

A New Algorithm of Course-scheduling Problem for the New College Entrance Examinations based on Scheduling Model

Huijian XU*

Zhejiang Institute of Economics and Trade, Hangzhou 310018, China

*Corresponding Author

Abstract:

In order to meet the flexible and diverse needs of intelligent Course Scheduling in the new college entrance examination teaching mode, this paper proposes a new college entrance examination course scheduling algorithm based on scheduling model. In the scheduling algorithm, constraints are divided into two categories by constraint programming, which are event-related constraints and resource-related constraints. Genetic algorithm based on scheduling model is used to make all hard constraints satisfied, while the sum of soft constraint conflicts is minimized, and then the optimal solution of the scheduling problem is obtained. The experimental results show that the algorithm can significantly reduce the scheduling time than the standard genetic algorithm, and has better versatility.

Keywords: Mobile teaching system, Intelligent course scheduling, Hard constraints, Soft constraints, Genetic algorithm.

I. INTRODUCTION

Since the State Council issued the document titled Opinions on Deepening the Reform of the Examination and Enrollment System, the new college entrance examination policy has been implemented in provinces like Shanghai, Zhejiang, Beijing, Tianjin, Shandong, Hainan and other provinces in succession, and the pilot reform of the examination and enrollment system has also been comprehensively launched. The new college entrance examination policy highlights the mechanism of classified examination, comprehensive evaluation and multiple admission, and it effectively increases students' independent selection of the subjects for the college entrance examination, thus dispersing students' examination pressure to a great extent [1, 2]. According to the principle of “compulsory examination” and “elective examination”, the new college entrance examination system has basically formed two modes: “3+3” and “3+1+2”. Taking Zhejiang Province as an example, the new college entrance examination has implemented the “3+3” mode, that is, Chinese, mathematics and the foreign

language are the exam-compulsory subjects, and the selected subjects are the three subjects arbitrarily selected from Politics, History, Geography, Physics, Chemistry, Biology and technology. In this way, students can have 35 combinations in the college entrance examination. In order to better adapt to the new college entrance examination system and promote the better personalized development of students, a flexible optional class system appears for students to choose classes.

The optional class system of the new college entrance examination breaks the traditional mode of natural teaching classes. It has high flexibility and can meet students' personalized education needs. Students can flexibly choose learning based on their own specialties and interests [3, 4]. However, the system also exposes many problems in practice while it brings flexibility and convenience to teaching. Firstly, the looseness and flexibility of the optional class system pose great challenges to management. Secondly, the intersection between the administrative class and the teaching class makes the course scheduling more complicated. The utilization of teaching resources becomes more demanding, and the constraints on the course scheduling also grow. Therefore, it is urgent for the school to establish an instructional management system based on “3+3” or “3+1+2” of the new college entrance examination. With the help of the system, students can better adapt themselves to the reform of the new college entrance examination, get rid of the constraints of the administrative classes and flexibly choose the suitable courses. Meanwhile, it further solves a series of problems in students' course selection, the optional class system, the course scheduling, management, etc. in the new college entrance examination.

II. BASIC THOUGHTS OF COURSE SCHEDULING IN THE OPTIONAL CLASS SYSTEM

The core element of the instructional management system is to meet the flexible and diverse demands of the course scheduling in the optional class system. In essence, it is a problem of the typical multi-class resource combination optimization. Regarding the solution of such problems, only reasonable and satisfactory solutions can usually be found rather than the real optimal solution [5, 6].

The course scheduling of the optional class system focuses on the overall combination and coordination of time, space, students and teachers, and gives priority to the connotations of the following three aspects. Firstly, the hard constraint conditions are met. That is, students and teachers should avoid conflicts in the utilization of resources. For example, a class can only correspond to one course, one teacher and one classroom at a time. Secondly, the soft constraint conditions are met. That is, the course time should be optimized to achieve the optimal solution. For example, under the “3+3” mode of the new college entrance examination, the courses can be scheduled according to the respective focus of the compulsory and optional subjects of examination so as to ensure that the weekly teaching hours and teachers' weekly class hours are reasonably and evenly distributed. Third, the special needs of some courses are met. For example, technical courses need to be taught in computer rooms or multimedia classrooms. Therefore, when it comes to the course scheduling of the optional class system, it is

necessary to comprehensively consider the time and space requirements of the course scheduling activities, namely, the hard constraint conditions. It is also necessary to consider whether some soft constraint conditions meet the teaching rules. After that, the two are effectively combined. The research on Constraint Satisfaction Problem (CSP) is a branch in the field of artificial intelligence [7, 8]. Generally, a CSP problem can be defined as the triple group $P = (V, D, C)$, where V is the set of n variables ($V_1, V_2, \dots, V_i, \dots, V_n$); D is the set of n fields ($D_1, D_2, \dots, D_i, \dots, D_n$). D_i is the set of possible values of V_i . C is the set of constraint relations between variables V [9].

