Design and Research of Course Arranging System Based on Niche Improved Genetic Algorithm

Yazi Wang[1], *Aijv Lei*[2,*]

[1.] *ZhouKou Normal University, Henan, 466001, China*
[2.] *Zhoukou Central Hospital,Henan,466000, China*

Abstract: Aiming at the deficiencies in the commonly used arranging algorithms, this paper analyzes that the single arranging algorithm based on genetic algorithm has many problems, such as influencing the curricular factors, difficult to carry out optimal combination and low arranging efficiency, etc. The niche technology is introduced to make niche. The combination of technology and genetic algorithm to construct a hybrid genetic algorithm to improve the deficiencies of the single genetic algorithm in solving the problem of class scheduling, and a complex process is decomposed into two stages to ensure the rationality and improvement of the arranging results. The efficiency of automatic class scheduling is conducive to the manual adjustment of later courses.
Keywords: Niche technology; genetic algorithm; scheduling system

1. INTRODUCTION

CDIO is a model based on project learning. It is a kind of engineering education "middle school". In the project life cycle as the carrier, it cultivates the students' academic knowledge, professional ethics and the ability to use knowledge to solve problems[1]. The education model launched by the Massachusetts institute of technology (MIT) and the royal Swedish institute of technology have been used by dozens of universities and they are recognized by the industry. It is very necessary or the education reform and the introduction of CDIO education mode [2].

Education is a professional education, which is based on the completion of medium education. It is a social activity to cultivate senior professional talents[3]. Professional education is the training of knowledge and skills are necessary for the education to engage in a certain occupation. Therefore, professional education also refers to professional technology education or industrial education. "Higher professional education" is a combination of "higher" and "professional education".

2. THEORETICAL METHODS

(1). Niche technology

In biology, Niche refers to a living environment in a specific environment. In the course of its evolution, creatures always live together with their own species and breed their offspring together; they also all live in a particular geographical area. For example, tropical fish cannot survive in colder areas, and polar bears cannot survive in the tropics. This provides the possibility for the formation of new species and is one of the fundamental reasons for the biological community to maintain near infinite diversity.

In several typical niche implementation techniques, the following niche technology is used in this paper: By comparing the Hamming distances between individuals in two pairs, when the distance is within a specified distance L, the fitness is smaller. The individual imposes a strong penalty function, which greatly reduces its fitness to maintain the diversity of the group, thus achieving good results in overcoming premature convergence.

(2) Principle of Genetic Algorithm

The use of genetic algorithm to establish the weight distribution model includes the following steps:

(A) constructs the evaluation matrix by applying the normalization process to the indicators of the parking decision factors. From the aspects of scientificity, simplicity and applicability, and according to the specific situation of the parking decision model, the corresponding index system is established[4]. The fuzzy evaluation matrix of the single evaluation index of the relative membership degree is established. The purpose of the fuzzy comprehensive evaluation is to make a good comparison of several alternative parking lots, so as to choose the relatively good Parking lot, this good and bad comparison with the parking lot other than the alternative parking lot has nothing to do, according to this comparison finally gives the driver to determine the optimal decision parking decision. Therefore, assuming that there are n impact decision factors $\{(i, j) \mid \mathrm{i} = 1 \sim n, j = 1 \sim m\}$, the influence of the parking decision factor value $x(i, j)$ are non. In order to standardize and normalize the decision data of parking decision-making factor, we need to standardize the sample data set $x(i, j)$ which affects the parking decision factor, so that the established parking decision model is universal. To maximize the impact of the decision on the impact of

parking decision factors indicators of change, the impact of parking decision factors as the larger the better the type of standardized formula shown as shown below Equation 1.

$$r(i,j)=\frac{x(i,j)}{x_{\max}(i,j)+x_{\min}(i,j)} \tag{1}$$

The smaller the better decision-making factor for the decision-making factor, the better the factor is shown in Equation 2.

$$r(i,j)=\frac{x_{\max}(i,j)+x_{\min}(i,j)-x(i,j)}{x_{\max}(i,j)+x_{\min}(i,j)} \tag{2}$$

Where, $x_{\max}(i,j)$ is the maximum value of the parking decision-making factor, and $x_{\min}(i,j)$ is the minimum value of the parking decision-making factor, which affects the parking decision-making factor. (i,j) for the impact of the decision after the impact of parking decision factors, that is, the impact of the decision-making factor on the j-th alternative parking lot relative membership value, $i=1\sim n, j=1\sim m$. The fuzzy evaluation matrix $R(r(i,j))_{n\times m}$ formed by the $r(i,j)$ value obtained by the above standard is an element.

