Article

School Timetabling Optimisation Using Artificial Bee Colony Algorithm Based on a Virtual Searching Space Method

Kaixiang Zhu[1*], Lily D Li [2*] and Michael Li [3, *]

[1] CQUniversity; k.zhu@cqu.edu.au
[2] CQUniversity; l.li@cqu.edu.au
[3] CQUniversity; m.li@cqu.edu.au
* Correspondences: KZ k.zhu@cqu.edu.au, Tel.: +061 0466 307 631; LL l.li@cqu.edu.au; ML m.li@cqu.edu.au

Abstract: Although educational timetabling problems have been studied for decades, one type, the STP, has not developed as quickly as the other two types due to its diversity and complexity. Also, most of the STP research has only focused on the educators' availabilities when studying the educator aspect, and the educators' preferences and expertise have not been taken into consideration. This paper proposes a conceptual model for the school timetabling problem considering educators' availabilities, preferences and expertise as a whole, and chooses a common real-world school timetabling scenario to study. A mathematical model is presented. A Virtual search space for dealing with the large search space is introduced, and the artificial bee colony algorithm is adapted and applied to the proposed model. The proposed approach is simulated with a random-generated large dataset. The experimental results demonstrate that the proposed approach is able to solve the STP and handle a large dataset in an ordinary computer hardware environment, which significantly reduces computational costs. Compared to the traditional CP method, the proposed approach is more effective and can provide more satisfactory solutions in considering educators' availabilities, preferences, and expertise levels.

Keywords: educational timetable; school timetabling; constraint satisfaction problem; optimisation; artificial bee colony algorithm

1.Introduction

The problem of timetabling can be defined as a computational problem that allocates resources in given periods under particular constraints to achieve desirable goals [1]. Educational timetabling is one of the fundamental tasks affecting educational institutes' operations and productions. Educational timetabling problems (ETPs) are constraint satisfaction problems involving multiple aspects, such as educators, students and educational resources. Based on the types of educational activities, ETPs can be roughly categorised as course timetabling, examination timetabling and school timetabling problems [2]. Course timetabling arranges educational facilities, such as classrooms and laboratories, ensuring no single student takes more than one course at the same time and no classroom hosts more than one class at a time. Examination timetabling prevents an individual student from taking more than one exam simultaneously but allows exams to share rooms and/or invigilators. Generally, the problem of the school timetabling is to allocate educators to classes with the considerations of their availability and expertise [3], which is the focus of this study.

School timetabling problems (STPs) have been recognised as a non-polynomial (NP) complete problem as its variables vary from one educational institute to another [4]. Although STPs have been studied since the 1960s [5], this area has not been developed as quickly as course timetabling and examination timetabling problems, probably because of the isolation of studies in particular schools [6]. In addition, the wide variety of school

timetables complexifies the problems. For example, some schools [7, 8] treated courses, educators and rooms as resources to be allocated, whereas some schools bound courses and educators as a pair to be assigned [8, 9].

To solve STPs, many computational intelligent methods and approaches have been applied, which mainly are categorised as heuristic approaches and novel approaches [10, 11]. Heuristic approaches consist of metaheuristics and hyper-heuristics. Metaheuristics approaches are inspired by natural phenomena aiming at seeking better solutions rather than the best solution [12, 13]. Hyper-heuristics uses metaheuristics methods to select a metaheuristic to solve generalised solutions [14]. Novel approaches include hybrid approaches, fuzzy logic approaches and MAS. Hybrid approaches employ different methods to solve a problem in order to alleviate the weakness of a single method [15, 16]. Fuzzy logic approaches aim to solve those problems which are hard to be quantitated and modelled as they do not have a precise classification [17, 18]. MAS deploys serval computational intelligent methods as an agent to play different roles to collaboratively fulfil a common goal [19].

Based on the abovementioned approaches, many applications are developed to solve STPs. In the metaheuristics field, Odeniyi, Omidiora, Olabiyisi and Aluko proposed a modified simulated annealing approach for Fakunle Comprehensive High School in Nigeria [20]. With the aid of the annealing scheme through the temperature parameter introduction, the approach successfully reduced the convergence time and computational cost brought by the annealing algorithm in dealing with large search spaces. A simple genetic algorithm (SGA) was adopted by Sutar and Bichkar [21] to solve an STP with knowledge augmented operators and probabilistic repair in the crossover step. The result of the modified approach against the OR-Library dataset [21] suggested SGA could produce faster solutions for GA-based optimization problems than the conventional GA method. In the hyper-heuristics area, Ahmed, Ozcan, Kheiri [22] combined five different selection hyper-heuristics with three-move acceptance methods to challenge the ITC2011 instances. The outcome indicated the approach was better than evolutionary algorithms [23] and an adaptive large neighbourhood search algorithm [24] but not when compared to hybridized simulated annealing (SA) and stagnation-free late acceptance hill climbing[25]. Hybrid approaches are combinations of various approaches. Those combined approaches for solving STPs include but are not limited to cat swarm optimization (CSO) with a local search algorithm [26], particle swarm optimization (PSO) with hybrid artificial fish swarm (AFS) [27]. Babaei, Karimpour and Oroji employed a fuzzy c-means clustering algorithm to solve STP for Islamic Azad University [28] to reduce redundancy and consider lecturers' preferences. When school timetabling is being planned, it will get multiple stakeholders involved, such as educators, heads of school and administrators, to negotiate. Therefore, MAS is often applied to simulate the process and the parties of the negotiation. Oprea [29] demonstrated that MAS could handle the negotiations between faculties and minimize resource conflicts, while Tkaczyk, Ganzha and Paprzycki [30] emulated the school timetabling workflow of the University of Gdansk with MAS.

