Personalized Course Evolutionary Based on Genetic Algorithm

WenZhi Han

Sichuan Vocational and Technical College,
Suining 629000, China,
562212867@qq.com

Abstract

The paper presents an evolution of personalized courses based on genetic algorithms (PCEGA). The genetic algorithm are successfully applied in the dynamic update process of the course during the whole learning process. Under this framework of this algorithm, the target user model updates dynamically, and the courses evolve during the process. It provides a good general purpose and scalable framework that addresses the personalized course generation in an online learning environment.

***Keywords:** Personalized learning, Recommendation algorithms, Genetic algorithm, Knowledge map, Course generation*

1. Introduction

Genetic algorithm (GA) is a kind of from the biological evolution law (survival of the fittest, the survival of the fittest genetic mechanism) randomized evolved search method. The fundamental theorem of genetic algorithm -- the schema theorem was first proposed by Professor Holland in 1975 USA [1]. Its main feature is to operate directly on the structure of the object, there is no limit derivation and the continuity of function. With global implicit parallelism inherent and better search ability; the probability optimization method can automatically acquire and optimize the search space, to adaptively adjust the search direction, not need to determine the rules. These properties of the genetic algorithm has been widely used in combination optimization, machine learning, signal processing, adaptive control and artificial life etc. Since 2007, L.deMarcos will learn sequence is defined as a typical constraint satisfaction problem (CSP), GA began is applied to find the adaptation of the course in the solution space learning path [2-3], and is dynamic personalized courseware learning sequence [4], and is to generate test problems, it has reached the purpose of personalized learning [5].

Application of genetic algorithm may further research on Personalized Course Dynamic Evolutionary Algorithm in the course process, proposes the personalized course based on evolutionary genetic algorithm, namely Personalized Course Generation and Evolution Based on Genetic Algorithm. In this system, the concept of course knowledge in the field of genetic algorithm is as in the concept of the sequence of gene, personalized course is as chromosomes. Constraint relationships are between concepts based on the GA generation for solution space, learning relationship generation are concept of object and the concept of learning resources based on two-dimensional structured class. In GA, the difficulty of learning content, time, the knowledge level, cognitive level, local target characteristics are defined to the concepts and learning object metadata. And in the student model constructs of fitness function, and using adaptive order statistic ranking tournament selection method based on individual survival of the fittest, similar to the cross section of double cross point matching crossover algorithm crossover operation.

This paper proposes a personalized course dynamic evolution system based on GA. This

system generates a series of different characteristics of students learning to adapt to the different stages of the learning process throughout the course.

In this process, Build personalized knowledge state diagram for each student, and build dynamic student model. In order to make the generative course fully adapt to the personality characteristics of students, the system will combine features of students' cognitive ability, goals into the fitness function, and proposes specific fitness standards. Experiment and application results show the effectiveness of the fitness standard.

2. Realization of Personalized Course Generation and Evolution Based on Genetic Algorithm (PCEGA)

Because all the core concepts must appear in the chromosomes, so the chromosome coding is adopted in this paper to mark the location of each gene sequencing to determine the concepts in the course.

2.1. Gene Metadata

In order to make the individual needs of students reflected into the fitness function, gene metadata is expressed as:

$$Concept : (id, t^e, d, d_w, r)$$

Where, id is a concept of identification, t^e is the time parameter. d is a difficulty parameter concept, t_w is time weight, d_w is the difficulty of weights, r is a comprehensive feature gene values.

2.2. The Algorithm Flow

This section introduced based on genetic algorithm and process according to each individual student needs and evolution of personalized course. In general, the implementation process of the genetic algorithm can be divided into four stages: initialization, selection, crossover, mutation, For the optimization problem of curricular content, the solving process is a series of knowledge abstraction of concept for representation [6], it is optimal for learning content of the organic combination of the process. Course of evolution is as shown in Figure 1.

2.3. Realization of PCEGA Algorithm

2.3.1. The Initialization of Population. The initial population of genetic algorithm directly affects the accuracy of the scale and the performance of the algorithm. In general, the initial scale selection is decided according to the complexity of the problem. According to the results of prior [7-8], the initial value is between 20 and 100. Obviously, the initial population size is larger, the algorithm is executed the more slowly, but increase the probability of optimization [9-10]. Taking into account the course generation scale, we will be the initial population size is set to 50.

