp.171-177, 2006.



Bing Jia JinLin Province, China. Birthdate: Februry, 1985. studies for the doctorate in the Computer Science College of Jilin University. And research interests on artificial intelligence and active service.



Yongjian Yang JinLin Province, China. Birthdate: 1960. is vice president and PhD supervisor of the computer College of Jilin University. And research interests on artificial intelligence and active service.

Multi-satellite Monitoring SST Data Fusion based on the Adaptive Threshold Clustering Algorithm

Hongmei Shi[1], Han Dong[2], Lingyu Xu[3], Cuicui Song[4], Fei Zhong[5], Ruidan Su[6]
 [1, 3-5]Department of Computer Engineering & Science, Shanghai University, Shanghai, China, 200072

[2]The National Marine Information Centre, Tianjin, China, 300171

[6] College of Information Science and Engineering, Northeastern University, Shenyang, China 110004

Email:{ [1] shmei1988@163.com , [3] xly@shu.edu.cn }

Abstract—This paper proposes a method which describes the information precision with a soft fusion model, instead of the traditional rigid fusion method. The method is divided into two steps, the pretreatment model and fusion center model. Each forms a relative independent model, and the two models have a progressive relationship. The former is used for consistency evaluation, data cleaning and invalid data eliminating, while the latter provides fusion results and variable precision fusion expression by the adaptive threshold clustering algorithm. Experimental results show that the fusion method can not only give every SST data a different precision, but also carry more information to describe precision multiple distribution, which make users get high-quality data and enjoy more rights.

Index Terms—information fusion, sea surface temperature (SST), variable precision, threshold, clustering, adaptive threshold

I. INTRODUCTION

Multi-source remote sensing information fusion, one of the intense topics of international remote sensing over several years, is becoming a key technology to process multi-source massive data fusion. The development is transforming from research into function, but the previous papers on the remote sensing sea surface temperature (SST) data fusion were only limited in some localities. In order to overcome the bottleneck of single technology, and achieve the complementary advantages, foreign scholars have started to comprehensively study on marine monitoring regarding information fusion since 2000. There are typical composite fusion test platforms which are Marine environment monitoring and information integrated system (SEAWATCH) developed by Norwegian, Marine environment remote-control measurement and detection system (MEAMAIID) developed by Germany [1]. MISST [2] (Multi-sensor Improved Sea Surface Temperatures) can get improved, high resolution, global coverage, near real-time SST data by fusing the infrared and microwave sensor data. Besides, Bovith et al. [4] have detected SST elements through multi-remote sensing fusion image processing. Oesch et al. [5] used multi-scale remote sensing fusion to observe tepidity

situation in lake. Kozai et al. [6] used multi-satellite remote sensing data to fuse the sea surface wind.

In China, the state 2nd oceanic administration Ma [7] pointed out that it's difficult to practically operate the common assimilation method in the specific application due to their respective shortcomings, so he proposed to use fusion method to deal with the assimilation problems of the common Marine observation data. Guan [8] adopted the objective analysis to fuse TMI, AVHRR, TRMM/VIRS satellites. Shi [9] first proposed the concept of soft fusion quantitative precision in remote sensing and assimilation product fusion. Xu [10] first gave the fusion products and the precision analysis in East China seas by using multi-remote sensing fusion method.

However, the research results mentioned above are all used traditional rigid fusion method, which means the rigid processing method is used through the fusion process to filter out fuzzy and inconsistent information, and finally certain results are got, but it doesn't have a precise description to the accuracy and reliability of the fusion point, and user can't learn a conflict among the data sources just from those products, they have no idea of which points are fused in certain circumstances, and which are in hesitation. This paper proposes a more advanced fusion method with the adaptive threshold clustering algorithm, and it's divided into two steps, the pretreatment model and the fusion center model, each forms a relative independent model, and the two models have a progressive relationship. The former is used for consistency evaluation, data cleaning and eliminating invalid data, while the latter provides fusion results and variable precision fusion expression by the adaptive threshold clustering algorithm. The first part of this paper introduces the classic fixed threshold clustering algorithms; The second part proposes the adaptive threshold clustering algorithm to make up the insufficient of the fixed threshold clustering algorithm; The third part is the case analysis of the adaptive threshold clustering algorithm; The fourth part is fusion result comparison and analysis.

II. CLASSIC FIXED THRESHOLD CLUSTERING ALGORITHM

A. The Idea of Classic Algorithm

The basic process of classic fixed threshold clustering algorithm is determine the threshold which is fit for all data samples according to the priori knowledge, and then do clustering analysis to data samples under the threshold, and finally obtain the clustering results.

B. The Description of the Algorithm

Input: the $N(N \geq 2)$ unspecified data samples X_N , threshold T

Output: M class data results

Steps:

Step1 Sort the unspecified data samples in ascending order ($X_1 \leq X_2 \leq \dots \leq X_N$).

Step2 Select the first member X_1 as the first group data for the first class C_1 .

Step3 Calculate the distance d_{11} between X_1 and X_i in dataset $X_i (i = 2, \dots, N)$ orderly, which $d_{11} = |X_1 - X_i|$. Classify $X_1 \sim X_{k-1}$ as Class C_1 until a data point X_k that makes $d_{1k} \geq T$, and classify X_k to a new Class C_2 , and take X_k as the first element of Class C_2 as well.

Step4 Similarly, calculate the distance between X_k and the data after X_k . Get all the elements of the class and the first element of next class.

Step5 Repeat step3 and step4 operations, until finishing all the samples and get M kinds of data totally.

C. The Adaptability Analysis of the Algorithm to SST Observation

The advantage of the algorithm is its simple calculation, and if there is a priori knowledge of the sample distribution for the threshold value selection, we can get reasonable clustering results quickly. However, this algorithm also has a series of problems. In practice, it is difficult to obtain accurate priori knowledge for high dimensional model samples. Therefore, we can only choose different threshold to tempt, and thus the clustering results largely depend on the choice of threshold T. Different thresholds lead to different clustering results, large threshold will get smaller cluster number, and small threshold may get more cluster number relatively, but there is no criterion to determine whether the large threshold effect, or small threshold effect.

III. THE ADAPTIVE THRESHOLD CLUSTERING ALGORITHM

A. The Thought of the Algorithm

When the user lacks priori knowledge of clustering data, it is difficult to determine the appropriate threshold, and often needs to execute many times of experiment and compare the test results to find the most suitable threshold. For the fixed threshold clustering, once the user changes the threshold value, it often needs to restart the clustering computing. Each-time clustering would cost a lot of time, and sometimes the clustering efficiency is very poor. Therefore, on the basis of the existing fixed threshold, the adaptive threshold algorithm is proposed.

B. The Description of the Algorithm

Input: SST data of thirteen satellites

Output: inside-class data and outside-class data after clustering, variable precision fusion expression, and fusion reliability.

Steps:

1. Extract thirteen satellites' temperatures of a certain day in the same latitude and longitude.

2. Input the minimum threshold and maximum threshold.

3. Calculate the distance D between any two of thirteen temperature data (totally $13 * 13$ distance), discard those temperature data which the calculated distance are greater than the minimum threshold, and find out the largest frequency temperature data involved in the calculation among the remaining distances, remember the temperature and those temperatures which the distance D less than the threshold as inside-class temperature, the rest of the temperatures as outside-class temperature.

4. If the number of inside-class temperature less than the outside-class temperature, then increase step length (increase one-tenth of the difference of the maximum threshold and minimum threshold) and return to step 3. If it is still not satisfied conditions until the threshold value greater than the maximum threshold, then exit algorithm and the fusion is failed. Otherwise, if the number of inside-class temperature greater than the outside-class temperature, then jump to Step5.

5. The inside-class temperature data recorded as (1), outside-class temperature data noted for (0), then according to the inside-class temperature data, calculate the fusion expression, the weighted center temperature, the center temperature, and the fusion reliability.

