Research on Intelligent Test Paper Generating System Based on Improved Genetic Algorithm

Qilong Yin1, *

1 East University of Heilongjiang, Harbin, Heilongjiang, China, 150066

*Corresponding author e-mail: yinqilong163@dfxy.net

Abstract. Intelligent test paper generation belongs to an important research topic in the field of computer-aided education. Among them, Genetic Algorithm (GA), as an algorithm with high efficiency and performance, is widely used in the intelligent test paper generating system. But generally speaking, traditional GA always has some problems, such as local optimal solution and prematurity. The traditional genetic algorithm is improved in the aspects of the initial population, fitness function and some genetic operators to make it more in line with the requirements of intelligent paper generation, improving the efficiency and success ratio.

Keywords: Intelligent Test Paper Composition, Genetic Algorithm, Examination System

1. The present situation of intelligent test paper generating system

Question Paper generating refers to the user requirements for entering the target test papers according to the requirements. These requirements are transformed into the constraint conditions of generating test papers. Afterwards, the questions are extracted from the bank according to the corresponding algorithm to form the target test paper. A test paper has many constraint attributes conditions, such as the examination time, the score of the students who join in this test, the question numbers contained in various fields of questions, the types of questions included in the paper, and the score of every question. The reasonable value of these conditions will determine the quality of a test paper.

In recent years, colleges and universities have been committed to promoting the development of campus information construction. The academic Affairs Office is the core department of teaching management in colleges and universities, which is responsible for almost all teaching management matters in colleges and universities, especially the examination work of all majors in the school at the end of each semester, from the beginning of arranging teachers to issue questions, to examination paper inspection, test paper printing, sealing and binding. The whole examination management process is complicated and trivial, the workload is huge and it is easy to make mistakes. This traditional, purely manual way of organization and management has many defects. With the development of Internet technology, there is an urgent need to design and develop an intelligent test paper database system to replace the traditional manual operation, which can achieve the separation of teaching and examination, and evaluate teachers' teaching quality objectively and impartially.

The intelligent test paper generating system has some basic requirements:
(1) The system needs to enter the examination questions of all subjects into the database. The database can organize the test papers according to the requirements, or import the existing test papers into the test paper database according to the format requirements. Then the test papers can be converted into documents that can be typeset and printed normally.

(2) Through the use of network database technology, the system should realize decentralized data collection, unified management, shared use and automatic processing.

(3) The operation of the system should be simple and convenient, and when users operate illegally, the system should be able to give friendly prompts to users. Also, the system should be able to run stably for a long time, have good response ability and supporting ability, and have good expansibility and maintainability[1].

Figure 1. The functional structure of the examination paper generating system.

2. Advantages and disadvantages of genetic algorithm in generating paper system
The algorithm of the system is the core of the question bank and the main basis to measure the quality of a test question bank system. At present, the main methods of extracting test papers from the question bank are random selection algorithm, backtracking trial algorithm, genetic algorithm and so on[2].

2.1. Random selection algorithm
The method of randomly selecting test papers can produce ideal test papers according to the requirements of users only when the number of questions is small and the condition is reasonable. Otherwise, the selection procedure will be trapped in the non-ideal test area for repeated cyclic selection of questions, and will be inability to produce papers that fit the claims of users.

2.2. Backtracking trial algorithm
The backtracking test paper algorithm needs a lot of retrospective trial operations in the test paper selection, making it no longer correspond with the requirements of the network examination paper proceedings, which often changes the questions in the question bank. The feature of the backtracking trial algorithm is that the random selection process is repeated continuously until the test paper is finished or fails. This repeated process has a serious impact on the efficiency of generating paper, also affects the rate of success[3].

2.3. Genetic algorithm (GA)
GA is a random searching process, in which optimization method can be found through simulating the basic mechanisms of natural selection and genetic mutation in biology to solve complex problems. It uses the natural law of survival of the fittest to select individuals. GA helps produce the next generation population through mating and variation, and evolve generation by generation until the conditions are met.
Evolutionary computation does not have to describe all the characteristics of the problem very clearly, but only produces new optimal solutions according to natural laws. GA uses simple coding techniques to represent a variety of complex data structures, and determines the search direction through simple genetic operations and natural selection mechanisms.

The group searching strategy of GA gives a favorable solution for multi-objective optimization. Generally speaking, this type of problems does not have an optimal decision, and all possible outcomes are considered as non-inferior solutions. Traditional techniques always obtain one of the non-inferior solution sets every time, but more best solutions can be acquired when using genetic algorithm to achieve the goals.

Because the GA implements the parallel search in the worldwide which is really large, and in the process of the search, this space is constantly adjusted to the correct direction, making it easier to find the optimal outcome.

Compared with the traditional algorithm, it can quickly find the solution that meets the multiple constraints of the test paper, and has better applicability and efficiency[4].

On the other hand, although genetic algorithm has many advantages, it has some problems, such as low success rate, local optimal solution and precocious.

3. Applications of improved GA in paper generating system

Based on some disadvantages of traditional genetic algorithm, researchers have improved it in recent years.

