

Structure of Automatic Judge System Mode Based on the Project Design of Independent Type

Wu Da Qin

Department of Information Engineering, Jilin Business and Technology College, Changchun, China

Email: liulan2003@sina.com

Yan Hui

Department of Information Engineering, Jilin Business and Technology College, Changchun, China

Email: liulan2003@sina.com

Abstract—Based on CDIO engineering teaching idea which emphasizes “learning while doing” and “project education”, the essay constructs automatic judge system model of software project design in network teaching platform with software testing technology. The essay solves the problem of automatically generated black-box test cases with genetic algorithm. In particular, two methods have been adopted to construct and calculate values of the fitness function, namely “cartridge construction method of branch function” and “construction method of the real time state of software”. The problem that structure test of abnormal operation program is solved with the method of key points matching by the pattern matching of regular expression.

Index Terms—Genetic Algorithm, CDIO, B/S Model, Black-box Test, Regular Expressions, Pattern Matching, static disassembling

I. INTRODUCTION

CDIO (Conceive — Design — Implement — Operate) engineering education mode is the concentrated generalization and abstract expression of Learning while doing and basing on project education. The mode integrates all curriculums systematically through project design. All contents that need to be learned and mastered center on the project design that is the core and fuse together with the core, forming a whole. Using CDIO engineering education mode to develop talents, every profession must have crafted conception, design, implementation and operation project to guide students to develop learning interest in core courses, improve their abilities and have a sober overall understanding of the major. The CDIO project is divided into 3 levels according to the scale and scope. Level 1 shows the projects that contain core curriculums and ability requirements in the major; Level 2 shows the projects that contain a set of relevant core curriculums and ability requirements of the project; Level 3 shows the projects that are set for enhancing the ability and understanding of the course within this individual course. The establishment of Level 3 projects and the forms of them are determined by the syllabus according to need. Each major has at least two times Level 1 projects. Level 1 project can be made in grade-one or grade-two students to

introduce core contents of the major and mobilize interest of students. The graduation fieldwork + graduation design should be made as the level 1 project for the second time. Each Level 1 project as a course is scored individually based on 3-5 Level 2 projects, strengthening the study and application of core courses in the major^[1-2].

With the development of network technology, the platform of engineering education has already expanded to rich resources and constantly updated web space. The teaching mode has changed from the traditional tutorial teaching to the diversified teaching of network independence and exploring type. Under CDIO engineering education idea, independent learning of students still need to focus on project design, such as the individual design or team design that students make in the laboratory or practice base. When students are making project design, they are out of teachers' guidance, which helps cultivate their comprehensive, practical and innovative ability. But the evaluation and guidance of the key steps in design is crucial, which can guide design achievements to correct, perfect and high quality direction.

So in professional computer teaching, the system which apply CDIO engineering education idea in network teaching platform enables students to test and evaluate the current design achievements anywhere at any time and give feedback on time to realize the quality performance of project design. Meanwhile, online testing system can reduce the workload of teachers and improve teaching efficiency. This essay mainly aims at the project design of Level 3. Teachers in projects design syllabus specify multiple project design topic for students to choose, and provide detailed performance requirements.

This essay solves the problem that how to give valuation and guidance for the student's uploading program, and the program maybe executed normally or maybe not. The problem of automatically generated black-box test cases can be solved with genetic algorithm. In particular, two methods have been adopted to construct and calculate values of the fitness function, namely “cartridge construction method of branch function” and “construction method of the real time state of software”. The problem that structure test of abnormal operation

program is solved with the method of key points matching by the pattern matching of regular expression^[3-4].

II. OVERALL STRUCTURE ANALYSIS AND DESIGN SCHEME

Online judge system model in the paper adopts B/S mode, as shown in Figure 1 below. Students login from the user interface of the network teaching platform and then upload the source code edited in the project design to the remote server of teaching platform. The back database receives the upload files and then compiles the program. System can support C, C++, Java and some other languages. Students can set the language for using when they submit programs. Evaluation system chooses corresponding compiler way according to the language that the users specify and compiles according to the set compilation command. If a program can generate an executable program after the compilation, links, then it will enter into the black box testing, namely ignoring the program's internal structure properties to test whether software performance meets the design standards. If the results of the program compilation have mistakes so that it can't run, then it will enter key tests, namely making source code the text to be processed and the key point of the program regular expression. Matching degree of key points can be got with the pattern matching method. Then give the evaluation for students according to the matching degree.