6. Computing formula of fusion expression: fusion expression = fusion temperature \pm error.

Set: the inside-class maximum temperature recorded as C_{Max} , inside-class minimum temperature recorded as C_{Min} , fusion temperature recorded as C_{fusion} , the error record as Error, fusion expression denoted by E, then:

$$\text{Fusion temperature } C_{fusion} = \frac{C_{Max} + C_{Min}}{2} \quad (1)$$

$$\text{Error} = \frac{C_{Max} - C_{Min}}{2} \quad (2)$$

$$\begin{aligned} \text{Fusion expression } E &= C_{fusion} \pm \text{Error} \\ &= \frac{C_{Max} + C_{Min}}{2} \pm \frac{C_{Max} - C_{Min}}{2} \end{aligned} \quad (3)$$

7. The center temperature calculation: the center temperature is the average value of all the inside-class temperature. Then, $W = \frac{W_1 + W_2 + \dots + W_n}{n} \quad (4)$

8. Calculate reliability $C_{reliability}$: the number of inside-class temperature number recorded as C_{num} , effective temperature denoted by C_{num}' ,

$$\text{Then, } C_{reliability} = \frac{C_{num}'}{C_{num}} \quad (5)$$

9. Weighted center calculation: set central temperature as C' , the weight P, the distance between all inside-class temperature and center temperature is $D = |C_1 - C'| |C_2 - C'| \dots |C_n - C'| \quad (6)$

$P_i = 1 - (|C_i - C'| / D)$ as to ensure the smaller distance from the center temperature C' , the greater the weight. Calculate weighted center:

$$W = \frac{P_1 * W_1 + P_2 * W_2 + \dots + P_n * W_n}{W_1 + W_2 + \dots + W_n} \quad (7)$$

C. Advantages

Compared to fixed threshold algorithm, the adaptive threshold clustering algorithm has many advantages. First, this algorithm can get satisfied clustering results, because the results of the fixed threshold clustering largely depend on the choice of threshold T , different threshold leads to different clustering result and clustering effect is not ideal. Sometimes there's only a few or no data within the class, making the clustering without any meaning. While the adaptive threshold clustering algorithm guaranteed the clustering results reasonable. Second, the clustering speed is fast, because we define a threshold interval, it only needs to traverse the data space one time to achieve clustering. Third, using the proposed adaptive threshold clustering algorithm, user needs less prior knowledge but two input threshold parameters. For the set size of threshold, the user can test several times to determine it, and finally obtain the satisfactory clustering results. When the user changes threshold to re-clustering, this algorithm can get all kinds of different particle size of clustering.

IV. EXPERIMENT AND ANALYSIS

A. Algorithm Example

According to the proposed adaptive threshold clustering algorithm, fuse the SST data on the 9th day of 2006. SST data products including AVHRR, MODIS, TMI, MCSST of 13 types.

Example procedure:

1. Read all the data of thirteen satellites on the 9th day of 2006. Take the temperature of the point (38.875 °C, 128.125°C) as example, the SST data is expressed in TABLE I.

TABLE I.
THE SST DATA IN (38.875 °C, 128.125°C)

Satellite name	SST data
Avhrr0.25	-
Amsre-A0.25	-
Amsre-D0.25	-
Clim0.25	5.712
Davhrr0.25	4.275
Modisast40.25	3.979
Modisastd0.25	3.978
Modisastn0.25	3.936
Modistsst40.25	4.269
Modistsstd0.25	-
Modistsstn0.25	4.292
Tmi-A0.25	-
Tmi-D0.25	-

2. Let the minimum threshold value be 0.15, and the maximum threshold value be 1.0.

3. According to Step3 of the algorithm, calculate the distance between any two of the seven valid temperature data, and write down the distance following in the corresponding satellites in order, which is shown in TABLE II.

TABLE II.
ALL THE DISTANCES OF THE SEVEN VALID SST DATA

Clim0.25	Davhrr0.25	Modisast40.25	Modisastd0.25	Modisastn0.25	Modistsst40.25	Modistsstd0.25	Modistsstn0.25
5.712	4.275	3.979	3.978	3.936	4.269	4.292	
0	1.437	1.733	1.734	1.776	1.443	1.42	
1.437	0	0.296	0.297	0/339	0.006	0.017	
1.733	0.296	0	0.001	0.043	0.29	0.313	
1.734	0.297	0.001	0	0.042	0.291	0.314	
1.776	0.339	0.043	0.042	0	0.333	0.356	
1.443	0.006	0.29	0.291	0.333	0	0.023	
1.42	0.017	0.313	0.314	0.356	0.023	0	

When threshold $T = 0.15$, according to Step3 of the algorithm, there are only three inside-class temperature while the outside has four, the result can be seen in TABLE III.

TABLE III.
THE NUMBER OF INSIDE-CLASS SST DATA WHEN THRESHOLD IS 0.15

Clim0.25	Davhrr0.25	Modisast40.25	Modisastd0.25	Modisastn0.25	Modistsst40.25	Modistsstd0.25	Modistsstn0.25	Inside-class data number when T=0.15
5.712	4.275	3.979	3.978	3.936	4.269	4.292		
5	9	8	6	9	2	4.29		
0	0				0.00	0.01	1	
		0	0.00	0.04	6	7	3	
		0.00	1	3			3	
		1	0	0.04			3	
		0.00	0.04	0.04	2		3	
		6	3	2	0	0.02	3	
		0.01			0.02	3	3	
		7			3	0		

According to step4 of the algorithm, a fixed step size (0.085) increased the threshold, then $T = 0.235$, empathy in step 3, obtained three data inside the class, and four outside the class; Then add another step, so $T = 0.32$, get six data within the class, and one outside the class, thus, the number of inside-class data greater than the outside-class, which fit the termination conditions. At this time, (1) is used to represent the six inside-class data, and the rest one is expressed by (0), as shown below in TABLE IV.

TABLE IV.
THE NUMBER OF INSIDE-CLASS SST DATA WHEN
THRESHOLD IS 0.32

Cli m0.	Dav hrr0.	Mod isass t40.2	Mod isass td0.2	Mod isass tn0.2	Mod istsst 40.2	Mod istsst n0.2	Inside- class data number when T=0.32
25	25	5	5	5	5	5	
5.71	4.27	3.97	3.97	3.93	4.26	4.29	
2	5	9	8	6	(1)	2	
(0)	(1)	(1)	(1)	(1)	(1)	(1)	
0	0.29	0.29			0.00	0.01	
0	0.29	6	7		6	7	1
0.29	0.00	0	0.04	0.04	0.29	0.31	5
0.29	0.00	1	3	3	0.29	3	6
7	1	0	0.04	0.04	1	0.31	6
0.04	0.04	2	2	2	4		3
0.00	3	0.29	0	0	0.02	0.02	5
6	0.29	0.29					
0.01	0.31	1			0.02	3	
7	3	0.31		3	0		
		4					

4. Calculate fusion expression, fusion expression = fusion temperature \pm error.

By the formula (1),

$$C_{fusion} = \frac{C_{max} + C_{min}}{2} = \frac{4.292 + 3.886}{2} = 4.114,$$

By the formula (2),

$$\text{Error} = \frac{C_{max} - C_{min}}{2} = \frac{4.292 - 3.886}{2} = 0.178,$$

Therefore, from formula (3),

$$\begin{aligned} \text{Fusion expression} &= E = C_{fusion} \\ \pm \text{Error} &= \frac{C_{max} + C_{min}}{2} \pm \frac{C_{max} - C_{min}}{2} = 4.114 \pm 0.178. \\ \text{Reliability calculation by formula (5)}, \\ C_{reliability} &= \frac{C_{num}}{C_{total}} = \frac{6}{7} = 85.7\%. \end{aligned}$$

B. The Experimental Analysis of Adaptive Threshold Clustering Algorithm

The proposed adaptive threshold clustering algorithm can apply to the project of multi-satellites monitoring SST data fusion. This project mainly studies the soft fusion method, which is divided into pretreatment and fusion center models, each forms a relative independent model, and two models have a progressive relationship. The former is used for consistency evaluation and data cleaning, while the latter provides fusion results and variable precision fusion expression by the adaptive threshold value clustering algorithm.