III. COURSE SCHEDULING ALGORITHM DESIGN OF THE OPTIONAL CLASS SYSTEM

Internationally, the high school curriculum schedule (HSTT) is a well-known problem. It consists of coordinating resources (e.g., teachers, classrooms), time and events (e.g., classes), and various constraints [10]. This paper studies the applicability of constraint programming in high school curriculum, and a new model for HSTT is developed by using a scheduling-based view. The course schedule quality is an important problem in the course scheduling algorithm design for the optional class system, because it directly affects the education system, the satisfaction of students and the teaching staff, and other matters. It takes at least one semester for each schedule to influence hundreds of students and teachers over a long period of time, which makes HSTT an extremely significant and responsible task. However, manual scheduling can be time-consuming, difficult, error-prone, and in some cases practically impossible. Therefore, it is critical to develop algorithms that automatically generate the best course schedule.

3.1 Constraint Conditions

A typical high school curriculum stipulates three main factors: time, resources and events [11]. Time refers to the discrete time units available, such as the first lesson and the second lesson on Monday. Resources correspond to the rooms, teachers, students, and others available. The main entity is the event, and its occurrence takes time and resources. An event can be a math class, which requires a math teacher (to be identified) and a specific group of students (both the teacher and the student group are considered resources) twice. Events will be scheduled into one or multiple solution events or subevents.

The purpose of the algorithm is to find a schedule by allocating time and resources to events, which makes all the hard constraints satisfied and minimizes the sum of the soft constraint conflicts [12, 13]. The currently given constraints for the high school curriculum problem are divided into event-related constraints and resource-related constraints.

3.1.1 Event-related constraints

(1) Time allocation --- A specified amount of time is allocated to a specified event.

(2) Priority time --- When time is assigned to events, the specified time is of high priority.

(3) related events---events that simultaneously happen.

(4) Spread events --- The designated events must continue throughout the week, e.g., Chinese and Math classes must be assigned the class hours every day.

(5) Split event allocation ---limit the number of subevents for a given event that may require a specific duration.

(6) Split events---limit the minimum and maximum duration of the given event as well as the number of the subevents. When it is combined with the distributed split events, the subevents can be better controlled.

(7) Ordering of events---The specified event must be arranged one after another with a non-negative time lag.

(8) Avoid the split allocation--- For all the subevents derived from an event, the same resources are allocated.

3.1.2 Resource-related constraints

(1) Resource allocation——The specified resources are allocated to the specified event.

(2) Avoid conflicts——Two or multiple subevents cannot use the specified resources simultaneously.

(3) Priority resources——When the resource is assigned to an event, the specified resource has priority over other resources.

(4) Avoid the unavailable time——The specified resources cannot be used in the specified time.

(5) Limit the load——The workload of the specified resource must be within the given values.

(6) Limit the busy hour——The use time of the resources should be within the given values in a specified time group.

(7) Cluster busy hour——The designated resource activities must all take place in the minimum and maximum time group.

(8) Limit the idle time——The idle time of the designated resources must be within the given values in the specified time group.

3.2 Improved Genetic Algorithm Design based on the Scheduling Model

In the course scheduling algorithm design, the key elements are a group of events E , a group of resources R and a group of time T , which are regarded as the integer $T = \{0, 1, 2, \dots | T | - 1\}$. Given the constrained forms of the problem, the resources used by each event are predefined. Regarding the early model of the general senior high school schedule problem, each pair of events and each time slot are explicitly manifested, which indicates whether the event occurred at that particular time. With the modeling method based on scheduling, each subevent is associated with two variables: the start time variable and the duration variable. Therefore, it can take advantage of both the discrete and the conventional global constraints. That is, the decision variables are described first and then each constraint is modeled.

Each event $e \in E$ has the maximum total duration $D(e)$. For each event e , $D(e)$ sub-events are created, and they are numbered from 0 to $d(e)-1$. Each subevent is associated with two variables that indicate its start time and duration. They are labeled as the start (e, i) and the end (e, i). The special start time $UN = |T|$ is used to represent the unused events. In this case, the corresponding duration is zero. Constraints may impose limits on the number and duration of subevents.