Although many efforts have been made in the educational timetabling research field for decades, there has still been no consensus about the standard formulation and data format [8]. In addition, most of the STP researches only focused on educators' availabilities when they studied the educator aspect [6, 31], and the attributes of educators' preferences and expertise were not taken into consideration. To address this issue, this paper chooses a common real-world school timetabling scenario to study and takes educators' preferences towards units (also known as 'courses' in some literature) and the corresponding expertise level into account. Based on the chosen scenario, this paper presents a model of the STP and proposes a modified ABC algorithm to solve the problem because the ABC algorithm has successfully solved ETPs in the course [32, 33] and examination [15, 34, 35] timetabling fields. In addition, this paper introduces a novel VSS method to reduce the search space. The proposed approach is simulated with a random-generated large dataset.

The experiment results demonstrate the proposed approach is able to solve the STP and handle a large dataset in an ordinary computer hardware environment.

The contributions of this research include a conceptual model and its mathematical formulation for the STP considering educators' availabilities, preferences and expertise as a whole, a novel VSS method for reducing the computational cost of handling a large searching space and a modified ABC algorithm to solve the proposed STP model.

The organisation of this article is as follows. Section 2**Error! Reference source not found.** presents the problem formulation, including a conceptual model of the STP, iden-tification of hard and soft constraints, and formulation of the objective function. Section 3 describes the proposed approach, consisting of a concept of educator allocation, a modi-fied ABC algorithm and a VSS construction method. Section 4 simulates the proposed approach with a case based on a local university's business scenario. A comparison study between the proposed approach and the Constraint Programming (CP) is presented in Section 5. Section 6 concludes the article and indicates future works.

2. Problem Formulation

As the descriptions and terminology of STP are dramatically different from study to study [6], this section firstly defines the STP studied and the terms used in this paper. After that, hard constraints and soft constraints of the STP will be identified followed by the objective function formulation.

2.1. Terminology

A *course* refers to an academic program that students need to learn to gain university degrees, for example, in a Bachelor of Mobile Application degree, Mobile Application is the course name.

A *unit* refers to the academic subject within a course, e.g., Java Programming is a unit of the course Mobile Application.

A *class* refers to the particular teaching activity being scheduled in a timeslot in a day, which could be a lecture, a tutorial, a workshop and/or other educational activities.

An *educator* refers to the educational staff who delivers lectures, tutorials, workshops, etc.

A *school week* is from Monday to Friday.

A *school day* is a day of a school week.

2.2. School Timetabling Concept Model

In this study, three parameters of educators have been taken into consideration: pref-erences, availabilities and expertise. The objective of the school timetabling is to allocate all the school educators to a scheduled course timetable that satisfies all the preferences, availabilities and expertise as much as possible. The conceptual model of the school time-tabling is presented in **Figure 1**. In the beginning, a course timetable is preconstructed according to all university course information. That is, all the class activities have been scheduled with their timeslots in a school week. The course timetable has ensured that every single student will not take more than one unit at the same time. Educators provide their availabilities in a school week and their preferences along with preference levels against each unit. After that, the head of school will generate a school timetable by allo-cating educators to the scheduled course timetable according to educators' availabilities, preferences and expertise. The level of expertise against units is predefined by the head of school.
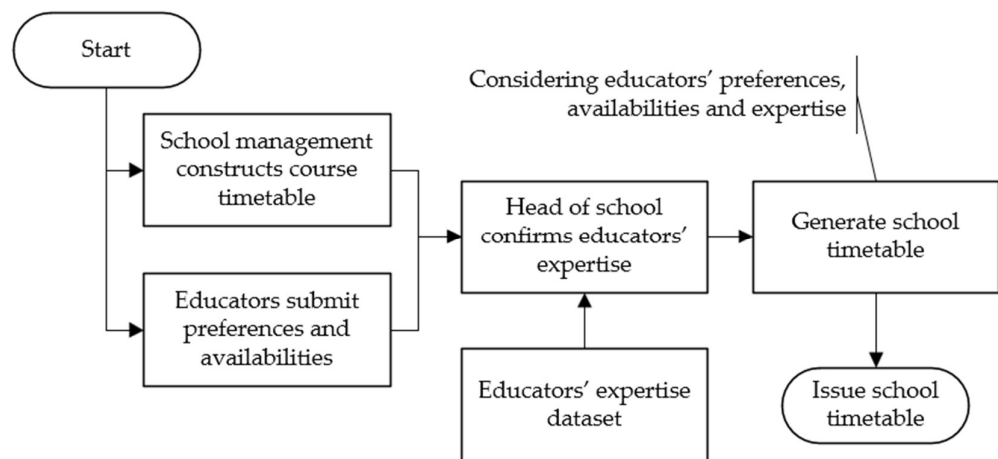
Figure 1. School timetabling concept model

2.3.Symbols and Notations

Parameters

L	Total number of educators.
O	Total number of units.
K	Total number of classes in a school week.
P	Maximum preference value.
E	Maximum expertise value.
V	Number of classes that an educator is allowed to deliver in a school week.
H	Number of hours in a school day (e.g., from 8 am to 6 pm, there are 10 hours).
D	Number of days in a school week.
M	Maximum number of hours a class lasts

Variables

d	Duration (in hours) of each class, $d = \{1, \dots M\}$
t	Set of educators, $t = t_1, \dots, t_L$
r	Set of units, $r = r_1, \dots, r_O$
c	Set of classes, $c = c_1, \dots, c_K$
p	Set of preference, $p = 0, \dots, P^*$
e	Set of expertise, $e = 0, \dots, E^*$
A	Set of timeslots of a school week, $A = \{1, \dots, H \times D\}$
u	Number of units that cannot be allocated with educators.

*When $p = 0$, it indicates an educator does not prefer a unit. When $e = 0$, it means the educator is not capable to deliver the unit. By default, all the p and e are 0. They need to be set by educators and the head of school.