The pretreatment model is mainly used for the basic data collection and preliminary quality control. Process the multi-format, multi-precision, multi-source remote sensing data in east China sea with AVHRR, MODIS, TMI, MISST products and etc. Considering the data loss of the multi-resolution, single monitor products and other factors, we preliminary sort the SST data and eliminate the noise.

According to above adaptive threshold clustering algorithm, read the SST data on the first day of 2006 as experimental data. Satellite temperature data including 13 satellites, which are AVHRR, MODIS, AMSR-E, TMI and etc. TABLE V. shows the pretreatment results.

TABLE V.
THE PRETREATMENT RESULTS ON THE 1st DAY OF 2006

Latitude	Longitude	SST data												
20.625	118.875	-	24.9	25.2	24.74	-	-	24.511	-	25.251	24.405	24.886	-	24.9
20.625	119.125	-	24.45	25.05	24.712	-	-	24.453	-	24.462	24.173	24.108	-	24.75
20.625	119.375	-	24.15	24.9	24.962	-	-	24.185	-	24.554	24.076	24.185	-	24.9
20.625	119.625	-	24.45	25.5	25.017	-	-	24.126	-	24.335	-	23.86	-	25.5
20.625	119.875	-	24.9	26.55	25.065	-	-	24.302	-	24.467	-	23.956	-	25.95
20.625	120.125	-	25.5	26.85	25.267	-	-	24.345	-	-	-	-	-	26.4

The pretreatment module mainly picks up the satellite data, eliminates the invalid data files of each satellite, finally generates the pretreatment files to make the following fusion operation more conveniently, and improves fusion efficiency as well.

In order to overcome the shortcomings of the fusion results ever express less information, and the poor transparency, this paper establishes a soft fusion model, explored with multiple precision/reliability distribution of continuous valued fusion algorithm to solve the multiple source conflict problem of three-dimensional monitoring. It can improve the quality as well as quantify precision further, and provide the user more right to know.

In order to overcome the poor effect by using the fixed threshold cluster method, we introduce the adaptive

threshold clustering method. First, according to the user's needs, input two thresholds, the minimum threshold and maximum threshold, to form a threshold interval. And make the threshold interval divided equally in a step length, compare the distance between every two satellite data circularly, it stopped until the number of the inside-class data more than half of effective data, and thus the clustering finished.

Since the threshold of each observation point is different, the error is also different, which eventually leads to different precision, resulting for each observation point a variable precision fusion expression, which can be expressed as the measured value \pm measurement errors.

Fusion thirteen satellite data on January 1, 2006, fusion results stored in fusion result file (TABLE VI.)

TABLE VI.
FUSION RESULTS ON JANUARY 1,2006

Latitude	Longitude	Final threshold	Cluster result						Variable precision fusion result	reliability
29.375	128.125	0.15	22.35 (1)	22.233 (1)	22.209 (1)	22.333 (1)	22.5 (1)		22.355±0.146	100%
26.875	123.375	0.15	22.05 (1)	21.9 (1)	20.691 (0)	21.9 (1)	21.75 (1)		21.9±0.15	80%
24.375	126.125	0.405	23.4 (1)	24.358 (0)	23.744 (1)	23.772 (1)	23.395 (1)	23.775(1)	23.585±0.190	83%
21.375	119.375	0.405	25.2 (1)	25.2 (1)	24.855 (1)	24.857 (1)	24.826 (1)	24.45(1)	24.825±0.375	100%
23.875	122.875	0.825	23.85 (1)	23.85 (1)	24.635 (0)	22.955 (1)	23.028 (1)	22.8 (1)	23.325±0.525	85%

In order to more directly see the advantages of the adaptive threshold clustering algorithm than classic fixed threshold algorithm, Figure 1 and Figure 2 show all the SST data of the two algorithms in the first day of 2006 respectively. Fig. 1 set the threshold interval [0.15,1.0], and from which get that the threshold is variable, different locations have different threshold. Finally, this algorithm obtained more accurate and reliable fusion results which based on point-to-point precision. While Fig. 2 only has a threshold value (0.15), and the result is equal precision that we can't determine which point is relatively reliable from the figure, and the transparency is very poor.

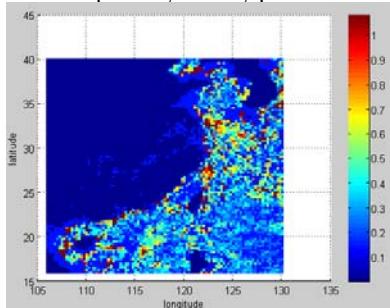


Figure 1. Threshold figure of adaptive clustering algorithm

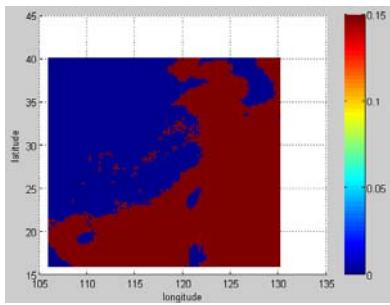


Figure 2. Threshold figure of classic fixed clustering algorithm

We use formula (7) calculate the weighted center value of all inside-class data in different latitude and longitude, to

describe the variable precision fusion result, and the result shows in Fig. 3.

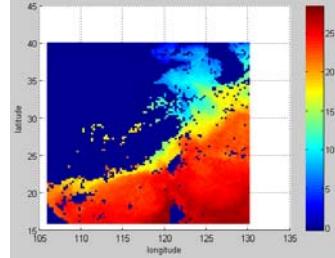


Figure 3. Variable precision fusion result of the weighted center

Similarly, Fig. 4 and Fig. 5 show the number of the inside-class data of the two algorithms, we can see that in some sea area, the number of inside-class data in Fig. 4 were significantly more than that in Fig. 5, which means that the reliability of the adaptive clustering algorithm has significantly increased than fixed threshold algorithm. The experiment shows that the algorithm proposed in this paper is more reliable.

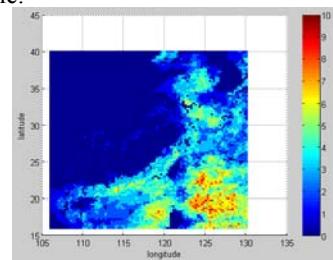


Figure 4. The number of the inside-class data using adaptive threshold clustering algorithm

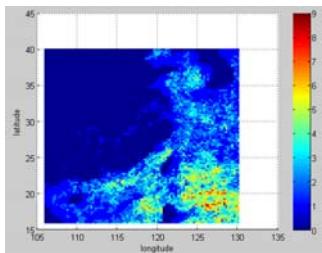


Figure 5. The number of the inside-class data using fixed threshold clustering algorithm

V. CONCLUSION

Adaptive threshold clustering algorithm are proposed on the basis of the existing fixed threshold clustering algorithm, the results of the fixed threshold clustering largely depend on the choice of threshold T , different threshold leads to different clustering result and clustering effect is not ideal. However, adaptive threshold clustering algorithm doesn't have this kind of problem, the threshold of different longitude and latitude is different, and then the final fusion expression is also different, thus realize the soft fusion which is put forward in this paper very well.

This paper proposes a soft fusion model which is divided into two models, pretreatment model and fusion center model, each forms a relative independent model, and two models have a progressive relationship. The former is used for consistency evaluation, data cleaning and invalid data eliminating, while the latter provides fusion results and variable precision fusion expression through the adaptive threshold clustering algorithm.