(B) According to the matrix obtained $R(r(i,j))_{n\times m}$ in step 1, the judgment matrix $B=(b_{ik})_{n\times m}$ is used to determine the weight of the decision factors that affect the parking decision factor. In the case of parking decision-making, if the influence of the parking decision factor index i_1, the sample series $\{r(i,j)\mid j=1\sim m\}$ is not the same as that of the selected alternative parking lot $\{r(i_2,j)\mid j=1\sim m\}$, the comprehensive evaluation information that affects the parking decision factor index i_1 is more than the influence of the parking decision factor i_2, which is more than the influence of the parking decision factor index i_2. For this reason, the standard deviation $\mathrm{CIC}(n)$ of the sample can be used to influence the decision factors of the parking decision, which reflects the influence of the decision-making factors on the parking lot decision-making, and the resulting sample difference as the element construction judgment matrix B.

$\bar{r}_i=\dfrac{\sum_{j=1}^{m}r(i,j)}{m}$ is the average of the series of samples under the influence of the decision factors. Therefore, the judgment matrix can be constructed according to the following equations:

$$b_{ik}=\begin{cases}\dfrac{s_{\max}(i)-s_{\min}(k)}{s_{\max}-s_{\min}}(b_m-1)+1, & S(i)\geq S(k)\\ \left(\dfrac{s_{\max}(i)-s_{\min}(k)}{s_{\max}-s_{\min}}(b_m-1)+1\right)^{-1}, & S(i)<S(k)\end{cases} \tag{3}$$

Where, $s_{\max}$ and $s_{\min}$ are the minimum and maximum values of $\{S(i),i=1\sim n\}$; the relative importance degree parameter value $b_m=\min\{9,\mathrm{int}[S_{\max}/S_{\min}]+0.5\}$.

(3) Calculate the weight $w_i(i=1\sim n)$ and the consistency check for the matrix B obtained in step 2, and satisfy the requirement: $w_i\geq 0$ and $\sum w_i=1$. According to the definition of decision matrix B, there is:

$$b_{ik}=\frac{w_i}{w_k}(i,k=1\sim n) \tag{4}$$

In this case, the matrix B has the following properties:

$$\begin{cases}b_{ik}=\dfrac{w_i}{w_k}(i,k=1\sim n)\\ b_{ii}=\dfrac{w_i}{w_i}=1\\ b_{ki}=\dfrac{w_i}{w_k}=\dfrac{1}{b_{ik}}\\ b_{ik}b_{kj}=\dfrac{w_i}{w_k}\cdot\dfrac{w_k}{w_j}=\dfrac{w_i}{w_k}=b_{ij}\end{cases} \tag{5}$$

According to the judgment matrix B obtained above, the weight value $\{w_i=i=1\sim n\}$ of the decision index of the parking decision is obtained. If the judgment matrix satisfies Equation 4, the parking decision maker can calculate w_iw_k , that is, $b_{ik}=w_iw_k$, If the B judgment matrix has complete consistency, then

$$\sum_{i}^{n}\sum_{k=1}^{n}|\,b_{ik}w_k-w_i\,|=0 \tag{5}$$

In the formula: | | is the absolute value. As the parking decision model is extremely complex, people like a variety of preferences, people's understanding of things in the subjective with one-sidedness and instability, to determine the consistency of matrix B The problem is in the application of the actual parking decision model is not fully satisfied, the problem is objective. If the judgment matrix B does not have satisfactory consistency, it should be some correction of the matrix. We suppose B correction judgment matrix is $Y=\{y_{ij}\}_{n\times m}$, the weight value of each element of the Y correction judgment matrix is $\{w_i=i=1\sim n\}$, then it can be considered that the optimal consistency judgment matrix of B is called the smallest Y matrix:

$$\begin{cases}y_{ii}=1\\ \dfrac{1}{y_{ki}}=y_{ik}\in|\,b_{ik}-db_{ik},b_{ik}+db_{ik}\,|\\ \sum_{i=1}^{n}w_i=1\end{cases} \tag{6}$$

$$\min CIC(n) = \sum_{i}^{n}\sum_{k=1}^{n} | y_{ik} - b_{ik} | / n^2 \quad (7)$$

In the formula: $\mathrm{CIC}(n)$ that the function of the consistency of the impact of parking decision factor index coefficient, d is a non-negative parameters, if the ordinary method is more difficult to deal with, according to experience $\mathrm{CIC}(n)$ selection range [0,0.25]. The other symbols are the same as before, where the upper triangular elements of $n(n+1)/2$ are the optimization variables for the correction judgment matrix $Y = \{y_{ij}\}_{n\times m}$. It is obvious that the higher the consistency of the matrix (6), that is, when $\mathrm{CIC}(n) = 0$, the correction judgment matrix Y = B, then the B judgment matrix has complete consistency, and in accordance with the constraint condition of (7) $\mathrm{CIC}(n) = 0$ is the only global minimum. Genetic algorithm for simulating biology chromosome exchange and mutation in two operating mechanisms and survival of the fittest is a method with global search ability, using genetic algorithm to solve (6) Is more effective and simple. According to experience, it is generally considered that the judgment matrix has satisfactory consistency when the judgment matrix is consistent with the influence of the parking decision factor index coefficient $\mathrm{CIC}(n) < 0$, the author thinks that the calculated The impact of parking decision factors indicators w_i weight value is more reasonable, or need to increased, until there is a satisfactory consistency so far.

(3). Niche Genetic Algorithm

Niche technologies and genetic algorithms have their own advantages and disadvantages. The combination of the two can complement each other, can effectively overcome the premature, improve the efficiency of convergence, and get a more satisfactory course schedule[5].

(4). Course Arrangement Elements and Constraints

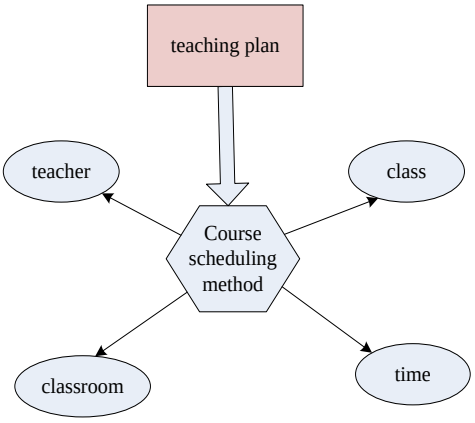

Figure 1. Logical diagram of arranging class problems

The problem of arranging classes is a multi-distribution problem, and its logical model is shown in Figure 1. Class schedules must meet the following hard constraints.

(1) Each teacher can only arrange one course at the same time;

(2) Each class can only be safe for a period of time.

(3) The type and capacity of classrooms used in each course must satisfy the requirements of the course; (4) Courses in each class can only be scheduled once and cannot be repeated; the courses should also satisfy the following soft constraints as much as possible .

(1) The course is arranged in the morning as far as possible, followed by the afternoon, and the second is the evening. (2) It is best for a teacher to take only one course a day; (3) It is best to take courses every other day in the course of the week.

3. SYSTEM DESIGN

(1). System design ideas

Choosing the operator: According to the number of choices expected by the individual in the modern age, choose the children and grandchildren who produce the next generation. Individual fathers with higher fitness values will generate one or more descendants in the next generation. During the selection process, the fitness value is the deciding factor for the individual to be selected or eliminated.