Notation

A_{t_j} Set of timeslots that educator t_j is available in the school week. If educator t_1 is available in a whole school week, then $A_{t_1} = \{1, \dots, H \times D\}$

A_{c_i} Set of timeslots that class c_i is scheduled to in the school week. If class c_1 is scheduled in the first three hours of the school week, then $A_{c_1} = \{1,2,3\}$

2.4.Hard Constraints

Hard constraints are those conditions that the solutions have to satisfy. In this study, the hard constraints listed below are binary values. If a hard constraint is violated, it will be given value zero, otherwise, value 1 will be assigned. The hard constraints are notated as g and mathematically modelled as below:

$g1$: No an educator can deliver more than one class simultaneously.

$$\left(A_{t_j} \supseteq A_{c_i} \ \& \ A_{t_j} \supseteq A_{c_v} \ \& \ A_{c_i} \cap A_{c_v} = \emptyset\right) = True \tag{1}$$

where $i \in \{1,2,\ldots,K\}; v \in \{1,2,\ldots,K\}; v \neq i; j \in \{1,2,\ldots,L\};$

$g2$: An educator will only be assigned to the timeslots of a class when s/he is available. Meanwhile, the class must be fully allocated to an educator, meaning that the allocated educator must be available during the whole period of the class.

$$A_{t_j} \supseteq A_{c_i}, i \in \{1,2,\ldots,K\}; j \in \{1,2,\ldots,L\} \tag{2}$$

$g3$: One class only can be allocated with one educator.

$$\sum_{j=1}^{L} x_{ij} = 1, i \in \{1,2,\ldots,K\} \tag{3}$$

where i is the index of classes and j is the index of the educators; x_{ij} are decision variables, which is explained in Equation 6.

$g4$: one educator is allowed to teach no more than V classes in a school week

$$\sum_{i=1}^{K} x_{ij} \leq V, j \in \{1,2,\ldots,L\} \tag{4}$$

where i is the index of classes and j is the index of the educators; x_{ij} are decision variables, which is explained in Equation 6.

2.5.Soft Constraints

Soft constraints are those conditions that the solutions do not have to satisfy but are preferably satisfied. The following soft constraints have been identified:

- Educators will be assigned with the units most preferred.
- Units will be allocated with the educators who are more capable to teach.
- All units must be allocated with educators.

In practice, some units cannot be allocated with educators. For example, when a unit has not been favoured by any educator, or when two units are only favoured by one educator, but these two units are time conflicted.

2.6.Objective Function

The objective of this study is to maximise both all educators' satisfaction and all units' quality of teaching, in other words, to let the units be taught by the most qualified educators and to let the educators teach their most preferred units. When a unit is allocated with an educator, its allocation quality will be decided by the educator's expertise level and associated preference level, subject to the hard constraints being satisfied. This study defines the objective value q of a class allocation as shown in Equation 5, which is the multiplication of preferences (p) and expertise (e) values given to class i.

$$q_{ij} = p_{ij} \times e_{ij} \ \ (i \in \{1,2,\ldots,K\}; j \in \{1,2,\ldots,L\}) \tag{5}$$

where p_{ij} refers to the preference of educator j toward class z; p_{ij} represents the expertise of educator j to class i.

According to the concept model, the course timetable is preconstructed. The STP can be transferred to an allocation problem. **Figure 2** is the bipartite graph for the STP.
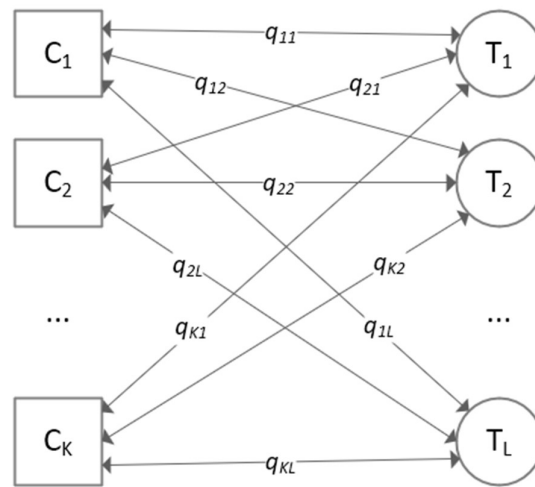
Figure 2. Bipartite graph for the STP

The objective function for school timetabling is to maximise the sum of every class's objective value (q).

Therefore, the STP optimization model can be presented as:

$$max \sum_{i=1}^{K} \sum_{j=1}^{L} q_{ij} x_{ij} \Big/ u \tag{6}$$

where $\qquad\qquad\qquad q_{ij} = p_{ij} \times e_{ij}$

x_{ij} are decision variables

$$x_{ij} = \begin{cases} 1 & allocated \\ 0 & unallocated \end{cases}$$

$$u = \begin{cases} K - \sum_{i=1}^{K} \sum_{j=1}^{L} x_{ij}, & if \quad K \neq \sum_{i=1}^{K} \sum_{j=1}^{L} x_{ij} \\ 1, & otherwise \end{cases}$$

Where $i \in \{1,2,\ldots,K\}; j \in \{1,2,\ldots,L\}; u$ is the number of unallocated units (if all the units are allocated, then u will be set to be 1).

Subject to:

$g1. \left(A_{t_j} \supseteq A_{c_i} \quad \& \quad A_{t_j} \supseteq A_{c_v} \quad \& \quad A_{c_i} \cap A_{c_v} = \emptyset \right) = True$

$g2. \; A_{t_j} \not\supseteq A_{c_j}$

$$g3. \; \sum_{j=1}^{L} x_{ij} = 1$$

$$g4. \; \sum_{i=1}^{K} x_{ij} \leq V$$

where $\quad i \in \{1,2,\ldots,K\}; v \in \{1,2,\ldots,K\}; v \neq i; j \in \{1,2,\ldots,L\};$

3. The Proposed Approach

This section firstly presents a basic concept of allocating educators to the scheduled course timetable. Based on the concept, a modified ABC algorithm is introduced. After that, a novel method named VSS is proposed to solve the gigantic searching space issue.

3.1. The Basic Concept of Educator Allocation

The research aims to allocate educators to a scheduled course timetable and attempts to satisfy their availabilities, preferences and expertise as much as possible. The allocation concept can be illustrated in Figure 3.

