A Study on the Genetic Algorithm for Class Combination Problem

Linjing Yan and Nan Wang

ABSTRACT

On the basis of the analysis of the joint-class problem, this paper established a math-model. With the expression of chromosome, it proposed a genetic algorithm for the class combination problem. Experimentation indicate that the arithmetic can find the optimal or nearly optimal solution.¹

INTRODUCTION

In many scheduling systems, the teaching staff always want to compile the most reasonable schedules. This will make full use of the school’s existing teaching resources. Among them, the problem of Class combination is a very critical issue, which directly affects whether the results of class scheduling are reasonable or even the success or failure of class scheduling. This issue can be described as follows: There are several classes that require the same course in most cases, it is impossible to teach a class separately. Therefore, some classes need to be arranged in the same classroom for teaching. This is the class combination problem. So how to reasonably arrange these classes in various classrooms, In other words, we consider which classes to combine, it’s the class combination problem we want to study. It’s the class combination problem we want to study. Under the premise that the number of lectures in the two courses is known, and the number of classes accommodated in each lecture platform is known, the class combination algorithm has obtained a better digital combination plan [1]. On the basis of it, with the goal of minimizing the difference in class professionality at each podium as a whole, in this paper, the genetic algorithm is used to solve the class combination problem. This includes establishing a reasonable mathematical model for the class combination problem,

¹Linjing Yan, Nan Wang. Beijing Polytechnic, Beijing, China
and designing genetic manipulation methods to construct a complete genetic algorithm.

**MATHEMATICAL MODEL OF CLASSCOMBINATION PROBLEM**

**The General Digital Combination Plan Obtained By the Joint Algorithm Is As Follows**

The program is a joint program for the two courses of \( x \) and \( y \). \( x_1, x_2, \ldots, x_n \) of them are the podiums of the first course. Its value indicates the number of classes accommodated. \( y_1, y_2, \ldots, y_m \) Of them are the podiums of the second course its value indicates the number of classes accommodated \( a_{ij} \). is the result of the digital joint program, It indicates the number of classes arranged on the \( x_i \) and \( y_j \) podiums.

<table>
<thead>
<tr>
<th>( y_i )</th>
<th>( a_{11} )</th>
<th>( a_{12} )</th>
<th>( a_{13} )</th>
<th>( \ldots )</th>
<th>( a_{1n} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_1 )</td>
<td>( a_{21} )</td>
<td>( a_{22} )</td>
<td>( a_{23} )</td>
<td>( \ldots )</td>
<td>( a_{2n} )</td>
</tr>
<tr>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
</tr>
<tr>
<td>( y_m )</td>
<td>( a_{m1} )</td>
<td>( a_{m2} )</td>
<td>( a_{m3} )</td>
<td>( \ldots )</td>
<td>( a_{mn} )</td>
</tr>
</tbody>
</table>

Figure 1. Joint program.

**Mathematical Model of Class Combination Problem**

It is known that the above one digital joint plan has \( M \) class, which belongs to \( m \) majors. Now it is distributed in \( n \) grids, which are divided into \( n \) categories.
Hypothesis:

$L_k$ — The total number of classes in the $k$ major of the $m$ majors;

$L_j$ — The number of classes in category $j$ of the $n$ categories (i.e. the value of $a_{ij}$);

$x_{kj}$ — The number of classes from the $k$th major in category $j$.

In order to ensure that the class professional differences in the overall table are minimal, there are the following objective functions:

$$f(x) = \min \left( \sum_{j=1}^{n} \left( \sum_{k=1}^{m} \text{sgn} x_{kj} \right) \right)$$

Restrictions:

$$s.t \quad \sum_{j=1}^{n} x_{kj} = k=1,2,3\ldots m \quad (1)$$

$$L_j = \sum_{k=1}^{m} x_{kj} \quad j=1,2,3\ldots n \quad (2)$$

GENETIC ALGORITHM FOR CLASS COMBINATION PROBLEM

As can be seen from the above model, solving the class combination problem is a reasonable assignment to each class that meets the constraints. In general, the classes in each podium are the least of the same number of majors. That is to say, the sum of the non-identical numbers of the classes in each podium is the smallest. Therefore, the genetic algorithm can be constructed as follows:

**Construct a Chromosome to Produce an Initial Population**

According to the specific situation of the joint work problem, this problem should be sequence coded. That is, "0482913576" can represent a chromosome. The meaning is the class number in the order of the digital work plan. The number is unique, The length of the chromosome is the number of all classes, then according to the serial number corresponding to the specific class, Put the class of the corresponding number of grids in the digital joint plan in the order just in the order, put it into the grid, that is to get the class combination plan.

The following uses the factorial method to generate the initial population: ① Numbers are assigned to $M$ classes, the serial number is unique, and each serial number corresponds to a specific class; ② Randomly select a number from 0 to $M-1$, which corresponds to the class as the first chromosome; ③ Renumber the remaining classes from 0 to $M-2$, Renumbering does not change their relative
order; ④ Randomly select a number from 0 to $M-2$, which corresponds to the class as the second digit of the chromosome; ⑤ Renumber the remaining classes from 0 to $M-3$, and renumber them without changing their relative order; ⑥ Randomly select a number from 0 to $M-3$, which corresponds to the class as the third digit of the chromosome; ⑦ And so on, until the last digit of the chromosome is selected, a chromosome is determined; ⑧ In the same way, other chromosomes are determined until the initial population size, i.e., the initial population generation.

Determination of Fitness Function

Finding the objective function value for each chromosome in a generation of populations, because the fitness function requires the largest, non-negative conditions, Therefore, the fitness function is determined to be $f_k = 1/f(x)$, Where $f(x)$ is the objective function. The class with the largest fitness function value corresponding to the objective function with the smallest objective function is the one we ask for[2].