$EV(r) \subseteq E$ shows the event set of the required resources r . The auxiliary Boolean variable busy (r, t) is introduced to show the event $e \in EV(r)$ of the required resource $r \in R$ happens at the time t . The time group $TG \subseteq T$ is a group of fixed time, and the event group $EG \subseteq R$ is a group of fixed events.

$$\begin{aligned} within(x, l, u) &= (x \geq l \wedge start(e, i) + dur(e, i) > t, x \leq u) \\ busy(r, t) &\Leftrightarrow \exists_{e \in EV(r), i \in \{0 \dots D(e)-1\}} start(e, i) \leq t \wedge start(e, i) + dur(e, i) > t \end{aligned} \quad (1)$$

(1) Time allocation:

The event must be allocated the specified time.

$$\sum_{i \in \{0,1,\dots, D(e)-1\}} dur(e,i) = D(e) \quad (2)$$

(2) Priority time

The subevent of the specified event e can only start within the specified time group T_{Ge} . If the optional parameter D is given, the constraint applies only to the subevents during the particular duration.

$$\forall i \in i \in \{0,1,\dots, D(e)-1\} \text{ start}(e,i) \neq UN \Rightarrow \text{start}(e,i) \in TG_e \quad (3)$$

(3) Related events

Some events must happen simultaneously. If EG is an event group composed of linkage events, all the events will have the same total duration TD , namely, $\forall e \in EG, D(e) = TD$. The global constraint “all equal” is used. It is required that the input variable must be allocated the same value.

$$\begin{aligned} & \forall i \in i \in \{0,1,\dots, TD-1\} \\ & \text{all_equal}([\text{start}(e,i) | e \in EG]) \\ & \text{all_equal}([dur(e,i) | e \in EG]) \end{aligned} \quad (4)$$

(4) Spread events

The starting frequency of the events in the specified event group is limited in the given time group. The event and the time group are the sets of events and durations, respectively.

$$Z = \sum_{e \in EG, i \in \{0,1,\dots, D(e)-1\}} \text{within}(\text{start}(e,i), \min(TG), \max(TG)) \wedge \text{within}(z, \min e, \max e) \quad (5)$$

(5) Split event allocation

The number of subevents within the given duration is limited within the allowable scope.

$$a = \sum_{(i \in \{0,1,\dots, D(e)-1\})} (dur(e,i) = d) \wedge \text{within}(a, \min ds_e, \max ds_e) \quad (6)$$

(6) Split events

The number of subevents between different events are adjusted.

$$\begin{aligned} \forall i \in \{0, \dots, \min s_e - 1\} : start(e, i) \neq UN \wedge dur(e, i) \neq 0 \\ \forall i \in \{\max s_e, \dots, D(e) - 1\} : start(e, i) = UN \wedge dur(e, i) = 0 \\ dur(e, i) \leq \max d_e \wedge dur(e, i) \neq 0 \Rightarrow dur(e, i) \geq \min d_e \end{aligned} \quad (7)$$

(7) Ordering of events

Regarding the given pair of events (e_1 , e_2), it is required in the constraint conditions that e_1 must happen prior to e_2 . In addition, the minimum and maximum time interval unit must be available.

$$\begin{aligned} Oe = \min \{ start(e_2, i) \mid i \in \{0 \cdot D(e_2) - 1\} \} - \max \{ start(e_1, i) + dur(e_1, i) \mid i \in \{0 \cdot D(e_2) - 1\} \} \\ \wedge wint hin(eo, \min_{ep}, \max_{ep}) \end{aligned} \quad (8)$$

(8) Avoid conflicts

In any give time, a resource is only used by an event at most.

$$\begin{aligned} disjunctive([start(e, i) \mid e \in EV(r), i \in \{0 \cdot D(e) - 1\}], [dur(e, i) \mid e \in EV(r), \\ i \in \{0 \cdot D(e) - 1\}]) \end{aligned} \quad (9)$$

(9) Avoid the unavailable time

The resource cannot be used at the specified time. For each resource r and the forbidding time t , it is encoded by creating a pseudo-event that requires r and has a fixed duration of 1 at time t . Newly created events will be added to the event and will be considered in avoiding conflict constraints. For soft versions, the duration of these virtual events is 0..1. If the duration 0 is used, the constraints will be violated.