In Figure 3, the left side is an example of a scheduled course timetable listing the classes along with their weekdays, start times and durations. The educator ID column is to be filled with the IDs of the educators who are allocated to the classes. Weekdays and start times indicate which weekday the class is allocated to and at what time the class will start. Duration shows how long the class will last. On the right side of Figure 3, the educator list is presented including educator ID, preference and expertise, and availability columns. The educator ID column lists educators' IDs $t_j (j \in \{1, 2, \ldots, L\})$; the preference and expertise column stores a two-dimensional table, exampled in Table 1, which has K rows and three fields storing all the unit IDs, and their preferences and the expertise levels of an educator against those units. If an educator has no interest to teach a class, for example c_k in Table 1, then the preference and expertise will be filled with 0; otherwise, the corresponding level number will be put. The availability column contains a one-dimensional array converted from the two-dimensional school week timetable exampled in The basic process of allocating educators is to match each educator against a unit. Figure 3 can be used as an example to explain the process. Assuming that t_1 is capable ("capable" means the expertise value e is not equal to zero) and available to teach c_1, t_2 and t_3 are capable to teach c_2 and c_3, but only t_3 is available to teach c_2. Firstly, the program will consult the first unallocated educator from the educator list, which is t_1, and allocate it to the first unallocated class, which is c_1, as t_1 is capable and available. Then, the corresponding cell on t_1's availability table will be marked with "N". After that, t_2 becomes the first unallocated educator. However, t_2 is not available to teach c_2, therefore the consecutive unallocated educator, that is t_3, will be consulted. As t_3 is capable and available, so t_3 will be assigned to c_2. When c_2's allocation is finished, t_2 becomes the first unallocated educator and will be consulted for the c_3. As t_2 is capable and available to teach c_3, thus t_2 takes c_3. When all the educators and their profiles have been consulted, a timetable solution would be outputted.

Table 2. The number of elements in the availability array is $H \times D$. The availability column indicates the availabilities of each educator in a school week. If an educator is available at, for example, 8 a.m. on Monday in Table 2, then the first cell of the array will be marked with a Y, otherwise, an N will be stored.
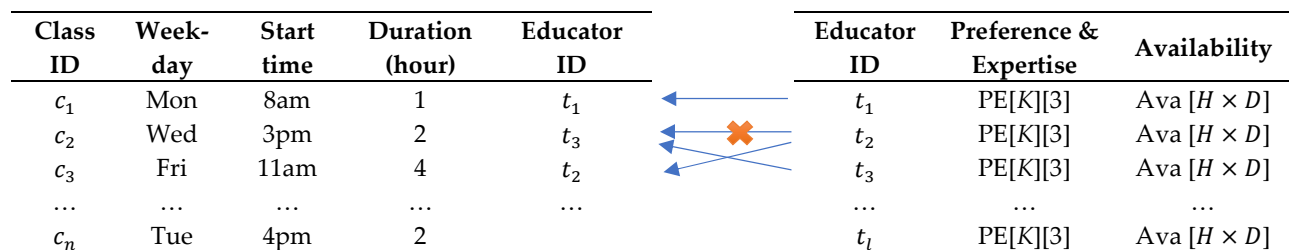
Class ID	Week-day	Start time	Duration (hour)	Educator ID		Educator ID	Preference & Expertise	Availability
c_1	Mon	8am	1	t_1		t_1	PE[K][3]	Ava [$H \times D$]
c_2	Wed	3pm	2	t_3		t_2	PE[K][3]	Ava [$H \times D$]
c_3	Fri	11am	4	t_2		t_3	PE[K][3]	Ava [$H \times D$]
…	…	…	…	…		…	…	…
c_n	Tue	4pm	2			t_l	PE[K][3]	Ava [$H \times D$]

Figure 3. Concept of allocating educators to units

Table 1. Example of Preference & Expertise table

Class ID	Preference	Expertise
c_1	2	3
c_2	1	1
…	…	…
c_{31}	3	2
…	…	…

c_K	0	0

The basic process of allocating educators is to match each educator against a unit. Figure 3 can be used as an example to explain the process. Assuming that t_1 is capable ("capable" means the expertise value e is not equal to zero) and available to teach c_1, t_2 and t_3 are capable to teach c_2 and c_3, but only t_3 is available to teach c_2. Firstly, the program will consult the first unallocated educator from the educator list, which is t_1, and allocate it to the first unallocated class, which is c_1, as t_1 is capable and available. Then, the corresponding cell on t_1's availability table will be marked with "N". After that, t_2 becomes the first unallocated educator. However, t_2 is not available to teach c_2, therefore the consecutive unallocated educator, that is t_3, will be consulted. As t_3 is capable and available, so t_3 will be assigned to c_2. When c_2's allocation is finished, t_2 becomes the first unallocated educator and will be consulted for the c_3. As t_2 is capable and available to teach c_3, thus t_2 takes c_3. When all the educators and their profiles have been consulted, a timetable solution would be outputted.

Table 2. Example of educator's availability

	Mon	**Tue**	**Wed**	**Thu**	**Fri**
8am	Y	N	Y	N	N
9am	Y	N	N	N	N
10am	Y	Y	N	N	N
11am	Y	Y	Y	N	N
12pm	N	Y	Y	N	N
1pm	N	N	Y	N	Y
2pm	N	N	Y	N	Y
3pm	N	Y	Y	N	Y
4pm	N	N	Y	N	Y
5pm	Y	N	Y	N	Y
6pm	Y	N	Y	N	Y

note: Y refers to available, N means unavailable. This table will be converted to be a one-dimensional array. The total number of cells of the array equals to the multiplication of $H \times D$. The index 1 refers to the first timeslot of the first day. The index $H \times D$ indicates the last timeslot of the last day.