ACKNOWLEDGMENT

This work is supported by The National Natural Science Foundation of China. (No.40976108)

This work is supported by Shanghai Leading Academic Discipline Project in China. (N0. J50103)

The Ocean Public Welfare Project of The Ministry of Science and Technology under Grant No.201105033.

This work is supported by Innovation Action Plan supported by Science and Technology Commission of Shanghai Municipality (No.11511500200)

REFERENCES

- [1] Shi Suixiang.Study on Soft Fusion Strategy of Multi-Channel Incertitude Information in Digital ocean[D].Northeastern University,2005:7-10.
- [2] http://www.misst.org/
- [3] Zhang hui. Merging AVHRR and AMSR-E Sea Surface Temperature Data Based on Wavelet Transform[D].qingdao: Ocean University of China,2006:4-9.
- [4] Bovith, Thomas, Nielsen, Allan Aasbjerg, Hansen, Lars Kai, Overgaard, Soren, Gill, Rashpal S.. Detecting weather radar clutter by information fusion with satellite images and numerical weather prediction model output. In proc: 2006 IEEE International Geoscience and Remote Sensing Symposium, 2006, 6: 511-514.
- [5] Oesch, D., Jaquet, J. M., Klaus, R. Schenker.. Multi-scale thermal pattern monitoring of a large lake (Lake Geneva) using a multi-sensor approach. International Journal of Remote Sensing, 2008, Vol.29, No.31, pp: 5785-5808.
- [6] Kozai K., Ishida K., Shiozaki T., Okada Y.. Wind-induced upwelling in the western equatorial Pacific Ocean observed by multi-satellite sensors, Advances in Space Research, 2004, Vol.33, No.7, pp: 1189-1194.
- [7] Ma Jianpu,Huang Daji,Zhang Benzhao.Blending method and its application in data assimilation[J]. Acta Oceanologica Sinica, 2003, Vol.25, No.2, pp: 33-41.
- [8] Guan lei, Chen Rui, He mingxia. Validation of Sea Surface Temperature from ERS-1/ATSR in the Tropical and Northwest Pacific[J]. Journal of Remote Sensing,01,63-69 (2007).
- [9] Shi Suixiang,Xia Dengwen,Yu ge.Method of fusing the data reversed from satellite remote-sensing with the assimilation data for sea surface temperature based on D-S theory[J]. Acta Oceanologica Sinica,2005, Vol.27, No.4, pp: 31-37.
- [10] Zheng Jinwu,Xu Dongfeng,Xu Mingquan.A review of merging methods of all covered high resolution SST. Journal oF Tropical Oceanography , 2008, Vol.27 No.4, pp: 77-81.
- [11] Sun Ji-Gui, Liu Jie, Zhao Lian-Yu, “Clustering Algorithms Research”, Journal of Software, Vol.19, No.1, January 2008, pp.48–61
- [12] Lingyu Xu, Jun Xu, “Satellite Remote Sensing SST Quality Evaluation Based on Consensus Measurement”, 2009 International Conference on Web Information Systems and Mining, pp807~810.
- [13] Zhong fei,Liu na,Liu yang,Xu Lingyu.New SST correction method from multi-satellite based on the coefficient of variation. Journal of Shanghai University (EnglishEdition),2011,15(5):463-466.

Description to Fe-C Alloy Film Fiber Corrosion Sensors by Fractal Corrosion Images

Lihua Cai

School of Mechanical and Electrical Engineering, China University of Mining and Technology, Xuzhou, Jiangsu, China

Email: cailihua1984@163.com

Wei Li

School of Mechanical and Electrical Engineering, China University of Mining and Technology, Xuzhou, Jiangsu, China

Email: liweicumt@163.com

Abstract—In this paper, the fractal dimensions are proposed for characterizing the optical fiber corrosion sensors, which is applied to measure the steel corrosion based on Fe-C alloy film. Because corrosion was a complex random phenomenon and the corrosion surface of Fe-C alloy film of the optical fiber corrosion sensor possessed fractal characteristics, the image fractal dimensions is as a quantitative index of the sensor corrosion degree. The experimental results showed that the complexity level of the sensor surface morphology increased and the optical output power increased along with the increase of fractal dimensions. And the sensing law of the thicker Fe-C alloy film is better than that of the thinner film. Therefore, it is feasible that the fractal dimension is used to characterize the Fe-C alloy film optical fiber corrosion sensors and it is also optimistic to the future prospect of the Fe-C alloy film optical fiber corrosion sensors.

Index Terms—Fe-C alloy film; fiber corrosion sensors; light output power; fractal dimensions

I. INTRODUCTION

For the past few years, with the development of the optic fiber sensing technology, the optic fiber sensing technology is also researched continuously and deeply to monitor metal corrosion condition. And as one kind of which, the optic fiber corrosion sensor (OFCS) based on the Fe-C alloy film plated in the fiber core is used to monitor the corrosion condition of reinforcing steel bars and underground pipeline[1-3]. Many researchers have carried out large numbers of research to this kind of sensor and obtained some primary sensing performance[4]. However, the corrosion test of the plated film is not thorough and the test design is simplex.

Like most of natural phenomenon, metal corrosion in

nature is complex and erratic, so the corrosion morphology and images could not be definitely identical even in the case of same material and environment. Meanwhile, corrosion images are irregular and unrepeatable. Therefore, fractal is one of the efficient ways to describe the corrosion behavior of the Fe-C alloy film. Now, many workers attempted to use fractal in corrosion researches. J.M. Costa presented a preliminary account on fractal properties of steel corrosion pitting[5]. And the fractal analysis of electrochemical noise helps to evaluate the inhibitor protection performance under the mild steel corrosion conditions tested[6]. Shaniavski and Artamonov calculated fractal dimensions for fatigue fracture surfaces[7]. Weng and Li use fractal dimension to modify the random fluctuation of average corrosion rate of car-bon steel in soils and obtain more accurate expression about soil corrosion[8]. Therefore the fractal dimension could be an important parameter to characterize the corrosion extent[9]. The author presented that the corrosion surface of Fe-C alloy film OFCS satisfied statistical fractal characteristics through plenty of testing data. In this article, the fractal dimension of corrosion morphology images of Fe-C alloy film OFCS is as a quantitative index of the sensor corrosion degree, which puts forward a new method for characterizing the OFCS.

II. THE PRINCIPLES OF FE-C ALLOY FILM OFCS.

The conductive optical signal through the fiber core generates an influence when Fe-C alloy film over the fiber core is corroded. Hence, we can obtain the corrosion information of steel materials by measuring real-time changes of output optical signal. The sensing principle is as follows:

According to the optical waveguide theory, when the angle of incidence θ meets the full-refraction condition with the incident light reflected by the interface of two mediums, the light beam occurs full-refraction phenomenon at the interface. The principle can be expressed as:

This work is supported by the National Nature Science Foundation of Jiangsu, China (No.BK2006037) and the Technology Support Project of Jiangsu, China (No. SBE20100378). Corresponding author is Lihua Cai (cailihua1984@163.com)

$$\theta > \theta_c = \sin^{-1} \frac{n_2}{n_1} \quad (1)$$

Where θ_c is the critical angle when the light waves are reflected by the interface of two mediums; n_1 and n_2 is the refractive index of two mediums, respectively. The light wave guide theory is applied to the OFCS. Fe-C alloy film which is similar with the composition of steel materials takes the place of the cladding of a certain part of the optical fiber, where is as a sensing field and became the metal waveguide. Because the absorption of Fe-C alloy film makes the optical energy loss larger inside the optical fiber in this field, which makes the emergent light energy lower. If the Fe-C alloy film is corroded, the sensing field is surrounded by corrosion medium and becomes a new waveguide environment, and the light beam satisfies the full-refraction condition in this sensing field, which makes the optical energy increase gradually. Therefore we can obtain the corrosion information of steel materials by measuring real-time changes of output optical signal. The optical fiber corrosion sensor is shown in Fig.1.