(10) Limit the busy hour

If a resource r is busy in a time group T G , the busy frequency in the time group will be limited between $\min br$ and $\max br$.

$$c = \sum_{t \in TG} (\text{busy}(r, t)) \wedge c \neq 0 \Rightarrow \text{within}(c, \min b_r, \max b_r) \quad (10)$$

(11) Cluster busy hour

The resource is busy in the time group if it is busy at least once in that time group. A time group TG is given by the constraint, and the total number of the time groups with the possible busy hours of the resource is limited between mint and maxt. For example, the teacher must finish the work within three days.

$$b = \sum_{TG \in TG} (\exists_{t \in TG} \text{busy}(r, t)) \wedge \text{within}(b, \min t, \max t) \quad (11)$$

IV. EXPERIMENTAL RESULTS AND ANALYSIS

4.1 Experimental Data

The test data used in this paper are based on the true desensitization data of the course scheduling in the three schools in Hangzhou in August 2020.

Experimental data 1: Grade 2 in the key senior high school includes 22 administrative classes and 79 teaching classes. Teaching classes and administrative classes are mixed. 40 class hours and 37 lessons are scheduled every week. In accordance with the technical optional class system, it is required that the lessons of communication technology and information technology are given in the same class but in different class hours.

Experimental data 2: In the foreign languages senior high school B, the courses are scheduled together in primary school, junior high school and senior high school. The optional class system is adopted for all the classes. Some foreign students have Chinese lessons in the classes of the junior high school, and they have other lessons in the classes of the senior high school. The individualization degree of students is rather high. 40 class hours and 39 lessons are scheduled every week.

Experimental data 3: In the key junior high school C, there are 37 classes of different grades in total. It has a rich curriculum. 37 lessons need to be scheduled in 40 class hours with little space for adjustment but rich requirements. The teaching and research groups, the lesson preparation groups, the grade groups, headteachers and students all have different needs. In addition, it is also obvious to meet the resource conflicts. If the elective classes are scheduled and given at the same time, there will be

obvious conflicts in terms of teachers and classroom resources.

4.2 Parameter Setting

Parameter setting are shown in Table I:

TABLE I. Specific parameter setting of the experiment

Parameter name	Parameter symbol	Parameter values and description	Specific parameter setting
population size	N	The values of N are between 100 and 150. If N value is too great, it will increase the computation cost. If N value is too small, the results are not ideal.	125
maximal evolutionary algebra	GEN	The values of GEN are between 300 and 500. If the value is too great, it will increase the computation cost. If the value is too small, the convergence fails or most of the value is still the default solution.	400
crossover probability	Pc	Pc value is between 0.8 and 0.9. A small Pc is bad for the convergence, while a big Pc causes the great fluctuations in the algorithm.	0.85
mutation probability	Pm	Pm value is between 0.01 and 0.05. A small Pm is bad for the solution of the global optimal, while a great Pm will causes the great fluctuations in the algorithm and even the convergence failure.	0.03
elitism retention parameter	r	The value of r is 3 or 4. Given that the course scheduling is in itself complicated, the great value will increase the computation cost, while the small value will make the selected pressure insufficient.	3

4.3 Improved Algorithm Comparison Experiment

The time standard is as follows: the runtime that satisfies 100% hard constraints and 80% soft constraints.

Based on experimental data 1, the scheduling programs of both the standard genetic algorithm and the improved genetic algorithm based on the scheduling model are used to compare the scheduling time, and the results are shown in Table II:

TABLE II. Results of Experimental Data 1

Experiment Number	Standard Genetic Algorithm (seconds)	Improved Genetic Algorithm Based on The Scheduling Model (seconds)
1	2123	1431
2	4312	1901
3	2136	1238
4	3212	2012
5	6512	2201
mean time	3659	1756.6

Based on experimental data 2, the scheduling programs of both the standard genetic algorithm and the improved genetic algorithm based on the scheduling model are used to compare the scheduling time, and the results are shown in Table III.

TABLE III. Results of Experimental Data 2

Experiment Number	Standard Genetic Algorithm (seconds)	Improved Genetic Algorithm Based on The Scheduling Model (seconds)
1	132	121
2	237	74
3	198	56
4	334	41
5	291	79
mean time	238.4	74.2

Based on experimental data 3, the scheduling programs of both the standard genetic algorithm and the improved genetic algorithm based on the scheduling model are used to compare the scheduling time, and the results are shown in Table IV.

TABLE IV. Results of Experimental Data 3

Experiment Number	Standard Genetic Algorithm (seconds)	Improved Genetic Algorithm Based on The Scheduling Model (seconds)
1	5772	4312
2	7892	2321
3	7263	4921
4	5461	3234
5	4813	2871
mean time	6240.2	3531.8

By comparing the runtime of the two algorithms, it can be seen that: the improved genetic algorithm based on the scheduling model takes less time than the standard genetic algorithm. This indicates that the improved algorithm proposed in this paper can indeed improve the running speed of the algorithm and shorten the time from the algorithm convergence to the optimal solution.