With the concept of educator allocation, it is known that a specific educator list sequence subjected to the course timetable will always output the same school timetable. For instance, in Figure 3, educator t_1 will always be assigned to the class c_1, even if educator t_2 might be more suitable (e.g., t_2 has a higher q value than t_1), as t_1 will be consulted before t_2. Therefore, an educator list can be considered as a solution and therein changing the educator list sequence can obtain different solutions, from which it can be reasoned that the number of possible solutions is the number of permutations of the educator list, in other words, is the factorial of the number of educators. All of the possible solutions form the entire solution searching space. However, the number of educators in a university could be over 100, meaning that the number of solutions will be gigantic and lead to the impossibility of seeking the best solution by traversing the whole searching space. Therefore, an artificial intelligence algorithm is needed.

3.2. ABC Algorithm

The ABC algorithm is chosen for the proposed STP for its ability to solve multivariable, multimodal optimisation problems efficiently. Also, the ABC algorithm can be easily implemented without requiring many parameters. The ABC algorithm is inspired by the behaviours of honeybees and was introduced by Karaboga [36]. It simulates the ways that honeybees forage for food sources, which helps bees efficiently and effectively seek better food sources in a vast area. The mechanisms the ABC algorithm uses include positive feedback, negative feedback, fluctuations and multiple interactions. These four

mechanisms help the bees to explore new food sources and to avoid over-populating a source, and also ensure information can be shared with each bee. Labour division is another feature of the ABC algorithm. Bees are categorised to be employee bees, onlooker bees and scout bees. Employee bees are responsible for new food source exploration; onlooker bees are in charge of food source analysis and exploitation; scout bees avoid food source exhaustion.

The ABC algorithm consists of four stages: population, employee bee, onlooker bee and scout bee stages.

In the solution population stage, the ABC algorithm randomly populates several solutions (bees) in a searching space with Equation 7.

$$s_i^d = s_{min}^d + random(0,1)\left(s_{max}^d - s_{min}^d\right) \tag{7}$$

where i is one of the nodes in the searching space, d is dimension. s_{max}^d and s_{min}^d is the upper bound and low bound for the dimension d.

In the employee bee stage, neighbours of the populated solutions will be looked for with Equation 8.

$$es_i^d = s_i^d + \varphi_i^d\left(s_i^d - s_k^d\right) \tag{8}$$

where k is randomly chosen from the searching space and $k \neq i$. φ_i^d is randomly generated in the range of [-1,1]. s_k^d is a neighbour of s_i^d. The probability of each employee bee es will be calculated with Equation 9.

$$p_i = \frac{fit(es_i)}{\sum_{n=1}^{SN} fit(es_n)} \tag{9}$$

In the onlooker bee stage, a random number [0,1] will compare to p_i for each employee bee, if the random number is better than a p_i, an onlooker bee will be sent to look for a neighbour of the employee. A parameter called trail will be used in this stage. Trail limits the times that a food source has been explored. If the exploration time reaches the trail and the better neighbour has not been found, the food source will be abandoned.

In the scout bee stage, the onlooker bee(s), whose food source(s) is abandoned, will become scout bee(s) and randomly populate a new solution. The equation scout bees use is similar to Equation 7.

3.3. A Modified ABC Algorithm for Proposed STP

The modified ABC algorithm is presented in Figure 4. To adapt the ABC algorithm, some modifications have been made to suit the investigated problem. The modifications are described as follows.

```
/*Initialization stage*/
01: Retrieve the scheduled course timetable and educators' profiles.
02: Define the neighbour search range, number of Traits and read parameters.
03: Initialize the food source and construct school timetable satisfying all hard constraints as in
    Equation 1, 2, 3 and 4.
04: Evaluate school timetable's objective values with Equation 6.
05: Send the employed bees to the current food sources.
06: Iteration = N
07: FOR (each iteration)
/*Employed Bee Phase*/
08:FOR (each employed bee)
09:        Seek new food source from neighbourhood in VSS satisfying Equation 1, 2, 3 and 4.
10:        Construct school timetables and evaluate their objective value with Equation 6, apply
           greedy selection
11:END FOR
12:  Calculate the probability p for each food source with Equation 9.
/*Onlooker Bee Phase*/
13:  FOR (each onlooker bee)
14:        Send onlooker bees to food sources based on p
15:        Find a new food source in its neighbourhood in VSS satisfying Equation 1, 2, 3 and 4.
16:        Construct school timetables and evaluate their objective value with Equation 6, Apply
           greedy selection. Set Trait+1 if applicable.
17:  END FOR
/*Scout Bee Phase*/
18:  IF (any onlooker bee becomes scout bee)
19:        Send scout bee to a randomly produced food source satisfying Equation 1, 2, 3 and 4.
20:  END IF
21:  Memorize the best solution achieved so far
22: END FOR
23: Output the best solution achieved
```

Figure 4. Pseudo code of the modified ABC algorithm

Before food source initialisation, a scheduled course timetable will be retrieved from the course database followed by the educator profiles data retrieval (Step 1 in **Error! Reference source not found.**). The examples of course timetable and educator profile are illustrated in **Error! Reference source not found.**. Since the boundary (Equation 7 and 8) of the searching space and the current neighbourhoods are needed to be known when bees are foraging for food sources, the searching space will be formed beforehand. However, as mentioned before, the number of solutions is gigantic, so this research proposes a VSS to tackle this dilemma. The VSS is detailed in Section 3.4. To construct the final solutions from the food sources that bees forage from the VSS (Step 03, 09, 15 and 19 in **Error! Reference source not found.**), the school timetable construction will be applied and detailed in Section 3.5. The objective values of constructed final solutions will be evaluated with Equation 6 (Steps 04, 10 and 16 in **Error! Reference source not found.**).

3.4.Virtual Searching Space Construction

VSS aims to provide an entire solution pool for the proposed algorithm rather than constructing a gigantic searching space. As discussed above, the searching space is formed through permutation manipulation with the magnitude of the factorial number of educators. Thus, it would be impractical to physically construct the whole searching space as it would exceed the memory capacity of ordinary computers not to mention the computational time for generating all the solutions for each educator list. To tackle this dilemma, we propose a novel VSS approach. VSS does not construct a searching space by the direct permutations. Instead, it "imagines" the solutions to be allocated in the searching space in the way as exampled in **Error! Reference source not found.**.