In order to improve the accuracy of the genetic algorithm, this article will automatically copy the individuals with high fitness in this generation to the next generation in the selection of the next generation group, and then perform hybridization on the individuals in this generation. The size of the crossover probability determines the frequency of crossover operations. The higher the frequency, the faster it can converge to the most promising optimal solution area. Therefore, generally choose a larger crossover probability, but too high crossover frequency may also lead to premature convergence, and can not perform a good overall search. The crossover probability is generally between (0.7, 0.9), and the crossover probability selected in this paper is 0.9.

The design of the system takes full advantage of the software engineering method, completes the design and development of the system through such steps as feasibility study, requirement analysis, outline design, detailed design code testing, and delivery and use. The overall workflow of the system is shown in Figure 2.

Academic management personnel enter basic information. The system administrator set the genetic algorithm scheduling parameters (multiple tests concluded) The teacher selected the course taught in this semester and entered the class conditions. The system automatically schedules classes. The hard constraint strategy is first met during the course scheduling process. Then meet the soft constraint strategy. Class schedules are generated at the end of the syllabus. The academic staff inquiries the required timetable according to different conditions.

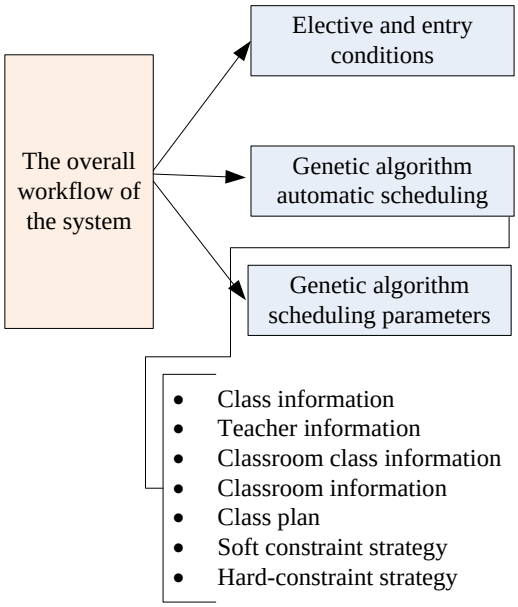

Figure 2. The overall workflow of the system

(2). The overall design of the system

The overall design function diagram of the system is shown in Figure 3.

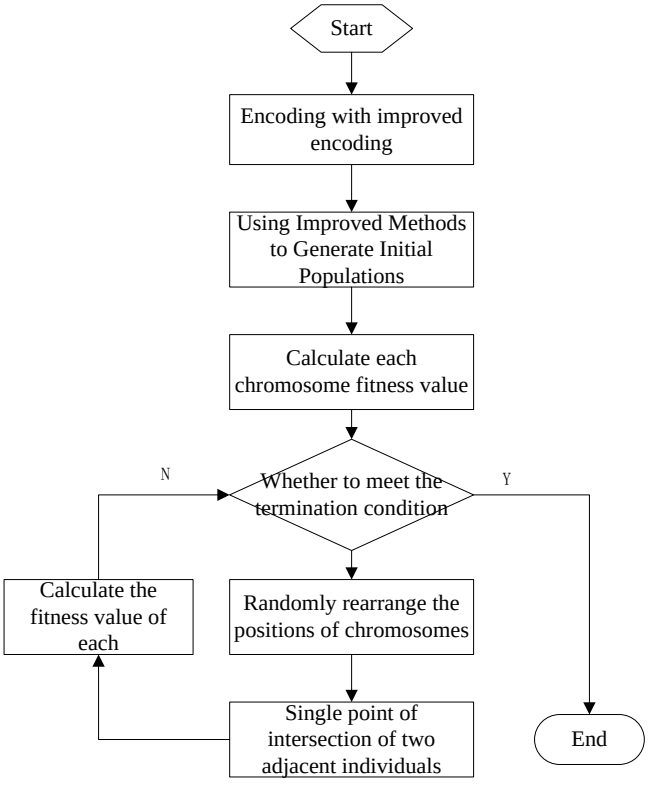

Figure 3. Teaching task schedule flow chart

First, the class to be scheduled is selected in batches, and a genetic algorithm is used to arrange the class for the selected class. The result is the optimal solution. Need to convert the results into a timetable. Convenient for late queries and other operations. The automatic arranging course of the genetic algorithm is shown in Figure 4, and the calling relationship of the class in the automatic arranging process code is shown in Figure 3. After the automatic class scheduling operation, the class information of the optimal solution is not obtained, and the system automatically derives the data, which needs to be manually adjusted .