Natural Selection

Using the adaptive value ratio selection and the elite retention strategy, the process is as follows: A total of $L$ chromosomes per population are ranked according to the fitness value $f_h$ from large to small, ($h=1,2,\hdots,n$), the top ranked individuals have the best performance. Copy it one and go directly to the next generation of stocks. The other $L-1$ chromosomes of the next generation population are generated by roulette from the $L$ chromosomes of the previous generation. The other chromosomes of the next generation population are generated by roulette from the chromosomes of the previous generation population, which can ensure that the optimal one survives to the next generation, and can avoid the chance of being selected due to the difference in the size of the individual. The disparity keeps the diversity of the next generation of individual populations, thus effectively improving the convergence speed of the algorithm[3].

Chromosomal Crossover

Selecting individual pairs by probability $P_c$ for new populations produced in 2.3, this paper chooses $P_c = 0.6$, because the class combination problem requires that the class within each chromosome is unique, a partial matching crossover method is used. The following examples illustrate, Let the two chromosomes of the father be $A = 4 6 9 2 5 7 1 3 0 8$, $B = 1 2 3 4 5 6 7 8 9 0$, According to the partial matching cross method, the intersection process is as follows:
\[ k_1, k_2, k_1, k_2 \]

A = 4 6 9 | 2 5 7 | 1 3 0 8 \rightarrow A_1 = 2 7 9 | 4 5 6 | 1 3 0 8

B = 1 2 3 | 4 5 6 | 7 8 9 0 \rightarrow B_1 = 1 4 3 | 2 5 7 | 6 8 9 0

Where the intersections \( k_1, k_2 \) are randomly selected, the fitness function values of each chromosome are calculated for the progeny generated after the crossover, and compared with the parent, and the individuals with large fitness function values are selected to enter the population.

**Mutation Operator**

In the population resulting from the above crossover, the chromosome is mutated at a mutation rate of \( P_m = 0.02 \), the mutation strategy is to randomly select two genes in the chromosome and let them exchange positions. Calculate the fitness value of the chromosome with successful mutation, compare it with the parent, and select the entry population with a larger fitness function value.

**Judge the Evolutionary Condition**

The final step in the design of genetic algorithms is to control parameters and algorithm termination conditions. The crossover rate and mutation rate have been given above, and the maximum evolution algebra is set to 150. The judgment process is as follows: Determine whether the algebra of the iteration is the maximum evolution algebra 150. If yes, stop the evolution, select the optimal chromosome as the best class combination scheme, or judge whether the fitness function value is the optimal value (that is, the classes of each platform are the same professional) The fitness function value), if it is, then stop evolution and determine the optimal chromosome as the best class combination scheme. Otherwise, proceed to 2.3.

**EXPERIMENT AND ANALYSIS**

This paper conducts experiments on a set of actual data from a university. The experimental results are as follows:

The two courses are based on advanced mathematics and computer science. There are 80 classes in which there are two classes, which belong to 14 majors. According to the actual situation of our school, there are 16 podiums in each course, but each class is distributed. The number of classes is different, the specific number is known, and the digital combination plan obtained by the combination algorithm and the class combination plan obtained by using the genetic algorithm of this paper are compared with the actual plan of our school as follows:
TABLE I. COMBINED ALGORITHM AND GENETIC ALGORITHM OPTIMIZATION SCHEME.

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>6</th>
<th>6</th>
<th>6</th>
<th>6</th>
<th>6</th>
<th>6</th>
<th>4</th>
<th>4</th>
<th>4</th>
<th>4</th>
<th>4</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>6/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>6/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>6/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>6/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6/3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4/1</td>
<td>2/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4/1</td>
<td>2/1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4/1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4/1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4/1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2/1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2/1</td>
</tr>
</tbody>
</table>

TABLE I shows the digital combination plan obtained by applying the combined algorithm and the class combination plan obtained by using the genetic algorithm of this paper. The first line indicates the number of classes in each stage of higher mathematics, and the first column indicates the number of classes of each platform of the computer foundation. The content part of the intermediate table is the result of the digital combined class plan, indicating the number of classes arranged in the corresponding two podiums; the denominator part indicates the class corresponding to the class in the class combination plan based on the above digital combination plan, using the genetic algorithm The number of majors. TABLE II shows the actual course plan of our school. The same numerator indicates the digital plan, and the denominator indicates the number of majors.

The data in the statistical table, using the genetic algorithm designed in this paper, the sum of the professional numbers is the objective function value of 21, and the actual course results of our school show that the objective function value is 30, it is obvious that the genetics designed by this paper The algorithm solves the result is superior.
At the same time, it is verified by full search. In this example, the probability of searching for the optimal solution by genetic algorithm is 100%. In addition, using genetic algorithm to solve other courses, the optimal probability of searching is above 98%. It can be seen that this method can find a satisfactory solution to the class combination problem, which is a better method. It is worth mentioning that its solution speed is much faster than full search, assuming that the total number of classes is $n$, the time complexity of full search is $O(n!)$, and the time complexity of the genetic algorithm is only $O(n)$.

**CONCLUSIONS**

In the course scheduling system, the importance of the teamwork problem is obvious. The solution of the class combination problem can completely solve the problem of the class. The research of the genetic algorithm is a feasible direction. Through the study of the class combination problem, this paper designs the genetic algorithm. The experiment proves that the algorithm has good performance and can
quickly find the optimal solution or satisfactory solution of the problem. It is a better solution to solve the class combination problem.

REFERENCES