3.4.1. How VSS works

Assume one bee is employed and four educators are to be allocated, then the searching space will be similar to **Error! Reference source not found.** with four regions (A, B, C and D) and 24 possible solutions (hereafter referred to as columns) numbered.

Table 3. Example of searching space.

A region						B region					
1	2	3	4	5	6	7	8	9	10	11	12
A	A	A	A	A	A	B	B	B	B	B	B
B	B	C	C	D	D	A	A	C	C	D	D
C	D	B	D	B	C	C	D	A	D	A	C
D	C	D	B	C	B	D	C	D	A	C	A
C region						D region					
13	14	15	16	17	18	19	20	21	22	23	24
C	C	C	C	C	C	D	D	D	D	D	D
A	A	B	B	D	D	A	A	B	B	C	C
B	D	A	D	A	B	B	C	A	C	A	B
D	B	D	A	B	A	C	B	C	A	B	A

In the food source initialisation phase (Step 03 in **Error! Reference source not found.**), VSS does not need to know the solutions' coordinators in the searching space as any educator list combination will be in the solution pool. VSS randomly generates an educator sequence B-A-D-C (Column 8).

In the neighbour seeking phase, VSS swaps the positions of A and C in Column 8 to obtain Column 10. The swap method can decide the distance, vector and boundary of neighbour solutions with the following rules:

Rule 1: Fixing top element(s) can confine the boundary of the neighbourhood. For example, when the first element "B" is fixed, the neighbourhood will be in B region. When "B-A" is fixed, there are only two neighbours, B-A-C-D (Column 7) and B-A-D-C (Column 8). This rule confines the upper bound and low bound for a dimension as s_{max}^d and s_{min}^d in Equation 7.

Rule 2: The relationship of swapping elements determines the orientation. For example, if the higher element (A) switches with the lower one (C) (in practice, higher/lower element refers to higher/lower indexed element, which could be alphabetic order or numerical order.), the orientation will be rightward, vice versa. This rule plays the rule as φ_i^d in Equation 8.

Rule 3: The positions and distance between swapping elements decide the distance of neighbours. For example, if swapping the last two elements, "D" and "C" in B-A-D-C (Column 8), the neighbour is right next to each other. If swapping higher elements, like "A" and "D" the neighbour is three steps away. This rule plays the role of k in Equation 8.

These three rules have been successfully applied to a large dataset with 150 educators experimented in Section 4.

The implementation of VSS in modified ABC algorithm

VSS uses swapping methods to locate neighbour solutions and confine the searching space boundary. The pseudocode of VSS implementation is detailed in.

01: Randomly select first educator (index is I_{first}) from the given educator list
02: Randomly generate neighbour search range r.
03: Randomly generate number b ($b = -1 \ or \ 1$).

04: Set the second educator's index $I_{second} = I_{first} + r \times b$

05: If $I_{second} > L \ or \ I_{second} < 1$, repeat Step 2 ($L$ is the number of educators)

06: Swap I_{first} and I_{second}

Figure 5. Pseudocode of VSS

3.5.School timetable construction

The proposed approach will not generate a timetable solution directly. Instead, it will randomly select educator lists from VSS as the food sources of the modified ABC algorithm. After that, the food sources will be passed to the timetable constructor to output complete solutions. The steps 03, 04, 10 and 16 in **Error! Reference source not found.** are the entries food sources passed to the school timetable constructor. **Error! Reference source not found.** demonstrates the construction process.

A class will be selected from the scheduled course timetable subsequently to match an unallocated educator in the given educator list. If the educator is allocatable due to her/his preference and availability, then the class will be assigned to that educator. Otherwise, the consecutive educator will be consulted. The scheduled course timetable will be traversed at particular times depending on the number of classes (parameter V) that an educator is allowed to deliver in a school week. If all the classes have been allocated with an educator, a constructed school timetable will be returned (steps 03, 04, 10 and 16 in **Error! Reference source not found.**). Otherwise, the unallocated classes will be revisited to seek the educators who are capable and available to teach those classes despite the limitation of the V. When the program reaches the end of the course timetable again, a school timetable will be returned regardless even if there may be some unallocated classes left. The extra revisiting aims to minimise the number of unallocated classes and provides suggestions to the head of school to negotiate with those educators who have the relevant expertise but do not prefer to teach the classes.

```
1 :   Receive an educator list and the number of units an educator can take
2 :   FOR (V)
3 :       FOR (K)
4 :           FOR (each unallocated educator in the educator list)
5 :               IF (the educator is available and prefer to teach the class)
6 :                   Allocate the educator to the class
7 :                   Modify the educator's availability
8 :   IF (all classes are allocated with an educator)
9 :       Return constructed School Timetable
10 : FOR (each unallocated class in the course timetable)
11 :       FOR (L)
12 :           IF (the educator can and is available to teach the class)
13 :               Allocate the educator to the class
14 :               Modify the educator's availability
15 : Return constructed School Timetable
```

Figure 6. Pseudocode of school timetable solution construction

4.Experiment

The experiment follows the proposed school timetabling conceptual model as illustrated in Figure 1. The scenario is detailed as below.

- The course timetable is prescheduled and fixed based on the course enrolment information. The class activities have been scheduled in the weekly timeslots.
- Before planning a school timetable, educators will submit their expression of interest to the administration, including the units they want to teach and the preferences towards each unit.

- Educators need to provide their unavailability form in the school week. The number of unavailable hours cannot be more than 3 hours.
- The head of school will confirm and adjust each educator's expertise against a unit.
- The goal of the school timetabling is to satisfy educators' preferences and classes' qualities and attempt to ensure every unit is allocated with an educator.

4.1.Experimental settings

According to the proposed conceptual model in **Error! Reference source not found.**, this research set the following parameters to randomly generate a dataset consisting of a predefined course timetable and an educator roster along with their availabilities, preferences and expertise levels. After that, the generated dataset will be simulated with the proposed approach.