This essay mainly constructs automatic judge system model which is at the server of teaching platform. When the source code edited in the project design is compiled, it maybe turn into the executable program, but maybe not. On the basis of the result, it can select the testing mode such as black box testing or key points testing. This essay generates black-box test cases with genetic algorithm. The fitness function is essential to making genetic algorithm effective, so how to constructs and calculates values of the fitness function is the important problem. This essay solves the problem with two method, namely, "cartridge construction method of branch function" and "construction method of the real time state of software". One of the two methods, cartridge construction method of branch function, is in terms of the cartridge rules of branch function of Tracy by plugging in code at the source code. But the other method, construction method of the real time state of software, is only executing the executable program. The key points testing is also need to use the source code of the design into matching by regular expressions. The shift of testing selection is shown in Figure 2.

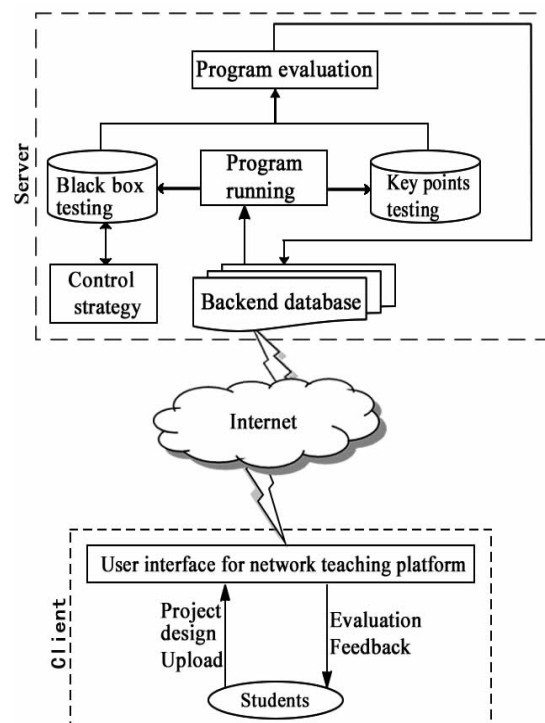


Figure 1. Overall Structure

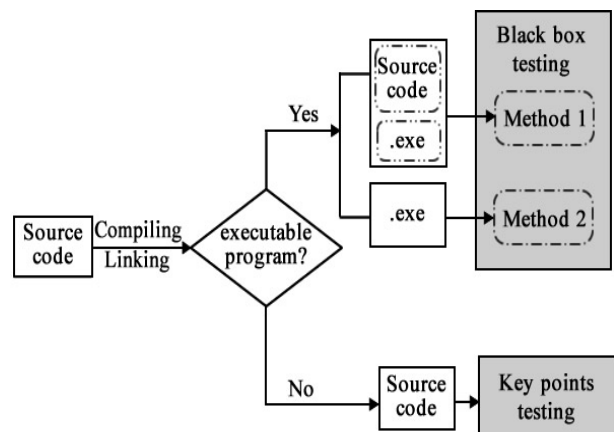


Figure 2. Overall Structure

III. BLACK-BOX TESTING BASED ON GENETIC ALGORITHM

Teachers will pre-put the project design topics and performance requirements specified in project design syllabus to the test database. If the compiled source program submitted by students can run, then it will turn to the black box testing. Online system black box testing work principle is: according to the requirement of project design that students selected, choose and automatically generate some typical test cases. The corresponding parameter scope and useful parameters will be extracted and fitness function shall be constructed so as to make the leg penetration of measured functions completed, then concrete pile test cases can be generated. Use these test cases in programming interface to check whether the program function can be realized and then evaluate the program.

This paper generates test cases through genetic algorithm. It selects few typical test cases from all the available test cases to achieve the maximum of test coverage, which plays a key role in black box testing.