III. EXPERIMENTAL SET-UP

A. Fe-C Alloy Film Preparation

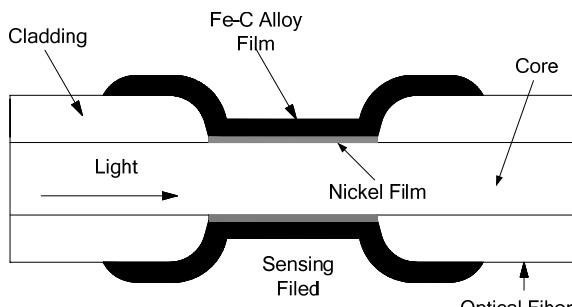


Fig.1 Optical fiber corrosion sensor with the corrosion-sensitive metal film

To plate Fe-C alloy film on the optical fiber surfaces, chemical plating nickel should be proceeded as a metal conducting layer. Before chemical plating nickel, surface treatment, such as removing fiber cladding, sensitization and activation, is prepared to obtain uniform and continuous nickel film.

(1) Surface treatment of optical fiber

Firstly, 15% hydrofluoric acid is selected to corrode the fiber cladding because of its strong volatile, and through repeated experiments, it is evidently linear relation between fiber diameter and corrosion time(the corrosion rate is $0.133 \mu\text{m}/\text{min}$). In order to completely corrode the fiber cladding off, the corrosion time is about 480 min and the fiber diameter is $61 \mu\text{m}$ (the original fiber core is $62.5 \mu\text{m}$), the process of which carries out fiber coarsening as well.

Then, the worked fiber is immersed into sensitizing solution to gain a layer of reduction substance on the fiber surface and offer catalytic activity metal ion for next activation. The recipe of sensitizing solution includes stannous chloride($\text{SnCl}_2 \cdot 2\text{H}_2\text{O}$ 10g/L) and hydrochloric acid(HCl 50ml/L).

Afterwards, optical fiber activation is that the Sn^{2+} on the fiber surface reacts with Pd^{2+} in the activation solution to form precious metal crystal nucleus with catalytic activity for chemical nickel-plating. The recipe of activation solution includes palladium chloride(PbCl_2 0.6g/L) and hydrochloric acid(HCl 5ml/L).

(2) Chemical nickel-plating

In this experiment, sodium hypophosphite is as reducing agent and by repeated experiments, the optimum craft condition of chemical nickel-plating is shown in Table 1.

By observing the fiber cross section with scanning electron microscope(SEM), the nickel film is widely distributed out of the fiber around and the average

TABLE I.
THE CRAFT CONDITION OF CHEMICAL NICKEL-PLATING

$\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$	32g/L
$\text{C}_2\text{H}_5\text{COOH}$	25ml/L
$\text{NaH}_2\text{PO}_2 \cdot 6\text{H}_2\text{O}$	26g/L
H_3BO_3	25g/L
pH	3.5~3.8
temperature	77°C~79°C
time	20min

thickness is approximately $0.2 \mu\text{m}$.

(3) Electroplating Fe-C alloy film

Electroplating experiment adopts constant current electroplating and 20 steel is used as the anode. In order to obtain uniform clad layer, the anode is coiled into cylindrical shape, where is put into metallized fiber core, and the axile wire of the fiber keeps coincidence with that of the anode. A good deal of experimental gropes is used to confirm the imposed current density of the cathode and the time of electroplating. the optimum craft condition of electroplating Fe-C alloy film is shown in Table 2. The plating solution temperature is controlled by high-accuracy constant temperature magnetic stirrers. Moreover, a small quantity of citric acid and ascorbic acid are mixed into the FeCl_2 solution, and its equation[10] is as follows:



TABLE II.
THE CRAFT CONDITION OF ELECTROPLATING FE-C ALLOY
FILM

FeCl ₂	40g/L
citric acid	1.2g/L
ascorbic acid	3g/L
pH	2.5~3.0
temperature	35°C~40°C

Through a quantity of electroplating experiments, current density plays a role in the quality of the coating. Therefore, we selects $i = 0.5\text{A}/\text{dm}^2$ first and electroplates 20min , then, we turns the current up to $i = 1\text{A}/\text{dm}^2$ and goes on eletroplating and the electroplating time decides the thickness of Fe-C alloy film. When viewed by SEM and shown in Fig.2, we

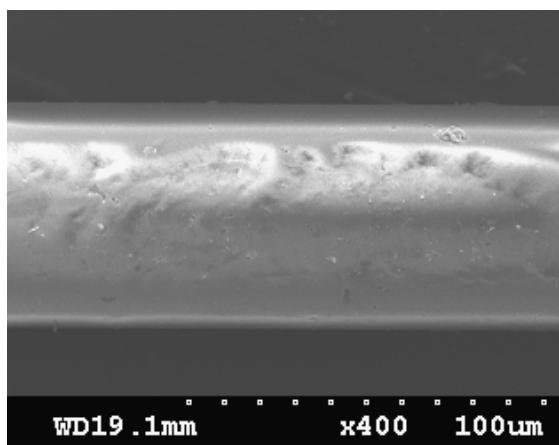


Fig.2 Surface morphology of fiber core electroplate with Fe-C alloy film

basically confirmed with the request.

B. Corrosion Sensing Experiments

After the optical fiber was plated with Fe-C alloy film, the sensing performance must be found out through corrosion test, which can be seen in Fig.3. The principle of the corrosion tests is very simple i.e., the optical fiber plated Fe-C alloy film is put into some kind of the corrosive medium and the change of the optical power transmitted in the fiber core is observed with the film layer gradually corroded. If we can find outthe accurate corresponding relation between the corrosion degree of the Fe-C alloy film and changes of output optical power, the corrosion degree can be judged with the change of output optical power.

In actual environment, different places have different corrosion condition with different corrosion rate. Therefore, because of the different corrosion rate, it is meaningless to record changes of output optical power with time as abscissas. And if we record changes of output optical power with percentage of corroded film as abscissas, although that would be not affected by corrosion rate, it is really tough to define the percentage of corroded film. Therefore, fractal dimension of metal-film corrosion morphology image is selected to express the corrosion degree. Corrosion is a complex random phenomenon, and at present, fractal dimension of the corrosion morphology image is used to express corrosion degree by many domestic and foreign researchers and the relationship is obtained among fractal dimension, corrosion morphology image and corrosion rate. The relationship is that the more complex the corrosion surface, the more irregular and the bigger the fractal dimension. And the faster the average corrosion rate, the deeper the corrosion surface hollow and the bigger the standard deviation, the bigger the fractal dimension. Therefore, the fractal dimension is an important parameter to characterize the corrosion degree. Author has analyzed that the corrosion surface of Fe-C alloy film OFCS satisfies with statistical fractal features by a large number of experimental data. In this paper, the