2.3.2. Fitness Function. As mentioned earlier, whether it is the best one chromosome eventually need, which is the measure of the concept of course provision and learning objects generated by the combination meets the individual needs of students. Whether to meet the need of students in the successful completion of learning tasks in difficulty. How to identify chromosome fitness? Determine the fitness function, so that the chromosome meets the

fitness function to solve the problem.

Gene (knowledge) characteristic value r is a comprehensive parameter of the concept of learning time and difficulty. Therefore, in this paper, genetic characteristics of the average value of r variance is to design the fitness function. The following describes the design of fitness function.
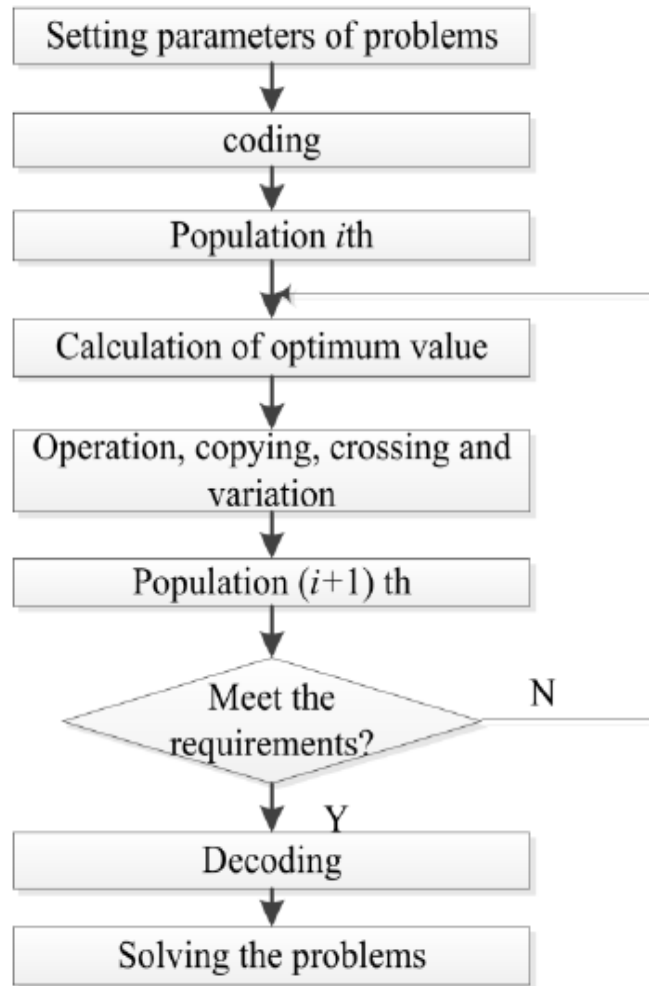
Figure 1. The Process of Course Evolution Algorithm

2.3.3. Selection Factors. The selection operation is the selection operator is used for the group, the purpose of study is to direct genetic outstanding individual to the next generation or by paired crossover generates new individual genetic to the next generation. Selection is based on the new individual fitness, but it does not mean completely by fitness level as a guide, because of simply selecting a higher fitness individuals may cause the algorithm rapid convergence to the local optimal solution rather than a global optimal solution, called early. As a compromise, based on the principles of course evolution algorithm: fitness is higher, the higher probability is to be selected, and adapt to the low degree, low probability is to be selected. The initial data can be obtained by this selection process to form a relatively optimal group. Then, the selected individuals get into the cross process.

In the learning process, according to the students' evaluation results obtained from learning content optimization, the smaller value of the fitness function of chromosomes were selected genetic to the next generation of probability is relatively high. In the practical design of the experiment, in the paper, according to the proportion of fitness assignment selection probability of the roulette wheel selection and ranking fitness assignment selection tournament selection method are compared. The results show that the former was measured when the genetic algorithm is crossover and mutation of one to two generations, the search range becomes very small, the population is 50, there are about half the same chromosome. Because the selection pressure is relatively large, the excellent chromosome obtained reproductive opportunities, leading to premature convergence. While the crossover probability and mutation probability are already quite high. If you continue to increase the mutation probability against the premature convergence problem, it will lead to the destruction of the optimal algorithm of information has been obtained. Therefore, in order to improve the above situation, makes the evolution regularity of population is more controllable, after repeated testing, this paper chooses the sort of fitness assignment method based on tournament selection method.