A. The structure of genetic algorithm

Genetic done is a search optimization algorithm based on biological simulation genetics and natural selection mechanism. The whole evolution process starts from initial population, determines fitness function, according to the features of tested software and assesses the fitness of each test case after the execution procedures. The higher the fitness is the closer between test cases and the expected effect. Improve the test cases with crossover, mutation and choice- three basic operations. Evaluate the fitness function. The process can be ended until the result is closest to the best expectation. Simple genetic algorithm is shown in Figure 3^[5-7].

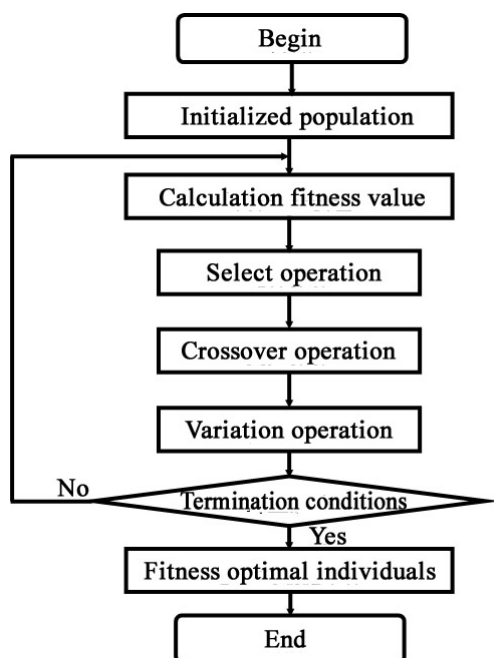


Figure 3. Simple Genetic Structures

B. Code of the Parameters

As a tested program in general it has one or more interface parameters with several same or different data types, such as integer type, string type, enumeration type etc. This paper codes the parameters with the real number coding method, namely, using problem parameter as chromo gene. Two groups of chromosome genes are set. Each line represents a test case and each column represents the input parameters of program or function and the type attribute of each column has corresponding parameter type.

$$w = (w_1 \dots w_m \dots w_n) \longleftrightarrow a = (a_1 \dots a_m \dots a_n)$$

The first m of genes are continuous genes, the later one is a discrete gene. Chromosome length is of the problem parameter and the value of the gene is the interval of

corresponding parameter value. Chromosome (individual) space and problem solution space are the same.

For example, the parameters of program interface are defined as follows:

X: Int range 0..40000;

Y: Enum (T1, T2, T3);

Z: Int range 0..255;

W: Short range 0..255;

The coding is shown in Table 1, every column represents a parameter. Every line represents a complete test case. Label the data type of each column to select the mutation and crossover operation of the corresponding data types.

TABLE 1 PARAMETERS OF THE CODE

Case serial number	X	Y	Z	W
1	0	T ₁	0	0
2	5	T ₁	5	10
...

C. The Selection and Improvement of Genetic Operations

In the project test, the goal is to choose the typical test cases to achieve the maximum coverage of the test domain. In order to achieve this, genetic algorithm begins in some random populations, which are some probable test cases of some substances called chromosome. Each chromosome is used as the input of measured code. Based on the result, an adaptive value is calculated to evaluate the pros and cons of chromosome for solving problems. Use this adaptive value to produce a new chromosome population, and the chromosomes with high adaptive value will also have high survival rate. Through crossover and mutation, the operation of the multiplicity of chromosome is endowed to produce the next generation again until selection and crossover reach the stop condition.

For the Black box testing of programs, this paper adopts the three basic operations provided by genetic algorithm: selection, crossover and mutation to improve efficiency^[8-9].

1) *Selection*: Real number coding has no effect on select operation. Therefore the individual quality can be decided by the fitness function such as linear accelerating fitness function. After calculating fitness value of the group, rule out the individuals with low fitness ability and select individuals with high fitness as the father of the new generation.