- $K = 300$
- $L = 150$
- $V = 5$
- $H = 8$
- $D = 5$
- $P = 5$
- $E = 3$
- $d = 1$ or 2
- Maximum unavailable timeslots an educator can have in a week: 3

The proposed approach has been experimented in the below environment.

- Operation System: Windows 10 Pro Edition
- Integrated development environment: IntelliJ IDEA 2020.3.2
- Programming language: JAVA
- The computer hardware system is Intel® Core™ i7-1065G7 1.30GHz Processor with 16.0 GB of memory

Ten samples have been configured in **Error! Reference source not found.** to experiment with the proposed algorithm with four ABC parameters: number of bees, neighbour range, number of iterations and number of traits. The number of bees presents the number of solution population; neighbour range confines the coverage of an exploration area of a bee; number of iterations is set to test whether increasing search times will improve the result; and number of traits determines the depth of exploitation of a food source. Each sample is tested ten times.

Table 4. Parameter settings for modified ABC algorithm

Sample	Num of bees	Neighbour range	Num of iterations	Num of traits
A	5	5	1000	10
B	20	5	1000	10
C	5	30	1000	10
D	5	5	10000	10
E	5	5	1000	30
F	20	30	10000	10
G	20	30	10000	30
H	5	5	10000	30
I	20	30	10000	30
J	40	60	20000	60

4.2.Experimental results

The experiment results for each sample listed in **Error! Reference source not found.** are shown in **Error! Reference source not found.**. Three key results demonstrating the performance of the proposed approach are chosen, which are values of the objective function, time spent and number of unallocated classes. Objective function values are calculated with Equation 6. Time spent presents computational time spent on the given samples. Unallocated class indicates the number of classes that cannot be allocated educators.

There are 27 unallocated classes, which has been manually reviewed and confirmed that those failures are due to educators' availabilities or interests. This is also because of the data quality resulting from all the experimental data being randomly generated. For example, two classes only can be taught by educator A, but these two classes share a timeslot with each other, or some classes have not been chosen by any educator. For this scenario, the educational administration will need to solve the issue manually, such as negotiating with educators and recruiting new educators. For simplifying the data comparison, the number of unallocated classes shown in **Error! Reference source not found.** and **Error! Reference source not found.** has been offset by 27.

Based on the result, the following conclusion could be drawn.

- The proposed approach can obtain a feasible solution in a gigantic searching space in a short time. In Sample A, the time spent for the best solution is 4.742 seconds.
- Deploying more bees can help improve objective value. Sample B has a better result than Sample A in objective value (43.4970 to 42.4299 on average) as Sample B populates 20 bees and Sample A deploys five bees.
- Increasing the number of program iteration can obtain a better result. Sample D has a better objective function value over Sample A (43.6596 over 42.4299 on average) as Sample D has 10,000 iterations and Sample A has 1,000 iterations.
- Increasing both the number of bees and the number of iterations can improve the result. Comparing Sample A and Sample F, the objective value improves from 42.4299 to 44.7148 on average, but the time spent also increased significantly (from 4.983 seconds to 197.513 seconds on average).
- Expanding the neighbourhood can slightly improve the result. Sample C expands the neighbourhood three times compared to Sample A, but the result only improves 0.85% (from 42.4299 to 42.7910 on average).
- Increasing the exploitation will not benefit the solution and instead provides worse results. Compared to Sample A, although Sample E triples the number of traits, the objective value decreases and the number of unallocated units increases. It is because that exploitation reduces the opportunity of exploration, which is also proved by the comparing the results between Sample D and Sample H, and between Sample F and Sample G.
- Sample J enlarges the variables. Although the improvement of the result is obvious, the time spent is significant. Compared to Sample A, the objective value is increased by 10.3% (from 42.4299 to 46.8359), but the time Sample J spends is 165 times than that of Sample A.

Table 5. Experiment result (ten times run)

Sample	Objective values		Time-spent(second)		Unallocated classes	
	Avg.	Best	Avg.	Best	Avg[1]	Best
A	42.4299	44.0740	4.983	4.742	0.5	0
B	43.4970	44.7037	18.8521	18.453	0.1	0
C	42.7910	43.7407	5.0719	4.834	0.5	0
D	43.6596	44.8518	48.843	47.238	0	0
E	41.9709	43.9256	4.9919	4.818	1	0
F	44.7148	45.3703	197.513	193.957	0	0
G	44.5255	45.1481	201.0127	194.826	0	0

H	42.5502	43.2222	50.3867	46.775	0.3	0
I	44.4592	45.8148	206.4878	197.141	0	0
J	46.8359	47.1481	831.6302	813.494	0	0

[1] The numbers have been offset by 27. Those 27 classes have been manually confirmed as non-allocatable.

Overall, although Sample A uses the least bees and iterations with a smaller neighbour range and traits, it achieves acceptable results within several seconds of execution. Therefore, the settings for Sample A are recommended for the proposed approach.

Besides, with the aid of VSS, the computational cost has been significantly released. Since VSS will not generate the entire searching space, the actual computational cost for forming the searching space is the cost of generating real solutions. The computational cost (cs) of forming the searching space can be calculate with the Equation 10.

Set the number of bees as b, number of iterations as w then the number of solutions that the program (detailed in **Error! Reference source not found.**) generates in one run is

$$cs = b \times w \times 3^* \tag{10}$$

(*each bee populates one solution and at most generates two neighbour solutions, so each bee at most generates three solutions in one iteration.)

As explained above, the scale of the entire searching space is the factorial of the number of educators, that is $L!$. Therefore, the reducing computational cost can be calculated with Equation 11.

$$\left(1 - \frac{cs}{L!}\right) \times 100\% \tag{11}$$

Given that the result of factorial generally is gigantic, computational cost will be considerably reduced.