(3). System test

The hardware environment of the system is: memory 6G, i7CPU, operating system: window7 64-bit system. There are 25 classrooms, 20 classes, 7 classes per class, 49 courses, 20 teachers as data, and special time constraints for individual teachers. And establish a class scheduling strategy. The test results are shown in Table 1 when the number of generations lost, the crossover rate, and the mutation rate are different.

All seven tests passed, and it can be seen that the fewer the number of generations lost, the faster the operation, the lower the crossover rate and mutation rate in the same environment and generation loss, the faster the operation speed, but the less the solution is obtained. According to the above test, the current loss was 50, the crossover rate was 0.2, and the mutation rate was 0.05, the most solution was obtained, and better results were obtained.

However, according to the actual operating environment, higher computers can be configured to increase the number of generations of losses. The more classes you choose to schedule, the longer it will take. It is recommended to select a certain number of classes for each class. Through the analysis of the data of 5 classes, it is found that the discharged curriculum has some details. The single-bi-weekly curriculum is not well-organized. For courses that require special classrooms, it is not ideal. When arranging the curriculum, special classroom courses are required. The entry of basic information is more troublesome and cannot be imported at one time.

Table 1. Generation Loss, Crossover Rate, Mutation Rate And Arrangement Effect Of Genetic Algorithm

No.	Loss of generation	Crossover rate	Mutation rate	Completion	Time cost
1	500	0.3	0.05	Yes	33.875s
2	300	0.3	0.05	Yes	15.338s
3	50	0.3	0.05	Yes	5.310 s
4	300	0.4	0.1	Yes	33.831s
5	300	0.3	0.03	Yes	34.470s
6	300	0.1	0.08	Yes	7.110 s
7	300	0.15	0.08	Yes	30.719s

4. CONCLUSION

In this paper, the use of genetic algorithm to solve course scheduling problems has the advantages of simple and practical, fast convergence, uniform distribution of time period, etc. The introduction of niche technology can effectively inhibit the shortcomings of genetic algorithms such as easy-to-premature and difficult to converge to the optimal solution. At the same time, the teaching task The arrangement of time slots and teaching venues are carried out in a step-by-step manner, avoiding the problem of too many factors to be considered by a single algorithm, as well as manual adjustments in the later stages, so as to ensure an optimal or at least feasible arrangement and not complete deadlock. This will make the arranging system more rationalized. The model and solution method have been applied in practice and have achieved good results. In addition, the combination of genetic algorithm and other algorithms to solve the scheduling problem is one of the hot topics. Designing a high-quality, versatile, and automated arranging system requires further research.

5. ACKNOWLEDGEMENTS

The authors acknowledge the Key scientific research projects of Henan Province: research of the Schedule arrangement problem on niche genetic algorithm (number：17A110038)

References

[1]Zhang J, Yang Q. An Improved Genetic Algorithm for Solving Multiple Sequence Alignment Problem Based on Niche Technology[J]. Journal of Changzhou Institute of Technology, 2016.

[2]Lo H Y, Liu G Z, Wang T I. Learning how to write effectively for academic journals: A case study investigating the design and development of a genre-based writing tutorial system[J]. Computers & Education, 2014, 78(259):250-267.

[3]Chen D, Zhou N, Cui T, et al. Control Strategy of Energy Storage for Frequency Coordination Dispatch Based on Improved Niche Genetic Algorithm[J]. 2017.

[4]Jin M S. Design and Realization of Laboratory Arranging Course System Based on Genetic Algorithm[J]. Research & Exploration in Laboratory, 2010.

[5]Wang F, Yu J L, Zhu P C, et al. RBF Network Based on Improved Niche Hybrid Hierarchy Genetic Algorithm[J]. Advanced Materials Research, 2012, 433-440:775-780.