2) *Crossover*: Crossover operation is a process that imitates genetic reconstruction and its role is to pass the good genes to the next generation and produce new individuals that contain more complex genetic structure. The design that realizes the crossover of the individual recombination is related with the specific problems in general. Any crossover needs to assure that the traits of excellent individuals in the prior generation can be inherited by the new generation. According to different

types of data, different types of crossover methods need to adopted.

a) *Single point crossover*: Select randomly from the generated individuals to make single point crossover and then get a new individual. Repeat the steps until all individuals are selected. Single point crossover method can be applied to all types of input parameters.

b) *Uniform crossover*: Uniform cross is the generalization of single-point crossover, which allows any number of cross points.

c) *Weighted crossover*: Weighted crossover adds weighted value on the basis of uniform cross. It enables individual genes with good fitness to be reserved when choosing chromosome of father generation according to size of weights. Usually the higher the individual fitness value of the father generation, the higher their corresponding weighted values.

d) *The mean crossover*: The mean cross over adds the father generation's gene values that are in the same position and then get the average of the values. The calculation methods for other types of data are the same with uniform crossover the mean crossover.

3) *Mutation*: Mutation is to select a test data (a_0, a_1, \dots, a_k), select a mutation position and replace the position with a random number. The mutation process will produce a new individual which will be added to a new generation. Variation introduces variability to the group, enhances the diversity and provides a method of escaping from local optimum.

D. Constructing Fitness Function

Fitness function (fitness function in genetic algorithm) is used to evaluate the good degree of individuals, to guide algorithm finally to find the function of test data which covers the specified path, which plays an important role in the success search of the genetic algorithm. The fitness function which is structured well not only enhances the possibilities of finding feasible solution but also produces better overall code coverage, and consumes less system resources at the same time.

Simple genetic algorithm usually selects one of the two methods below to transform objective function into individual fitness function, in order that individual fitness is negative^[10-11].

Method 1: The transformational method for solving the maximum optimized problem :

$$F(x) = \begin{cases} f(x) + C_{\min} & \text{if } f(x) + C_{\min} > 0 \\ 0 & \text{if } f(x) + C_{\min} \leq 0 \end{cases} \quad \text{Formula(1)}$$

Method 2: The transformational method for solving the minimum optimized problem :

$$F(x) = \begin{cases} f(x) + C_{\min} & \text{if } f(x) + C_{\min} > 0 \\ 0 & \text{if } f(x) + C_{\min} \leq 0 \end{cases} \quad \text{Formula(2)}$$

In smaller group scales, if certain genes have high fitness in the initial stage of the genetic algorithm running. Therefore, the high probability selection that they are selected will exist in simple selection method. The result can make the genes breed in large numbers. Most of the group become their offspring, so the diversity of the group has been reduced. This make the algorithm converges at the extremum pixels in local area. The global optimal solution can't be obtain. If most of genes have a high fitness value when the algorithm's running will come to a close, the highest fitness and average fitness of the group are similar each other. At this time, the genetic algorithm will select genes from the highest fitness and average fitness in the same probability. So the selection for genes with genetic algorithm is similar to selection in a random way. The fitness function is disable, at the same time, the algorithm is also invalid.

Base of the above statements, it is necessary for fitness function to be improved for feasibility and diversity of the genetic algorithm. The dispersed fitness values can be obtained, which make genes fitness different and keep the healthy and different competition. The genetic algorithm can gain excellent effect.

Some common conversion methods are listed below.

1) *Linear transformation*: the linear function to turn the optimal objective function into fitness function.

$$F' = aF + b \quad \text{Formula (3)}$$

2) *Power transformation*: the power function to turn the optimal objective function into fitness function.

$$F' = F^k \quad \text{Formula (4)}$$

3) *Index transformation*: the index function to turn the optimal objective function into fitness function.

$$F' = \exp(-\beta F) \quad \text{Formula (5)}$$

The fitness function is divided into Approximation Level, Distance Calculation or both. Approximation Level is the close program of the actual execution path and the local goal in which it will arrive. Distance calculation is to check the deviation between path in the branch node actually and the expected path, whose aim is to measure the close degree in predicate conditions of the specific test cases getting to expected paths. Application of the manner of distance calculations is taken in this paper, i.e. based on the deviation degree of actual execution paths in measured unit and designated logical path to construct the fitness function.