3.1. Improved method of genetic algorithm

The solving process of GA primarily covers coding and population initialization, calculating population individual fitness, genetic operation and preserving optimal strategy, in which genetic operation includes three operators: selection, crossover and mutation.

According to the disadvantages stated in the second chapter, the methods partially improve the crossover, selection, coding, and mutation of the traditional GA, alleviating the problems of prematurity and local optimal solution to some extent.

(1) When creating the initial population, according to the various requirements of the users and question paper for the total score, question type and answer time, filtering the test questions that do not fit the demand of the question bank can realize the number of iterations in the calculation of the selected test paper algorithm, as well as improve efficiency.

(2) The decimal segmented real number coding method is used for coding. Compared with the traditional binary coding, the decimal piecewise real coding omits the coding and decoding process. It can make the coding more practical, shorten the solving time, and reduce the length of individual coding. This method is convenient for segmented crossover and mutation, and segmented coding is also beneficial to the initialization of the population.

(3) In the selection of genetic operator, the number of possible selected questions is calculated by the expected value model. And then the Bernoulli test is carried out by taking the decimal part of the selected individual as probability. If the experiment is successful, the individuals with high fitness will be selected as genetic operators.

3.2. Application of improved GA Test Paper generating system

3.2.1. The coding method of chromosomes

In the traditional binary system, it is hard for teachers to accurately control the distribution situation of test questions. And because of the corresponding type conversion, this method will seriously reduce the operation efficiency. So, the use of real coding will greatly improve the efficiency of the algorithm.

The number of test questions is much less than that in the question bank, and the piecewise real coding greatly shortens the coding length. This strategy embodies the test paper structure constraints
in the coding, such as the type and number of questions in the requirements, avoiding the combination of test questions that do not conform to the constraints in the genetic operation, which is helpful to promote the efficiency of questions composition. When the initial population is generated, the codes in each segment are randomly generated within the range of the sequence number of the corresponding question type\cite{5}.

Therefore, the decimal system is used to encode the questions numbers in the paper. According to the requirements of generating papers inputted by users and the principle that the examination contents of the papers cannot be repeated, N questions are randomized to be selected to generate the initial group.

![Figure 2. Implementation process of GA in the test paper system.](image)

### 3.2.2. Fitness function

The selection of this function is the key which can affect the performance of GA greatly. Generally, the fitness function is switched from the target function. The relationships of functions and performance are as follows: the smaller the objective function, the better; while the larger the fitness function, the better. Because the exponential ratio not only allows very good individuals to maintain more replication opportunities, but also limits the number of copies so as not to quickly control the whole population, and improves the competition among similar individuals, the exponential proportional transformation method is used to transform it into a fitness function.

### 3.2.3. Genetic operator

When selecting the operator, firstly calculate the number of times that each item number individual expects to be selected, and carry on the Bernoulli test on the result before extracting the test question successfully. Repeat the extraction process until all the questions meet the requirements\cite{6}.

A new genetic operator is designed according to the improved algorithm:

First of all, make the operator crossover, and when there is an illegal solution, such as the number of a question in a certain type of question is not in the correct range of the system, re-crossover operation or choose to eliminate in the population.

Then mutate the operator. Add a certain mutation rate (usually very small) to the constraint condition, and use this condition to select chromosomes. Delete a position randomly, then randomly
generate a title number that is not in the set, and insert it into the corresponding position within the scope of the question type.

4. Conclusion
Paper generation is one of multi-constraint optimization processes. And traditional GA has higher possibility to produce precocious problems and local optimization. The traditional GA is improved in some aspects, including initial population selection, some genetic operators and fitness function. As expected, the application of these improved genetic algorithms to the intelligent paper generation system can reform the shortcomings of the traditional GA, improve the efficiency and accuracy of the system when searching in the question bank, and have better test paper performance.

Acknowledgments
1. The work is supported by Scientific Research Projects of Heilongjiang East University (HDFKY200111). General online examination and score big data analysis system based on SSM;
2. The work is also supported by Natural Science Foundation of Heilongjiang Province (LH2020F040) Research on robust single target tracking method based on twin network.

References
[1] Ge Xuefeng, Ma Lu, Li Jie; A System Design and Implementation of Intelligent Test Paper Item Bank [J]; Computer and Information Technology; 2015.
[2] Yang Qiongfang; Intelligent Test Paper Generation Based on Improved Genetic Algorithm [J]; Journal of Inner Mongolia Normal University (Natural Science Edition); 2016.
[3] Chen Liyan; The Research on Algorithm in Intelligent Grouping Test Paper [J]; Journal of Hangzhou Teachers College; 2014.
[4] Qu Xiaotang; Research of the Intelligent Test Paper Composition Based of the Genetic Algorithm [J]; Computer Knowledge and Technology(Academic Exchange); 2007.
[5] Zhang Duan, Su Jianmin; Research of the Intelligent Test Paper Composition Based of the Genetic Algorithm in the Construction of the Examination Database [J]; Techniques of Automation and Applications; 2010.
[6] Ding Weiping, Guan Zhijin, Shen Xuehua; Research and Applying of the Intelligent Test Paper Composition Based of the Genetic Algorithm in the Construction of the Key Examination Database [J]; Journal of Electrical & Electronic Engineering Education; 2005.