2.3.4. The Crossover Operation. The course content included rearrangement problem, namely the crossover operation, the purpose of which is to part structure of two parental chromosome replacement recombinant generates an excellent sub chromosome. In the genetic algorithm, the crossover operation is because of its global optimization ability as its main operation, plays a core role in the entire genetic algorithm. Through the cross, parental chromosomes are performance and genetic inheritance.

In practical design, how to choose the appropriate cross algorithm is crucial. Crossover operator is to ensure excellent individual performance of the parent in the offspring to get the possible genetic. In the relationship of genes can provide better performance and higher individual fitness, the single-point crossover is more suitable. In this paper, the object of strong randomness cannot guarantee this premise; multi-point crossover is very likely damage the good genes model parent, and thus is unable to obtain optimum results.

So this paper try to avoid damage has been good genes models father generation based on the principle of the genes in the chromosome, This paper is similar to double-cross point selected cross section cross-matching algorithm (PMX), and based on the actual content of the curriculum provision to design of the mapping relation.

3. Experiment Design and Discussion

Because of the selection operator and crossover operator of the genetic algorithm, making some excellent gene segment is lost prematurely, thus limiting the scope of the search, making the algorithm can only find local optima, and not the global optimal solution. To overcome the problem of premature convergence, the paper experimentally determine the relative good combination of parameters, these parameters include population size, crossover probability P_c, mutation probability P_m, select the basic crossover and mutation operators.

Control parameters have important effect on the performance of the system. Such as population scale, will influence the search space of the algorithm, unable to get the global optimal solution; the population size is too large then the algorithm reduces the searching efficiency. The crossover probability and mutation probability is bigger, stronger ability of detection algorithm, but easier to destroy the stability of algorithm, so as to detect the new hyper-plane, crossover and mutation probability is small, then the algorithm development ability is strong, can better maintain the suboptimal individual algorithm. However, in the

course of the generated application environment, need to learn practical data, determining the optimal combination of the parameters by experiment.

3.1. Determining the Crossover Probability and Mutation Probability

In the experiment, the initial population is set to 50, repeated tests to determine the probabilities of crossover and mutation, in order to achieve good results.

Firstly, the crossover probability is set to 0.5, the mutation probability for different values of, the population average fitness protocol and the best fit effect of chromosome generated. Figure 2-3 shows that the test and comparison results. Test results show that the influence of mutation probability, average fitness of the population is very small, average fitness and a sufficiently high degree will change the evolution trend of population.
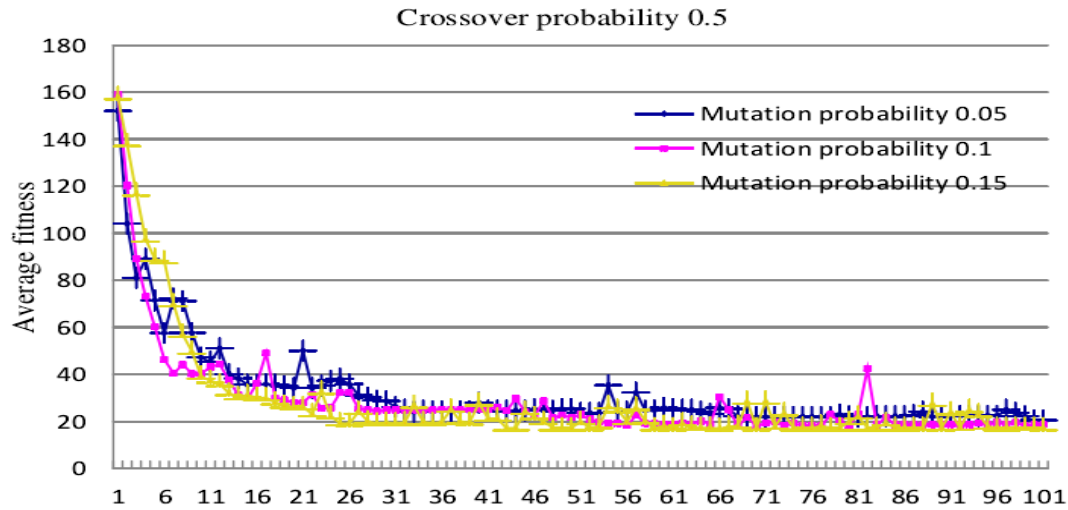