E. Generate Modules of Test Cases

Aiming at the source code, executable file of compiled binary, different generation methods to construct test cases are adopted in this paper.

1) *Cartridge Construction Method of Branch function*

In the last project source code, the cartridge of branch function shall be performed. According to the cartridge rules of branch function of Tracy, a real-valued function

fi (x1, x2, ..., xn) will be inserted before each branch point which the specific logic path of the interior of program unit passes, to construct branch function and evaluation function corresponding with branch predicates, the penetration of measured function can be completed.

Gaving the branch predicates cartridge inserted valve has two styles. One style is plugging when the if branch structure appears, the other is plugging when the while loop structure appears. The cartridge of the if branch structure is put cartridge statements next to the if branch sentences. These cartridge sentences can give the tester information which is displayed while the programme is executing.

For example, below is a paragraph of source code about gaining the max number from three numbers. It will be plugged-in cartridge statements.

```

if(x>y)
    if(x>z)
        max=x;
    else
        max=z;
    endif
else
    if(y>z)
        max=y;
    else
        max=z;
    endif
endif

```

The program above includes six branch predicates. We put the character lable identifying the branch behind every if branch predicates.

Plugged programm is shown below.

```

if(x>y)
    confit( 'a' );           // Cartridge
    if(x>z)
        confit( 'c' );       // Cartridge
        max=x;
    else
        confit( 'd' );       // Cartridge
        max=z;
    endif
else
    confit( 'b' );           // Cartridge
    if(y>z)
        confit( 'e' );       // Cartridge
        max=y;
    else
        confit( 'f' );       // Cartridge
        max=z;
    endif
endif

```

If the detector distribution is true, the lable will gain a "T" to identify. Or else the lable will gain a "F" to identify. Then an array will be build in order that the programm

has run the branch. If the programm has run the branch, the array's corresponding position will be written "T", or else it will be written "F". In the method, different paths can corresponding to different values.

When a set of test data drives the measured unit, these real-valued functions are calculated, and whose values will be reflected in the current group of test data. With the deviation degree of actual execution paths in measured unit and designated logical path, can the joint expression of function f1 be defined as $F = F[f1(x1, x2, \dots, xn), f2(x1, x2, \dots, xn), \dots, fm(x1, x2, \dots, xn)]$. Evaluation function F can be summed up as the function of x1, x2, ..., xn of shape parameter of measured unit, and its value is a good dimension to evaluate the actual value of shape parameter. About the specific form taken by F, a more simple even add form is chosen: $F = \sum f$. The experiment shows that the method of defining evaluation function has a good effect. This paper introduces the method of generation test case based on genetic algorithm, compared with the method of traditional generation test case by search optimization in time efficiency.

We can see that from Figure 4 below, with the increase of the branch number of measured program, the time delay of generated test cases of genetic algorithm increases linearly, the time delay of traditional search optimization algorithm increases exponentially, and generated test cases of genetic algorithm in time efficiency has obvious advantages.

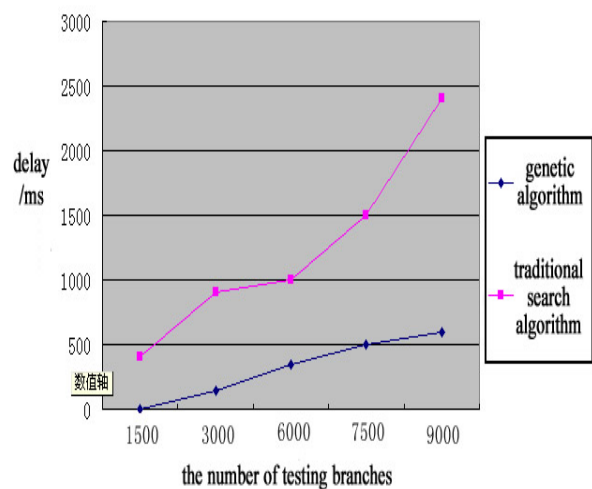


Figure 4. Algorithm contrast figure

2) Construction Method of the Real Time State of Software

In the previous introduced methods, it needs to program cartridge source code, and which is too complicated. The code block of dynamic record program of software in the running test cases needs to be adjusted. The concrete methods are: the source code submitted by students, after the process of compilation, links, so an executable program can be generated. And under the

