

An Intelligent Course Scheduling System of Military Academy Based on Improved Genetic Algorithm

Yanming Yang^{*}, Weituan Wu and Yue Teng

Naval Aeronautical and Astronautical University Qingdao Campus, Qingdao 266041, China

^{*}Corresponding author

Abstract—Timetabling problem is a multi-objective combination optimization problem with constraints, and also has been proved to be a NP (Non-deterministic Polynomial) problem. The genetic algorithm is a highly parallel, random and adaptive global searching algorithm which is derived from the theory of natural selection and natural genetic mechanism, and it can effectively solve NP problem. According to the formulate principles and characteristics of actual schedule in military academy, the course scheduling system is designed and implemented based on the improved genetic algorithm. First, the characteristics of course arrangement in military academy is analyzed and the corresponding mathematical optimization model is established; then, the course scheduling system is designed and implemented based on the improved genetic algorithm, the main improvement of genetic algorithm includes: three-dimensional code scheme, optimal preservation strategy, self-adaptive crossover probability and mutation probability design schemes; finally, the course scheduling system is tested and analyzed. With the actual arrangement data from a military academy, the course scheduling system is tested and the irregular aperiodic process is realized; through the test analysis of the efficiency of the improved genetic algorithm, the feasibility and effectiveness of the improved genetic algorithm is verified.

Keywords—timetabling problem; genetic algorithm; three-dimensional code; optimal storage strategy

I. INTRODUCTION

Course arrangement is one of the most important and complex management tasks in university educational administration. It is the key to implement the teaching plan and is the important guarantee for the normal order of school teaching. Course scheduling problem is the timetabling problem (TTP). The main goal is to arrange the teachers, classes, courses and classroom according to the school teaching plan in a certain period of time which does not conflict in a semester [1]. In 1963 Professor Gotlieb for the first time raised the mathematical model of scheduling problem [2]. In 1976, S. Even et al. proved that the scheduling problem was NP-complete, and declared the academic position and difficulty of the "combinatorial explosion" problem of time and space [3].

In Colleges and universities, the main solution is the regular periodic timetabling problem, namely the timetable for a week time unit (week schedule), and in the military academies, to solve the non-periodic irregular timetabling problem, namely the schedule to date for the unit of time (the

daily schedule). Therefore, the research and development has a strong advanced, practical and intelligent, and in line with the characteristics of military colleges and universities in the course scheduling system is imperative.

Because of the characteristics of intelligence, parallelism and robustness in solving multi-objective combinatorial optimization problems, this paper chooses genetic algorithm to solve the problem of course arranging.

II. MATHEMATICAL MODELING OF TIMETABLING PROBLEM

A. Influence Factors of Course Arrangement

Course scheduling problems are mainly related to the five factors, such as classes, courses, teachers, time and classroom. These factors are not exist in isolation, they are interrelated and restrict each other.

B. The Constraint Conditions of Course Arrangement

To obtain the optimal solution of the combinatorial optimization problem, we must set reasonable constraints [4]. After research and analysis, constraints can be divided into hard constraints and soft constraints of these two categories [5]. Among them, the hard constraint conditions is a measure of the feasibility of the curriculum plan is the standard, the soft constraint is a measure of the criteria for scheduling the pros and cons [6]. Based on the research and analysis of the characteristics of military academy curriculum, the constraint conditions that should be taken into account in the process of course arrangement are as follows.

Hard constraint conditions:

- H1: At the same time, a class can not have more than one course.
- H2: At the same time, a teacher can not have more than one course.
- H3: At the same time, a classroom can not have more than one course.
- H4: The allocated classroom should accommodate more students than the number of classes.
- H5: Classroom types should match the course requirements.

Soft constraint conditions:

- S1: The most important courses are arranged in the most efficient way for students to learn.
- S2: In the same timetable, the distribution of the course should be basically uniform.
- S3: According to the number of students assigned to the classroom, maximize the use of resources.
- S4: Every morning the first and second classes to avoid as much as possible empty.
- S5: As far as possible to meet some of the teacher's class time and place reasonable requirements.
- S6: Theoretical and practical courses as far as possible alternating arrangements, in order to facilitate students to improve learning efficiency.
- S7: When the same teacher holds a number of courses, the course schedule should be kept at a reasonable interval so that the teacher can prepare lessons.
- S8: Some courses have a cohesive relationship between the courses, preferably when one of the courses started, another course to start.
- S9: Experimental, practical exercises, training and physical education classes are best arranged in the afternoon.