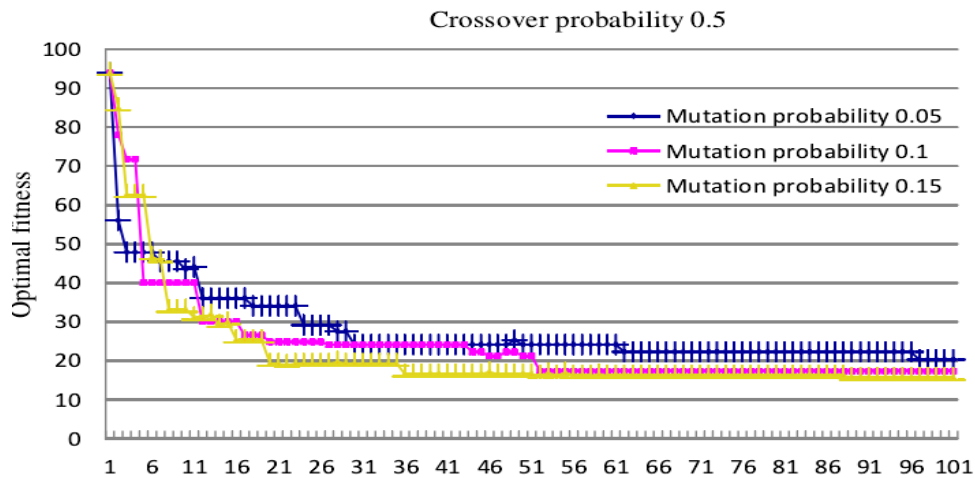**Figure 2. The Impact on the Mutation Probability for Average Adaptation**

Figure 3. The Impact on Mutation Probability for the Optimal Adaptation

Secondly, the mutation probability is set to 0.1, and the crossover probability take different values, the population average fitness protocol and the best fit effect of chromosome generated. Figure 4-5 shows the experimental results show that, the chromosome crossover probability larger easier to solve the optimal solution.

The above experiments show that, the larger the crossover probability and mutation probability can avoid premature convergence. However, a large number of new chromosomes, means the computation of large scale, the evolution of population will decline performance. Comprehensive consideration, after repeated trials, this paper determines the crossover probability is 0.5, the mutation probability is 0.15.
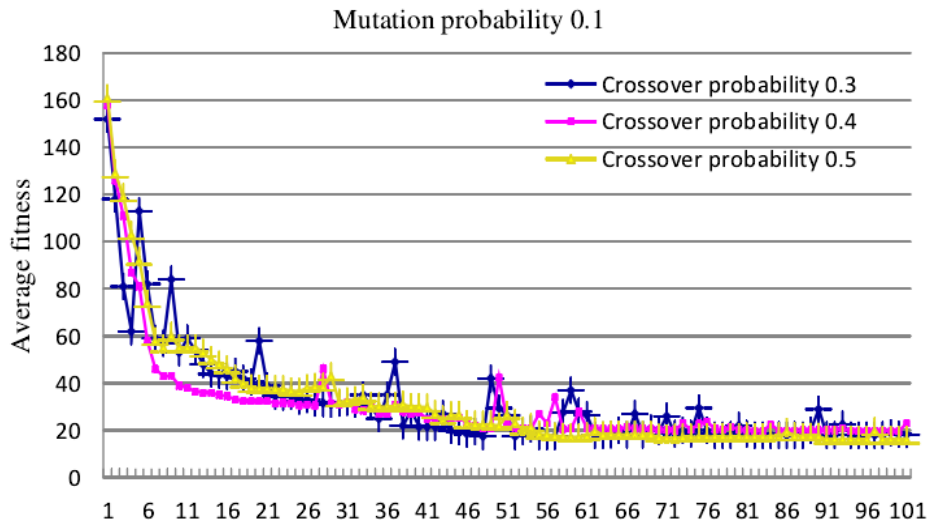