TABLE 2 PARAMETERS OF THE CODE

regular expression	source code	matching result
<code>((for\(*\){0,1}\(.\+=,+\?; .+\[< >\].+\?; .+\?))\{0,1\}</code>	<code>for(err=1; err<cou;err++)</code>	succeed
<code>((for\(*\){0,1}\(.\+=,+\?; .+\[< >\].+\?; .+\?))\{0,1\}</code>		fail
<code>(if\(*\){0,1}\(.\+=,+\?; .+\[< >\].+\?))\{0,1\}</code>	<code>if(a<b)</code>	succeed

The structure test of key point matching is similar to the artificial static review in some aspects, but this review is limited to the given specific program in the database in advance. But in the actual independent learning, students will have many implementation algorithms in specific project design, and if not matching the algorithm given in the database in advance, it will cause evaluation mistakenly. Therefore, the CDIO teachers shall input various algorithm solutions to the database in project algorithms, then the program codes of students can be matched the codes of many solutions, and finally matching scores can be obtained respectively.

V. CONCLUSION

Based on CDIO engineering teaching idea, the essay constructs online judge system model of software project design in network teaching platform with software testing technology. The essay solves problem of automatically generated black-box test cases with genetic algorithm and structure test of abnormal operation program with the method of key points matching, improving learning efficiency of students and reducing the workload of teachers. As to how to further apply network technology and software testing technique, the research is still needed to make the automatic evaluation of various types of project design more accurately and efficiently.

ACKNOWLEDGMENT

The authors gratefully acknowledge the funding of this study by The Ministry of Education humanities social sciences studies "Research on Engineering Technology Talent Cultivation" special task "The Application of Information Technology in Talent of Engineering Technology Cultivation"(10JDGC021) and Education and research project of Education Department of Jilin Province "The Research and Practice in Innovative Engineering Technology Talent Training Mode Facing Strategic Newly Arisen Industry "and Jilin Province Educational Science "The Twelfth Five-Year" plan topics "The Research and Application of Expert System in the Cultivation and Evaluation to CDIO Mode Engineering Technology Talent Based on the System in Petri Net"(2011327) and the Jilin Province Department of Education research project "Research and Practise of CDIO Intelligent Teaching Platform Base on Internet of Things ".

REFERENCES

- [1] Wan Jinhua, "The Comprehensive Training Exploring Program Based on CDIO Mode", *Modern Enterprise Education*, 2009 (4).
- [2] Edward F.Crawley, "The CDIO Syllabus:A Statement of Goalsfor Undergraduate Engineering Education[EB/OL]", <http://cdio.org/cdio2syllabusrept/index.html>,2008-07-1.
- [3] Lun Lijun and Ding Xuemei, "The Research of Testing Data Based on Genetic Algorithm", *Computer Engineering*, 2005, 31 (23) : 82-84.
- [4] Zhang Junke, "Characteristics of Project Teaching and Its Requirements for Teachers", *Journal of Shanxi Vocational Institute of National Defense Technology*, 2009, (3) : 46-48.
- [5] Cheng Hua, "Automatic Scoring Scheme Design of the Subjective Items in Computer Language Classes", *Computer Learning*, 2006 (1): 16-17.
- [6] Zhao Xingtao, "The Application of Regular Expressions in Automatic Identification Document", *Journal of Chinese People's Public Security University*, 2005, 4:33-35.
- [7] Grosso C D, "Improving Network Applications Security:a NewHeuristic to Generate Stress Testing Data", *Proc.of GECCO'05.Washington D.C.,USA:[s.n.],2005*.
- [8] Li Z, Harman M and Hierons R M, "Search algorithms for regression test case prioritization", *IEEE Transactions on Software Engi- neering*,2007,23(4):225-237.
- [9] Sun Yajuan, "The research of automatic generating the path testing data on genetic algorithm", *Journal of Hebei industry university*,2006:30-40.
- [10] Zhuo Xiao lan,Lin Ying andZhang Jun, "Comparison of selection strategy in genetic algorithm", *Proceedings of the 12th Chinese Automation & Computing Society Conference in the UK, Loughborough, England, 2006* : 245-250.
- [11] Wang Zhuo wei, Yin Bao lin, "A network of automatic program based on survey system", *Journal of Beijing University of Aeronautics and Astronautics* , 2005,30(6) : 502-505.
- [12] PRAVEEN RANJAN SRIVASTAVA, PRIYANKA UPTA, YOITARRAWATIA, "Use of Genetic Algorithm in Generation of Feasible TeData", *SIGSOFT Software Enineering Notes*, 2009 (34) : 1-4.
- [13] Lima L, Iyoda J, "Test case prioritization based on data reuse an experimental study", *Lake Buena Vista,FL:3rd International Symposium on Empirical Software Engineering and Measurement, IEEE*, 2009:279-290.
- [14] Fazlalizadeh Y, Khalilian A, Azgomi M A, "Prioritizing test cases for resource constraint environments using historical test case performance data", *Beijing:2nd IEEE International Conference on Computer Science and Information Technology, IEEE*, 2009:190-195.