C. Mathematical Description

Suppose the school has M classes, N teachers, H courses, P classrooms, Q time period, the model described in detail as follows:

- Class set is $C = \{c_1, c_2, \dots, c_m, \dots, c_M\}$, the number of classes are $\{x_1, x_2, \dots, x_m, \dots, x_M\}$.
- The teacher set is $G = \{g_1, g_2, \dots, g_n, \dots, g_N\}$. Each teacher corresponds to one or more courses, the corresponding number of teachers is $\{y_1, y_2, \dots, y_n, \dots, y_N\}$.
- The course set is $K = \{k_1, k_2, \dots, k_h, \dots, k_H\}$. Each course corresponds to one or more classes, one or more instructors. The number of classes for each course is $\{z_1, z_2, \dots, z_h, \dots, z_H\}$.
- The set of classrooms is $R = \{r_1, r_2, \dots, r_p, \dots, r_P\}$. The number of people in each classroom is $\{\omega_1, \omega_2, \dots, \omega_p, \dots, \omega_P\}$.
- The time set is $T = \{t_1, t_2, \dots, t_q, \dots, t_Q\}$. The Cartesian product of "time-classroom pair" is shown in equation (1):

$$S = T \cdot R = \{(t_1, r_1), (t_1, r_2), \dots, (t_1, r_p), (t_2, r_1), (t_2, r_2), \dots, (t_2, r_p), \dots, (t_Q, r_1), (t_Q, r_2), \dots, (t_Q, r_p)\} \quad (1)$$

Classes, courses, and teachers can be bundled together to form a teaching event (course tuple), since the appropriate course and teacher has already been assigned to each class in the start-up plan. The scheduling problem can be seen as finding a suitable "time-classroom pair" for each course.

III. THE DESIGN OF GENETIC ALGORITHM IN COURSE SCHEDULING SYSTEM

A. Three-Dimensional Coding Scheme

In the genetic algorithm, the encoding method directly affects the complexity of the algorithm and the efficiency of the implementation [7]. This paper adopts three-dimensional coding scheme, which is shown in figure 1, in order to balance the various factors of scheduling problem.

In Figure 1, the X axis represents the time axis, and each time coordinate on the X axis corresponds to a time period. The Y-axis represents the classroom axis, and each coordinate interval on the Y-axis represents a classroom. There are P coordinate values (Y1~YP). Each coordinate interval on the Z-axis represents a teaching event (course tuple). In this way, three-dimensional coordinates can uniquely identify a small cube, known as the individual's gene block, also completed a class a course once the arrangement. When all of the chromosomes have been filled, it shows that a schedule has been created.

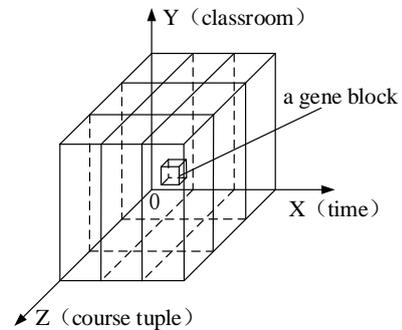


FIGURE I. THREE-DIMENSIONAL CODING OF CHROMOSOME.

B. Construct Feasible Initial Population

According to the three-dimensional coding scheme, the initial population is generated as follows:

- Step1: first build an empty three-dimensional array, and expressed as a form of chromosome, the array of all the elements initialized to 0.
- Step2: Determine the number of times (M1) that the first one in a teaching event lists is taught in a semester, according to the size of M1 in the first layer of the three-dimensional array of chromosomes (the layer perpendicular to the Z axis) M1 position is randomly selected and assigned a value of 1.

- Step3: Continue to the next one in the teaching event lists until all the teaching events have been processed;
- Step4: Continue to build the next three-dimensional array of chromosomes, and repeat Step2 and Step3, until the number of three-dimensional array of chromosomes to reach the pre-set size of the population so far.