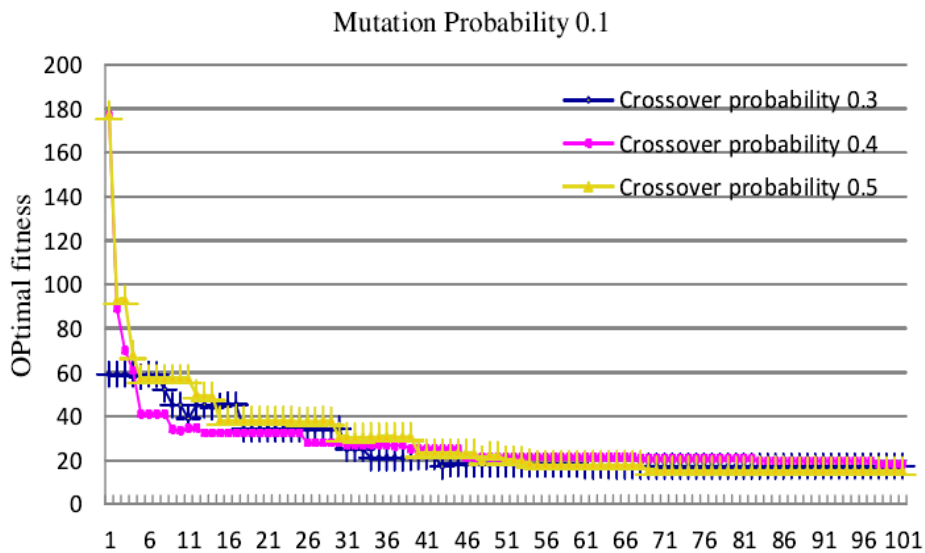**Figure 4. The Impact on the Crossover Probability for Average Adaptation**

Figure 5. The Impact on Crossover Probability for the Optimal Adaptation

3.2. Determining the Choice of Operator

In the experimental process operator choice, we compare the classic roulette wheel selection, Tournament Selection and order statistic competitive selection method.

Holland [11] proposed the roulette wheel selection (roulette wheel selection) is the basic principle to determine the probability of the individuals are selected according to the proportion of each chromosome, fitness, and establish the roulette wheel model to represent the probability. Individuals are selected with probability proportional to the fitness function value. Set the population size is N, the individual fitness of x_i is $f(x_i)$, the individual selection probability for x_i:

$$p(x_i) = \frac{f(x_i)}{\sum_{j=1}^{N} f(x_j)} \qquad (1)$$

Tournament Selection contains random and deterministic characteristics [12]. The method to select the best to propagate from some randomly selected in the chromosome. Wetzel proposed a stochastic competitive selection. The method uses the common method to calculate the probability of selection, and then the roulette wheel model has elected to the chromosome, with high fitness individuals into the new population.

Based on order statistics competitive selection method [13]. Set the population size is n, an individual is as sorted by grade ascending:

$$x_{p_1} \leq x_{p_2} \leq x_{p_3} \leq ... \leq x_{p_n}$$

Random variable $X_{i:n}$ represents the distribution of the corresponding x_{p_i}, the variable $X_{i:n}$ is called i-th order statistics. In the probability density $p_{i:n}(x)$ defined populations of statistics $X_{i:n}$ Select probability of i-th individual adaptation of x is calculated as follows:

$$p_{i:n}(x) = n\binom{n-1}{i-1} P(x)^{i-1}(1-P(x))^{n-1} \qquad (2)$$

Where, $P(x)$ represents the cumulative distribution function of x

In this paper, the above three methods are compared in the choice of operation, it is as shown in figure6-7. From the chart can be obtained, the standard roulette wheel method found in the measured when genetic algorithm crossover and mutation carry on one to two generations, the search range becomes very small, the population is 50, there are about half the same chromosome. Because the selection pressure is relatively large, the excellent chromosome obtained reproductive opportunities, leading to premature convergence. Competitive selection also encountered the problem of premature convergence, while increasing the probabilities of crossover and mutation, it still appear similar situation. If you continue to increase the mutation probability against the premature convergence problem, it will lead to the destruction of the optimal algorithm of information has been obtained. Therefore, in order to improve the above situation, makes the evolution regularity of population is more controllable, after repeated testing, this paper selects to order tournament selection based on statistic ranking method.
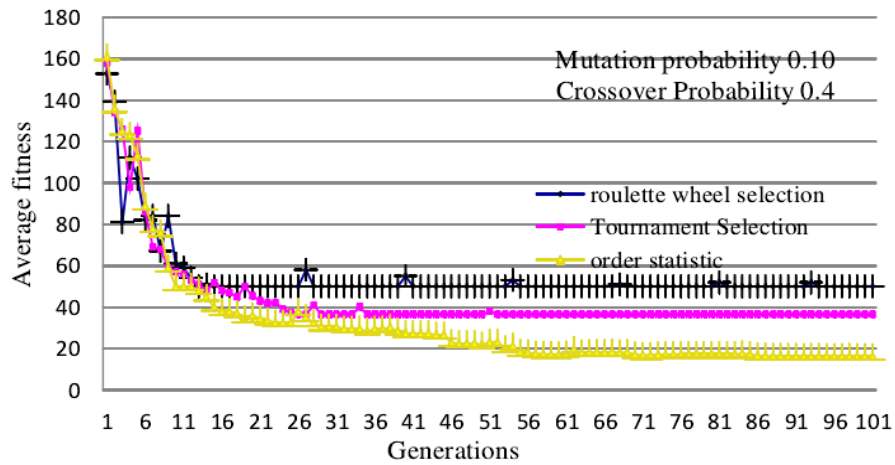
Figure 6. Impact on Different Select Operator for the Algorithm
($P_c = 0.4$, $P_m = 0.1$)
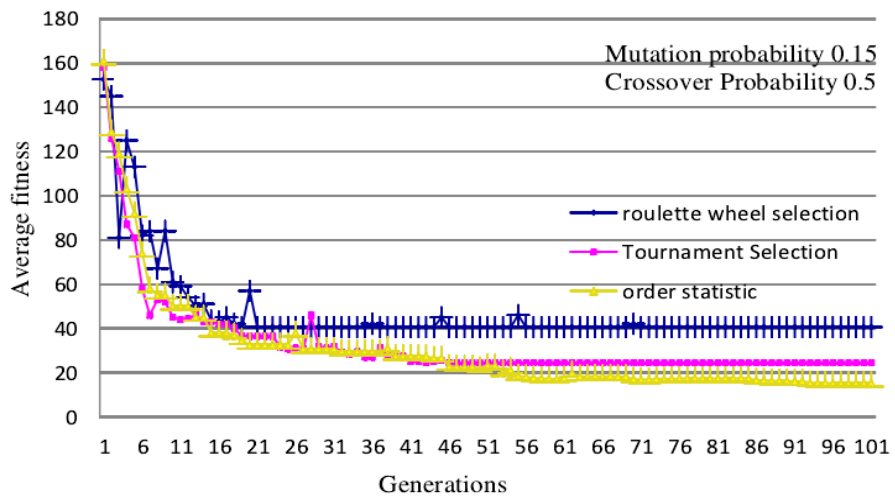
Figure 7. Impact on Different Select Operator for the Algorithm
($P_c = 0.5$, $P_m = 0.15$)

3.3. Determining the Crossover Operator

This paper try to avoid damage to have been based on the parent principle of good genes mode. Considering the chromosome gene must be present and are not repeated. This paper is similar to select the part of the double-crossing point cross matching cross algorithm (PMX), and in accordance with the practical problems of course to design mapping relation.

During the experiment, Data sets for the same student in the field of knowledge and subjects, respectively, of a single-point crossover, multi-point crossover and were compared with the experimental cross matching. Different crossover impact on the average fitness is as shown in Figure 8. When the set algebra was 1000, the crossover probability is 0.5, the convergence characteristics of the same data set is different. Convergence of single point crossover algorithm is 382 time, convergence of two point crossover is 396 times, and some

convergence matched crossover algorithm is 438 times. The results show that the proposed algorithm convergence times is higher than the traditional genetic algorithm, and has better stability and convergence, and ensures that the population in the optimal direction toward to genetic evolution.
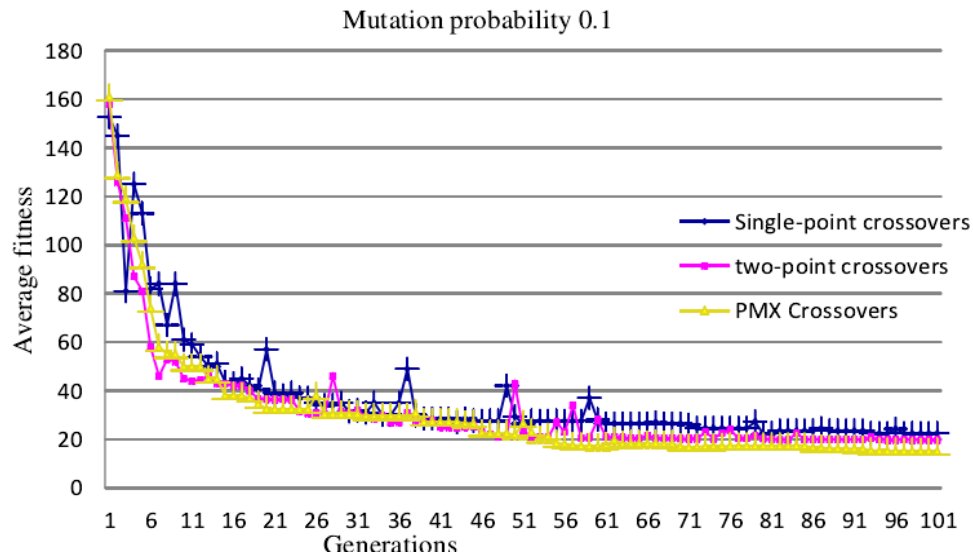