Wu Da Qin, a woman, was born in Songyuan city, Jilin Province on December 4th, 1979. She graduated from Chang Chun University where she earned the bachelor's degree of Computer Application in 2002. She received the M.S. degree in Computer Software and Theory area from Jilin university which is in ChangChun city, Jilin Province in 2007 year. She is a lecturer at Jilin Business

and Technology College in Chang Chun. Her research interests include Network, Multi-media and Artificial intelligence.

She has worked in computer instruction and researching for nine years at department of information engineering of Jilin Business and Technology College. She has taken charge of or participated in 8 national and provincial research projects such as taken charge of Jilin Province Educational Science "The Twelfth Five-Year" plan topics "The Research and Application of Expert System in the Cultivation and Evaluation to CDIO Mode Engineering Technology Talent Based on the System in Petri Net", published 7 scientific research papers, participated to write 2 books.

"Data Structure"(Beijing: Tsinghua University Press,2011)

"Study of assessment model on Educational for Applicative Talents Based on AHP" (Changchun: the journal of Changchun Teachers' College,2010)

Ms. Da Qin has awards the won the first prize of the provincial courseware contest and the second prize of the provincial Teaching Achievement Award.



Yan Hui , a woman, was born in Changchun city, Jilin Province on February 5th, 1971. She graduated from Jilin University where she earned the bachelor's degree of Computer Application in 1993. She received the M.S. degree in Computer Software and Theory area from Changchun university of science and technology which is in ChangChun city, Jilin Province in 2004

year. She is a professor at Jilin Business and Technology College in Chang Chun. Her research interests include Network, Data base and Artificial intelligence.

She has worked in computer instruction and researching for eighteen years at department of information engineering of Jilin Business and Technology College. She has taken charge of 12 national and provincial research projects , such as the Ministry of Education humanities social sciences studies "Research on Engineering Technology Talent Cultivation", "The Application of Information Technology in Talent of Engineering Technology Cultivation" and Education and research project of Education Department of Jilin Province "The Research and Practice in Innovative Engineering Technology Talent Training Mode Facing Strategic Newly Arisen Industry "and Jilin Province Educational Science "The Twelfth Five-Year". She published over 20 scientific research papers, led the compilation of 7 text books.

"Data Structure"(Beijing: Tsinghua University Press,2011)

"Computer Assembly and Maintenance"(Beijing: Tsinghua University Press,2009)

A novel bimodal identification approach based on hand-print IEEE(EI: 083911598935) 2007.12

Personal identification based on multi-modal hand-print features IEEE (EI: 083911600537) 2008.1

Evaluation and Study of Students Automatic Learning under Network Environment IEEE (EI: 20104413349500) 2010.7

Prof. hui is a president of information department of Jilin Business and Technology College, a engineer of CommerceNetwork technology, a venture consulting expert of Jinlin Province, a director member of Jilin Computer Federation, a executive director of Jilin Higher education Academic. She is teaching the provincial outstanding levels. She conferred the title of "the leaders of Jilin technology".