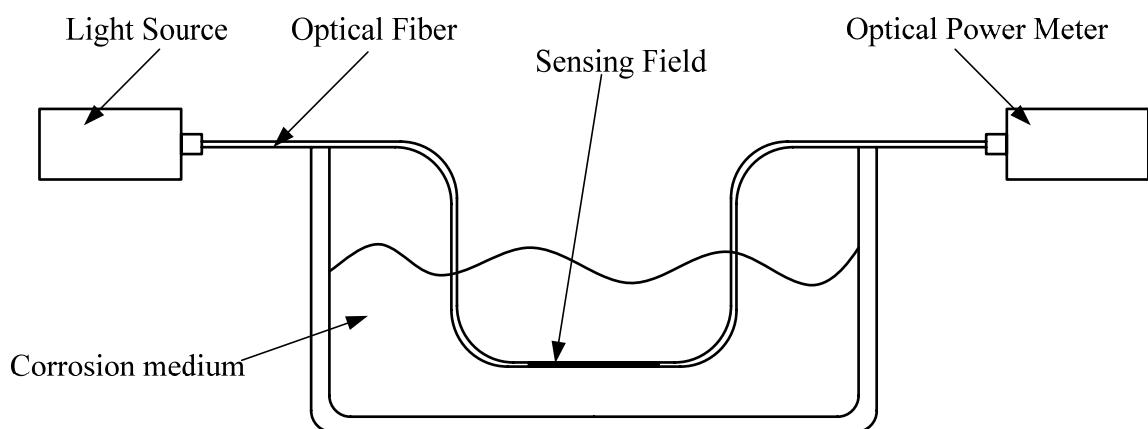


Fig.3 Principle of the corrosion test

can see that the quality of coating is widely distributed and the surface of that is relatively flat, which is

fractal dimension of corrosion morphology image of

Fe-C alloy film OFCS is a quantitative measure, which puts forward a new method to characterize OFCS.

Given a long corrosion process in reality, some corrosive solution is generally used to accelerate corrosion process of the coated optical fiber in the lab. Compared with using different concentration corrosive solution, it is more significant that sensing law is found out with the same concentration corrosive solution and the different thickness metal films. Based on the above analysis, the corrosion experiment test is as follows:

(1) Firstly, nickel-film is plated in the core as the middle conductive layer. Then, Fe-C alloy film is plated outside the nickel layer with different thickness, which is put in the same concentration corrosive solution respectively.

(2) Fe-C alloy film of each thickness needs to proceed corrosion test more than once to obtain as many test data as possible and avoids some accidental data, consequently, the sensing law is found out.

IV. EXPERIMENTAL ANALYSIS AND RESULTS

A. Fe-C Alloy Film Structure Microscopic Analysis

X-ray diffraction technique is used to assay the X-ray diffraction of Fe-C Alloy Film, comparing with that of 20 steel. The X-ray diffraction patterns are shown in Fig.4 and Fig.5.

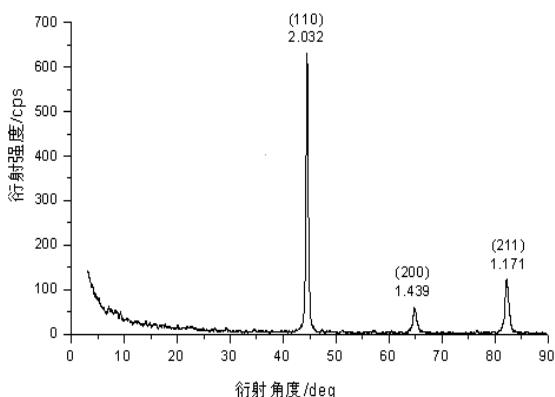


Fig.4 X-ray diffraction pattern of 20 steel

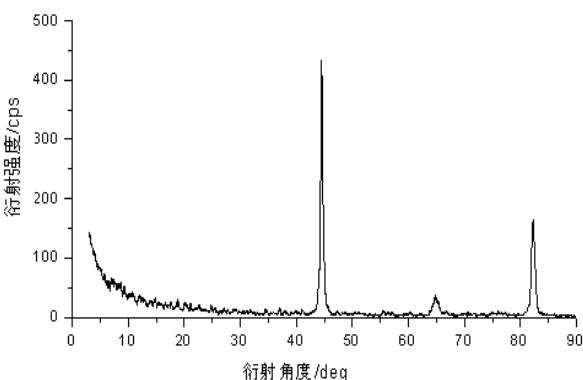


Fig.5 X-ray diffraction pattern of Fe-C alloy film

The experimental apparatus is D/Max-3B X-ray

diffractometer from Japanese Rigaku company. And the experimental parameters are as follows: copper target, $K\alpha$ radiation, graphite crystal bending monochromator, X-ray tube voltage is 35 KV, X-ray tube current is 30 mA, continuous scanning, scanning speed is $3^\circ/\text{min}$, and the sampling interval is 0.02° .

From the X-ray diffraction patterns of 20 steel, we can perceive that the interplanar spacing is 2.032, 1.439 and 1.171 respectively, and the interplanar indices is 110, 200 and 211 respectively, which indicates that the crystal structure is α -Fe. And from Fig.5, it can be perceived that electroplating Fe-C Alloy Film forms amorphous microstructure and the diffraction peak position is consistent with that of 20 steel. However, the diffraction strength of Fe-C Alloy Film is weaker at the diffraction peak of 110 and 200 than that of 20 steel, and is stronger at the diffraction peak of 211 than that of 20 steel. That suggested that primary crystal structure type of Fe-C Alloy Film is basically consistent with that of 20 steel and is just different from the crystal growth direction, which is relative to some technology conditions of electroplating.

B. Corrosion Sensing Analysis

Fe-C alloy films with different thickness are electroplated. If the film thickness is very thin, the corrosion time is too short to satisfy corrosion condition. However, if the film thickness is very thick, the fiber core is too thin to bear the weight of Fe-C alloy films. Hence, to obtain the preliminary sensing law of different thickness, optical fibers Plated Fe-C alloy film, with different thickness ($3.0 \mu\text{m}$, $7.6 \mu\text{m}$ and $12.8 \mu\text{m}$, respectively), is put into 10% nitric acid solution to conduct corrosion test, which is connected between light source and optical power meter. In the process of the corrosion, variation of the light output power and corrosion image is observed. When the light output power stops to change and stabilize for a period of time, we can terminate the experiment. The corrosion image and corresponding fractal dimension of Fe-C alloy film are shown in Fig.6. We can see that the corrosion images are more complex and the Fe-C alloy film corrodes more severely, the fractal dimension is bigger. Afterwards, the light output power was recorded with the variation of the fractal dimension.

Fig.7 shows that the relation is described between the three Fe-C alloy film with different thickness and the light output power. The results indicate that the curve of the $3.0 \mu\text{m}$ Fe-C alloy film is not better than that of the $7.6 \mu\text{m}$ and $12.8 \mu\text{m}$ Fe-C alloy film. And the $3.0 \mu\text{m}$ Fe-C alloy film is corroded plot by plot when the surface morphology is observed (the other two Fe-C alloy films are corroded gradually). The reason of that phenomenon should include two aspects: one is that the $3.0 \mu\text{m}$ Fe-C alloy film is too thin, the other one is that the bond between the alloy film and the fiber core is too weak. From the curve of the $7.6 \mu\text{m}$ and $12.8 \mu\text{m}$, we can perceive the corrosion process is basically smooth at the