C. Construct the Fitness Function

According to the above course scheduling system soft constraints analysis of military academies, construct the fitness function as shown in equation (2) :

$$F = \sum_{i=1}^4 (\theta_i \times f_i) \quad (1)$$

It is according to the target value weighted, where the value of q_i ($i = 1, 2, 3, 4$) can be defined by the management of their own, it represents the weight of course targets, respectively, the value is 4,1,2,3.

D. The adaptive Crossover Probability and Mutation Probability

In the genetic algorithm parameters, the crossover probability and mutation probability selection is very important, they affect the performance of the algorithm. The specific design idea of this paper is to prevent the optimal individuals from crossing and mutating in the contemporary population, but to make more optimal individuals cross and mutation operation. The crossover probability and mutation probability are shown in equation (3) and (4) :

$$p_c = \begin{cases} a_1 \sin\left(\frac{\pi}{2} \times \frac{f_{\max} - f'_c}{f_{\max} - f_{avg}}\right), & f'_c \geq f_{avg} \\ a_2, & f'_c < f_{avg} \end{cases} \quad (3)$$

$$p_m = \begin{cases} a_3 \sin\left(\frac{\pi}{2} \times \frac{f_{\max} - f_m}{f_{\max} - f_{avg}}\right), & f_m \geq f_{avg} \\ a_4, & f_m < f_{avg} \end{cases} \quad (4)$$

Where a_1, a_2, a_3, a_4 are constants in the interval (0,1), in general, let $a_1=a_2, a_3=a_4$, the specific value depends on the actual situation; f_{\max} is the fitness value of the best individual in the contemporary population; f_{avg} is the fitness value of the two cross individuals, f_m is the fitness value of the individual.

E. The Optimal Storage Strategy

Roulette selection method is simple, but in the latter part of the algorithm may appear premature convergence or stagnation of the phenomenon. Therefore, this paper adopts the optimal storage strategy in the roulette selection method, the specific implementation process is as follows:

Suppose the population size is M , the parent population is $F = \{x_1, x_2, \dots, x_i, \dots, x_M\}$, the initial state of the offspring population is $S = \{\}$, and $f(x_i)$ is the fitness value of the individual x_i .

- Step1: calculate the adaptive degree value of each individual in a population F , and find the adaptive degree value of individual x_k , that is $f(x_k) = \max(f(x_1), f(x_2), \dots, f(x_M))$, and according to the size of the adaptive degree value to arrange order of all individuals in a population, get $F_1 = \{y_1, y_2, \dots, y_i, \dots, y_M\}$;
- Step2: calculate the sum of all individuals in the population adaptive degree $\sum_{i=1}^M f(y_i)$;
- Step 3: calculate the probability of y_k is selected individuals in the population $P_{y_k} = \frac{f(y_k)}{\sum_{i=1}^M f(y_i)}$, $k = 1, 2, \dots, M$;
- Step4: calculation of individuals in a population is selected in the roulette wheel, the cumulative probability $PP_k = \sum_{j=1}^k P_{y_j}$, $k = 1, 2, \dots, M$;
- Step5: rotation M time table;
- Step6: in the population S to find the lowest adaptive degree value of individual z_t , expressed as $f(z_t) = \max(f(z_1), f(z_2), \dots, f(z_T))$;
- Step7: the highest adaptive degree value of individual x_k in the population F instead of the individual z_t in the lowest fitness value in the population S ;
- Step8: all individuals will be selected to save, and return.

IV. TESTING OF THE COURSE SCHEDULING SYSTEM

A. Parameter Testing

Test groups 1 to 3 were tested for population sizes of 30, 50 and 100. In each group of tests were randomly selected 10 times the results of operations to record. The time of the approximate optimal solution obtained by the test can be used as the standard to measure the efficiency of the algorithm. The results of the three groups are shown in Table 1.