4.4 Algorithm Performance Evaluation

In order to better reflect the performance of the improved algorithm proposed in this paper, this article focuses on the course scheduling programs of the standard genetic algorithm and the improved genetic algorithm based on the scheduling model, and conducts the performance evaluation comparison of the course scheduling quality from the following five dimensions: teachers' satisfaction with the course schedule, the main subject satisfaction of the class, similarity in the main subject schedules of teachers, course distribution dispersion, and satisfaction of teachers' individualized needs. The results are shown in Table V.

TABLE V. Results of algorithm performance evaluation

experiment on average	standard genetic algorithm	improved genetic algorithm based on the scheduling model
teachers' satisfaction with the course schedule	73	87
the main subject satisfaction of the class	82	87
similarity in the main subject schedules of teachers	65	89

course distribution dispersion	91	90
satisfaction of teachers' individualized needs	81	87

By comparing the experimental results of the two algorithms above, it can be seen that: the improved genetic algorithm based on the scheduling model improves a lot not only in runtime but also in the evaluation indexes of the course scheduling.

V. CONCLUSIONS

This paper proposes an improved genetic algorithm based on the scheduling model to solve the course scheduling problem of the optional class system in the new college entrance examination mode. By setting constraints and making the schedule modeling, the algorithm solves the problems of the traditional standard genetic algorithm, including the time consumption and the relatively low similarity in the main subject schedules of teachers. Both the experiments and the practices have proved its effectiveness and universality in the course scheduling of senior high schools. As the course scheduling rules are the key factors that influence the results of the course scheduling algorithm, the parameters of the specific course scheduling rules will be further studied, optimized and upgraded according to the requirements and environments in different areas and schools. The study aims to establish a diversified model of the rule base and improve the universality of the algorithm.

ACKNOWLEDGEMENT

Phasic research results of the Higher education Teaching Reform Research Project of Zhejiang Province in 2018: "Research and Practice of Dynamic Management Mechanism in College Major Classification Construction from the Perspective of Diagnosis Reform" (JG20180661); Phasic research results of Zhejiang Natural Science Foundation project: "Mobile Personalized Recommendation Model and Incentive Mechanism Research based on Participatory Perception and Privacy Concern" (LY18G020012).

REFERENCES

- [1] Meng DF (2016). Action research on the implementation of student growth tutorial system in small-class education in senior high school. Yunnan University.
- [2] Yun JY (2017). The design and implementation of the educational administration course-scheduling system in colleges and universities based on the genetic algorithm. Hunan University.

- [3] Liao BJ, Yao GZ (2016). Enlightenments of the foreign optional class system on the teaching reform of senior high schools in China. *Journal of Lanzhou Institute of Education*, 32(7):112-113.
- [4] Tan HZ, Zhao L (2018). Cloud workflow scheduling method combining Backfilling with the idle resource scheduling. *Journal of Southwest University (natural science edition)* 40 (6): 149-157.
- [5] Xu JX (2017). A queuing algorithm based on multi-influence factors. *Engineering Technology*, (23):315.
- [6] Sui ZL (2018). Cloud-computing resource scheduling based on chaotic particle swarm and chicken flock fusion algorithm. *Journal of Changchun University* 28 (1): 6-11.
- [7] Zhang G, Chen A, Hu J (2019). Study on course scheduling based on Monte Carlo Genetic Algorithm. *Experimental Technology and Management* 36(3): 170-174.
- [8] Liao et al (2017). Adaptive optimization method of platform resource scheduling in uncertain environment based on cloud genetic algorithm. *Journal of Air Force Engineering University (natural science edition)* 18 (4): 86-92.
- [9] Feng et al (2018). A research review of parallel genetic algorithms. *Computer Application and Software*, 35(11): 1-7, 80.
- [10] Chen LR (2017). Calm thinking on the heated “optional class system” in the background of the new college entrance examination. *Basic Education and Research*, 2017, (9).
- [11] Wu, Zhao & Yuan (2019). Solution of the constraint satisfaction problem by combining confidence propagation with simulated annealing. *Applied Research of Computers* 36(05):1297-1301.
- [12] Wang HX (2019). Research and design of the course scheduling algorithm in the optional class system. Central China Normal University, Wuhan.
- [13] Jiang J, Bai SX (2018). Genetic algorithm improvements and its application in the course scheduling *Journal of Nanchang University (science edition)* 42(4): 388-392.