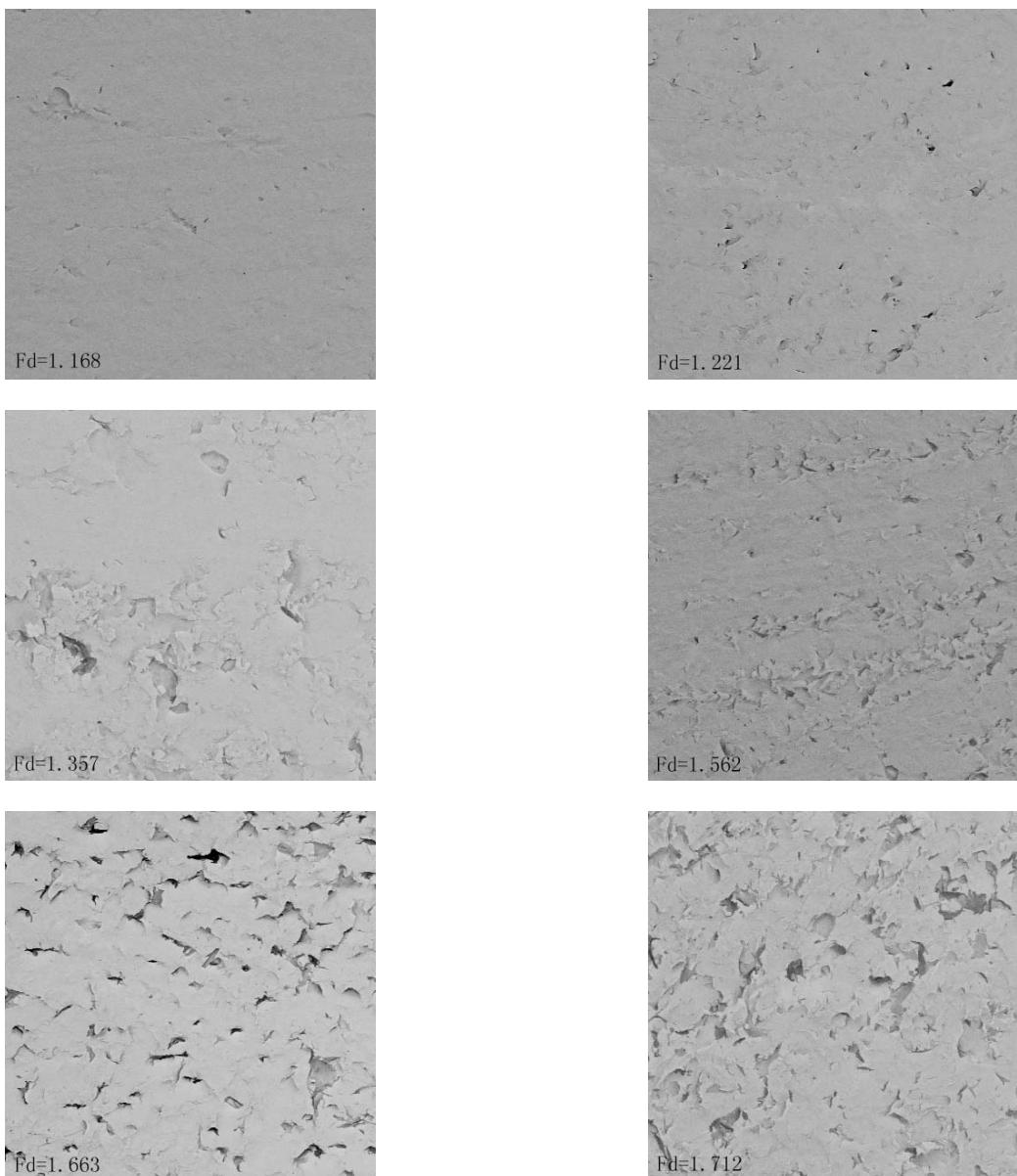
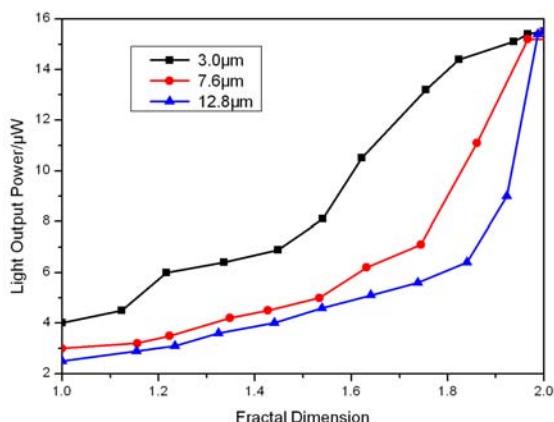


Fig.6 Fractal dimension of different corrosion images

beginning of corrosion, after that, the light output power starts growing. At last, the light output power increases rapidly because the Fe-C alloy film begins to drop plot by plot.

Fig.7 Variation of the light output power with fractal dimension of Fe-C alloy film corroded(10%HNO₃)

V. CONCLUSIONS

In this paper, we describe a novel method to characterize the Fe-C alloy film OFCS based on fractal corrosion images. Through reduplicative corrosion tests, the corrosion law of Fe-C alloy film OFCS is obtained basically. When the film thickness is larger, the corrosion process is smooth at the beginning and the light output power is increasing sharply at the end of corrosion, the phenomenon of which is because Fe-C alloy film begins to drop plot by plot at the end of corrosion.

In addition, we can find that the repeatability of experimental result is well in acid solution, and in the same conditions of the corrosion solution, the law between the light output power and the fractal dimension with the same thickness Fe-C alloy film is basically similar. Therefore, it is feasible that the fractal dimension is used to characterize the Fe-C alloy film OFCS and it is

also optimistic to the future prospect of the Fe-C alloy film OFCS.

ACKNOWLEDGMENT

This work is supported by the National Nature Science Foundation of Jiangsu, China (No.BK2006037) and the Technology Support Project of Jiangsu, China (No. SBE201000378).

REFERENCES

- [1] M. W. Woodruff, J. S. Sirkis: SPIE, 2191, 1994, pp. 511-515.
- [2] P. Rutherford, R. Ikegami, J. Shrader: SPIE, 3042, 1997, pp. 248-259.
- [3] P. L. Fuhr, D. R. Huston, A. J. McPadden, R. F. Cauley: SPIE, 2719, 1996, pp. 229-237.
- [4] Dong.Saying,Peng.Gangding,Luo.Yanan.Preparation techniques of metal clad fibres for corrosion monitoring of steel materials.Smart Materials and Structure, 16 , 2007,pp. 733-738.
- [5] J.M. Costa, F. Sagües, M.Vilarasa.Fractal patterns of corrosion pitting. Corrosion science, 32, 1991,pp.665-668.
- [6] E. Sarmiento, J.G. González-Rodríguez , J. Uruchurtu. A study of the corrosion inhibition of carbon steel in a bromide solution using fractal analysis. Surface & Coatings Technology,203,2008,pp.46-51.
- [7] Shaniavski, A.A., Artamonov, M.A.. Fractal dimensions for fatigue fracture surfaces performed on micro-and meso-scale levels. Internet J. Fract. 128, 2004,pp.17-24.
- [8] Weng, Yongji, Li, Xiangyi. Study on characterization for soil corrosion rate fluctuation of carbon steel. Chemistry, China, 7, 2005, pp.547-550.
- [9] S. J. Xu, Y. G. Weng: Pattern Recognition Letters, vol. 27,2006.
- [10] X. M. Li, W. M. Wei, Z. Q. Huang,et al. A new type of fiber optic corrosion sensor. Chinese Journal of Sensors and Actuators, 32,1999,pp.176-180.

Lihua Cai was born in 1984. She is a Ph.D. candidate of China University of Mining and Technology at Xuzhou city in Jiangsu Province China. From Sept. 2003 to Jul. 2007, she was Mechanical Engineering and Automation major at China University of Mining and Technology. After finished her bachelor, she was admitted to a ph.D. student without examination by China University of Mining and Technology in the field of Artificial Intelligence under the supervision of Prof. Wei Li. Her main research interests include automated reasoning and intelligent planning.

Wei Li was born in 1964. He got M.S degree in 1989 and Ph.D degree in 2004 at China University of Mining and Technology at Xuzhou city in Jiangsu Province China. He is a professor and doctoral supervisor at China University of Mining and Technology and he is the dean of school of Mechanical and Electrical Engineering at China University of Mining and Technology. His main research interests include distributed artificial intelligence, knowledge engineering and intelligent instrument,etc.

Call for Papers and Special Issues

Aims and Scope.

Journal of Computers (JCP, ISSN 1796-203X) is a scholarly peer-reviewed international scientific journal published monthly for researchers, developers, technical managers, and educators in the computer field. It provide a high profile, leading edge forum for academic researchers, industrial professionals, engineers, consultants, managers, educators and policy makers working in the field to contribute and disseminate innovative new work on all the areas of computers.