Figure 8. Impact on Different Crossover Algorithm for Average Fitness

4. Conclusion

This paper presents the evolution of PCE-GA model based on genetic algorithm personalized courses. In this system, the knowledge in the field of course concepts is as GA genes, the concept of individualized course sequence is as individual chromosomes, generate the solution space of GA based on constraint relations of concepts. In GA, the difficulty of learning content, time, the knowledge level, cognitive level, local target characteristics are defined to the concepts and learning object metadata, and in the user model, build adaptive function, according to the practical problems of course to design mapping relation. Through a series of experiments, from the effectiveness of personalized learning to start to determine the crossover rate, mutation rate, crossover parameters of genetic algorithm in the personalized learning content environment. Experimental results show that improving the efficiency and effectiveness of online learning based on PCE-GA personalized learning system.

References

[1] Y. Xiong, "Analysis of the telecommunications industry management system construction discussion", The accounting & OSS world, vol. 11, **(2004)**, pp. 43-44.

[2] J. Holland, "Adaptation in Natural and Artificial Systems", **(1975)**.

[3] C. Chen, "Intelligent web-based learning system with personalized leaning path guidance", Computer & Education, vol. 51, pp. 787-814.

[4] M. J. Huang, H. S. Huang and M. Y. Chen, "Constructing a personalized e-learning system based on genetic algorithm and case based reasoning approach", Expert Systems with Applications, vol. 33, **(2007)**, pp. 551-564.

[5] P. Arampiperis and D. Sampson, "Adaptive Instructional Planning using Ontologies", Proceedings of the IEEE international of conference on Advanced Learning Technologies (ICALT'04), **(2004)**.

[6] A. Menga, L. Yea, D. Royb and P. Padillab, "Genetic algorithm based multi-agent system applied to test generation", Computer & Education, vol. 49, no. 4, **(2007)**, pp. 1205-1223.

[7] L. Davis, "Handbook of genetic algorithms", Amsterdam: Van Nostrand Reinhold, **(1991)**.

[8] M. Bhaskar, M. M. Das, T. Chithralekha and S. Sivasatya, "Genetic algorithm based adaptive learning scheme generation for context-aware e-leanring", International Journal on Computer Science and Engineering, vol. 2, no. 4, **(2010)**, pp. 1271-1279.

[9] M. J. Huang, H. S. Huang and M. Y. Chen, "Constructing a personalized e-learning system based on genetic algorithm and casebased reasoning approach", Expert Systems with Applications, vol. 33, pp. 551-564.

[10] K. Jebari, A. EI Moujahid, A. Bouroumi and A. Ettouhami, "Genetic algorithms for online remedial education based on competency approach", 2011 International Conference on Multimedia Computing and Systems (ICMCS), **(2011)**.

[11] J. H. Holland, "Adaptation in Natural and Artificial Systems", MIT Press Cambridge, MA,USA. ISBN:0-262-58111-6, **(1992)**.

[12] D. Goldberg, B. Korb and K. Deb, "Messy genetic algorithms: motivation, analysis, and first results", Complex Systems, vol. 3, no. 5, **(1989)**, pp. 493-530.

[13] B. L. Miller and D. E. Goldberg, "Genetic algorithms, tournament selection, and the effects of noise", Complex Systems, vol. 9, no. 3, **(1995)**, pp. 193-212.

Author

WenZhi Han He received his bachelor's degree of engineering in Southwest University. He is an associate professor in Sichuan Vocational and Technical College. He is in the research of computer network and software.