TABLE I. EFFECT OF POPULATION SIZE ON ALGORITHM EFFICIENCY (UNIT: SECOND)

Number of Tests	Test Group 1 Population Size 30		Test Group 2 Population Size 50		Test Group 3 Population Size 100	
	Time	Fitness value of near-optimal solution	Time	Fitness value of near-optimal solution	Time	Fitness value of near-optimal solution
1	271.6	1549	315.4	1610	721.6	1702
2	257.9	1491	334.2	1599	733.1	1598
3	302.8	1583	371.5	1583	763.0	1662
4	284.2	1568	352.1	1627	752.8	1678
5	265.3	1531	326.7	1593	708.4	1582
6	276.0	1596	340.1	1614	736.1	1636
7	263.8	1574	338.6	1589	751.4	1699
8	281.1	1469	357.4	1623	732.7	1684
9	271.4	1551	308.2	1634	741.4	1622
10	259.3	1528	319.3	1576	709.1	1649
average	273.3	1544	336.3	1604.8	734.1	1651.2

It can be seen from Table 2 that the larger the population size, the longer the running time of the algorithm, and the higher the fitness value of the near-optimal solution. But the population size can not be increased unrestricted, otherwise it will lead to a sharp increase in algorithm running time, efficiency greatly reduced. The system is proved to be efficient when the population size is between 50 and 100 by several tests. Therefore, the population size is set to be 100.

B. Efficiency Testing

In this paper, the selection operation is to combine the optimal preservation strategy in the classic roulette selection method, and compare with the running efficiency of the references [8] is not integrated into the optimal preservation strategy. The results of 10 runs are randomly selected and the results are shown in figure 4. According to figure 2, the improved genetic algorithm is more efficient than the reference [8].

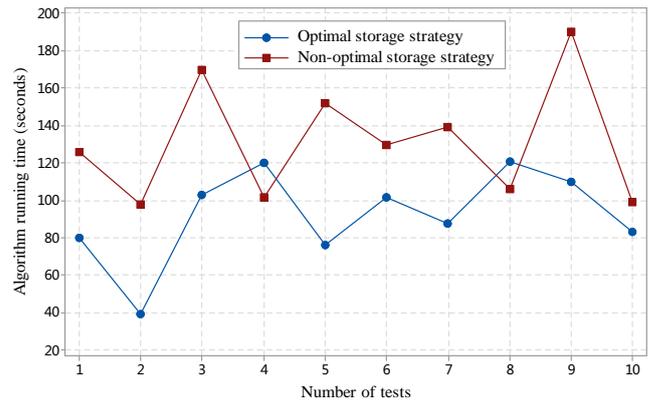


FIGURE II. COMPARISON OF ALGORITHMS WITHOUT INTEGRATION OF OPTIMAL STORAGE STRATEGY AND OPTIMAL STORAGE STRATEGY.

CONCLUSIONS

According to the characteristics of course scheduling system in military academy, this paper analyzes the influence factors and restraint conditions of course scheduling, and establishes the corresponding mathematical optimization model. And designs and realizes the military course arranging system with the improved genetic algorithm. The test results show that the proposed improved genetic algorithm can effectively solve the problem of arranging course in military academies. According to the result of the course arrangement, the course distribution is more uniform, and can realize the non-periodic irregular scheduling problem.

REFERENCES

- [1] Pillay N, Banzhaf W. A study of heuristic combinations for hyper-heuristic systems for the incapacitated examination timetabling problem[J]. European Journal of Operational Research, 2009, 197(2): 482-491.
- [2] Gotlieb C. The construction of class-teacher timetables[C]. Proc. IFIP Congress, 1963: 73-74.
- [3] Even S, Itai A, Shamir A. On the complexity of timetable and multi-commodity flow problems[J]. SIAM Journal on Computing. 1976(5): 691-703.
- [4] Abuhamdah A, Ayob M. Multi-neighborhood particle collision algorithm for solving course timetabling problems[C]. Proceedings of the 2nd Conference on Data Mining and Optimization, 2009: 21-27.
- [5] LI Hongchan, HU Gang, ZHU Haodong. Research of UTP based on population dominant GA[J]. Computer Engineering and Applications, 2011, 47(10): 233-236.
- [6] SUN Tong, GUO Qian-qian. Simulation on UTP based on new immune GA[J]. Computer Simulation, 2012, 29(2):386-391.
- [7] LI Ming, WANG Yan, NIAN Fuzhong. Intelligent Information Processing and Application[M]. Beijing: Electronic Industry Press, 2010: 106-107.
- [8] WANG Nuo. Optimization of Course Arrangement System Based on Genetic Algorithm [D]. Hebei University of Science and Technology, 2011.