JCP invites original, previously unpublished, research, survey and tutorial papers, plus case studies and short research notes, on both applied and theoretical aspects of computers. These areas include, but are not limited to, the following:

- Computer Organizations and Architectures
- Operating Systems, Software Systems, and Communication Protocols
- Real-time Systems, Embedded Systems, and Distributed Systems
- Digital Devices, Computer Components, and Interconnection Networks
- Specification, Design, Prototyping, and Testing Methods and Tools
- Artificial Intelligence, Algorithms, Computational Science
- Performance, Fault Tolerance, Reliability, Security, and Testability
- Case Studies and Experimental and Theoretical Evaluations
- New and Important Applications and Trends

Special Issue Guidelines

Special issues feature specifically aimed and targeted topics of interest contributed by authors responding to a particular Call for Papers or by invitation, edited by guest editor(s). We encourage you to submit proposals for creating special issues in areas that are of interest to the Journal. Preference will be given to proposals that cover some unique aspect of the technology and ones that include subjects that are timely and useful to the readers of the Journal. A Special Issue is typically made of 10 to 15 papers, with each paper 8 to 12 pages of length.

The following information should be included as part of the proposal:

- Proposed title for the Special Issue
- Description of the topic area to be focused upon and justification
- Review process for the selection and rejection of papers.
- Name, contact, position, affiliation, and biography of the Guest Editor(s)
- List of potential reviewers
- Potential authors to the issue
- Tentative time-table for the call for papers and reviews

If a proposal is accepted, the guest editor will be responsible for:

- Preparing the “Call for Papers” to be included on the Journal’s Web site.
- Distribution of the Call for Papers broadly to various mailing lists and sites.
- Getting submissions, arranging review process, making decisions, and carrying out all correspondence with the authors. Authors should be informed the Instructions for Authors.
- Providing us the completed and approved final versions of the papers formatted in the Journal’s style, together with all authors’ contact information.
- Writing a one- or two-page introductory editorial to be published in the Special Issue.

Special Issue for a Conference/Workshop

A special issue for a Conference/Workshop is usually released in association with the committee members of the Conference/Workshop like general chairs and/or program chairs who are appointed as the Guest Editors of the Special Issue. Special Issue for a Conference/Workshop is typically made of 10 to 15 papers, with each paper 8 to 12 pages of length.

Guest Editors are involved in the following steps in guest-editing a Special Issue based on a Conference/Workshop:

- Selecting a Title for the Special Issue, e.g. “Special Issue: Selected Best Papers of XYZ Conference”.
- Sending us a formal “Letter of Intent” for the Special Issue.
- Creating a “Call for Papers” for the Special Issue, posting it on the conference web site, and publicizing it to the conference attendees. Information about the Journal and Academy Publisher can be included in the Call for Papers.
- Establishing criteria for paper selection/rejections. The papers can be nominated based on multiple criteria, e.g. rank in review process plus the evaluation from the Session Chairs and the feedback from the Conference attendees.
- Selecting and inviting submissions, arranging review process, making decisions, and carrying out all correspondence with the authors. Authors should be informed the Author Instructions. Usually, the Proceedings manuscripts should be expanded and enhanced.
- Providing us the completed and approved final versions of the papers formatted in the Journal’s style, together with all authors’ contact information.
- Writing a one- or two-page introductory editorial to be published in the Special Issue.

More information is available on the web site at <http://www.academypublisher.com/jcp/>.

Multi-party Dialogue Games for Dialectical Argumentation <i>Jinping Yuan, Li Yao, Zhiyong Hao, and Yanjuan Wang</i>	2564
An Approach of Trustworthiness Evaluation of Software Behavior Based on Multidimensional Fuzzy Attributes <i>Zhen Li and Junfeng Tian</i>	2572
Personalized Web Search Using Clickthrough Data and Web Page Rating <i>Xueping Peng, Zhendong Niu, Sheng Huang, and Yumin Zhao</i>	2578
Study on Learner Modeling in Adaptive Learning System <i>Bing Jia, Yongjian Yang, and Jun Zhang</i>	2585
Multi-satellite Monitoring SST Data Fusion based on the Adaptive Threshold Clustering Algorithm <i>Hongmei Shi, Han Dong, Lingyu Xu, Cuicui Song, Fei Zhong, and Ruidan Su</i>	2593
Description to Fe-C Alloy Film Fiber Corrosion Sensors by Fractal Corrosion Images <i>Lihua Cai and Wei Li</i>	2599

(Contents Continued from Back Cover)

An Approach to Interactive Affective Learning Algorithms <i>Chong Su and Hongguang Li</i>	2425
A New Error-Correcting Transmission Method for Dual Ring Fieldbus in CNC System <i>Lei Yang, Hu Lin, Dongfeng Yue, and Tianrong Gao</i>	2432
Numerical Simulation of Heavy Rail Quenching Process <i>Gongfa Li, Yuesheng Gu, Jianyi Kong, Guozhang Jiang, Jintang Yang, Liangxi Xie, Zehao Wu, Siqiang Xu, Yaping Zhao, and Gang Zhao</i>	2439
A Redundant FPGA Based Controller for Subsea Blowout Preventer Stack <i>Zenkai Liu, Yonghong Liu, Baoping Cai, Fei Wang, Zhili Chen, Xiaojie Tian, and Yazhou Wang</i>	2446
A Non-Standard Approach for the OWL Ontologies Checking and Reasoning <i>Yingjie Song, Rong Chen, and Yaqing Liu</i>	2454
Process Goose Queue Methodologies with Applications in Plant-wide Process Optimization <i>Jingwen Huang and Hongguang Li</i>	2462
A New Sub-topic Clustering Method Based on Semi-supervised Learning <i>Xiaodan Xu</i>	2471
Multi-feature Fusion Face Recognition Based on Kernel Discriminate Local Preserve Projection Algorithm under Smart Environment <i>Di Wu, Jie Cao, Jinhua Wang, and Wei Li</i>	2479
Safety Separation Assessment in Free Flight Based on Conflict Area <i>Zhaoning Zhang, Chang Sun, and Peng Zhou</i>	2488
Design of Three-axis ED Milling Machine Based on the PMAC Motion Card <i>Fei Wang, Yonghong Liu, Zhili Chen, Renjie Ji, Xiaojie Tian, and Zenkai Liu</i>	2496
A Case Study of Model Checking Retail Banking System with SPIN <i>Huilin Shi, Wenke Ma, Meihong Yang, and Xinchang Zhang</i>	2503
The Investigation of Fault Diagnosis Based on GA-HPSO-NN <i>Bin Peng, Chang-an Hu, Rong-zhen Zhao, and Yu-qiao Zheng</i>	2511
A Novel Approach to Hardware/Software Partitioning for Reconfigurable Embedded Systems <i>Linhai Cui</i>	2518
A Study on the Control Methods Based on 3-DOF Helicopter Model <i>Junshan Gao, Xinghu Xu, and Chen He</i>	2526
High Level Synthesis using Learning Automata Genetic Algorithm <i>Huijing Yang, Chunying Wang, and Ning Du</i>	2534
Application of Particle Swarm Optimization in Figuring out Non-differentiable Point of Function <i>Yuanbin Mo, Xin-quan Zhao, and Shu-jian Xiang</i>	2542
A Robust Watermarking Against Shearing Based on Improved S-Radon Transformation <i>Minghui Deng, Qingshuang Zeng, and Xiuli Zhou</i>	2549

REGULAR PAPERS

Multiple Model Predictive Control of Component Content in Rare Earth Extraction Process <i>Hui Yang, Rongxiu Lu, Kunpeng Zhang, and Xin Wang</i>	2557
---	------