5.Comparison Study

With the purpose of verifying the performance of the proposed approach in solving STPs, a CP is implemented for the comparison study since STPs are considered constraint satisfaction problems and CP has been proven to successfully solve various problems, including timetabling [37]. The CP applied to the investigated problem is demonstrated in **Error! Reference source not found.**. The steps of testing data retrieval and solution population are the same as the ones detailed in the initialisation stage in step 1 of **Error! Reference source not found.**. After that, unassigned classes and the populated educator list will be consecutively visited. If the consulting educator satisfies the constraints of expertise and availability of the selected class, the educator will be allocated to the class. If none of the educators can be allocated to the current class, the class will remain unassigned and the consecutive class will be selected to look for an eligible educator. If the number of unassigned classes exceeds the limitation, backtracking will be triggered. If all the classes have attempted to be assigned an educator, a solution will be outputted and then its objective function value will be evaluated. If the backtracking reaches the first class of the course timetable, the current solution is not feasible and another solution will be populated.

To allow the modified ABC algorithm and CP be more comparative, the CP will be fed with Sample A, shown in **Error! Reference source not found.**, which can reach a better solution with the smallest population and iteration. As it is known that there are 27 unassignable classes in the dataset, to avoid over-backtracking, the compared approach sets up a backtracking trigger with the value 27, which is the N parameter in **Error! Reference source not found.**.
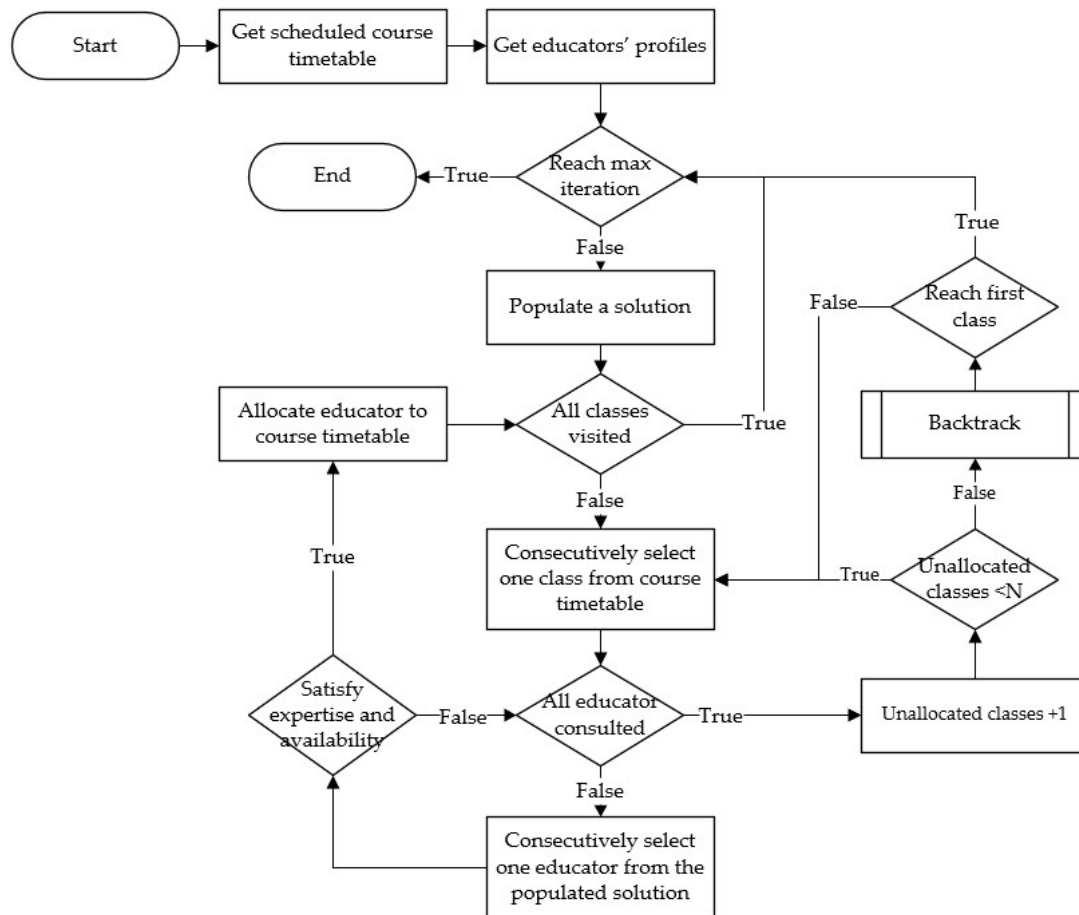
Figure 7. Constraint programming flowchart

This comparison study tests both the proposed modified ABC algorithm and CP ten times. The results of time spent in program execution, the number of unallocated classes and the objective function value will be compared. The comparison results are represented in **Error! Reference source not found.**.

Table 6. Comparison results of ABC approach and cp approach

Items	Proposed ABC		CP	
	Best	Average	Best	Average
Time-spent(second)	4.742	4.983	3.5	3.6108
Unallocated classes[1]	0	0.5	12	12
Objective function values	44.0740	42.4299	30.5128	30.1158

[1] The numbers have been offset by 27. Those 27 classes have been manually confirmed as non-allocatable.

From Table 6 the following findings can be discovered.
- CP is slightly faster than the modified ABC algorithm in solution-seeking (3.6108 seconds over 4.983 seconds on average).
- The modified ABC algorithm can allocate all the non-conflict classes. However, CP has 12 allocatable classes left.
- The modified ABC algorithm can seek a solution that is 40.88% better than CP in terms of objective function values gaining (42.4299 against 30.1158 on average).

Overall, the modified ABC algorithm can find out a better solution than CP in terms of hard constraint satisfaction and objective function optimisation. Although the proposed ABC is slower than CP in program execution, one second time difference for a practical scenario is insignificant.

6.Conclusion

This research aims at providing an effective solution for solving an STP with the considerations of educators' availabilities, preferences and expertise as a whole. STP is an ETP that has not advanced as quickly as the other two types due to its diversity and complexity. Most of the STP research only focused on educators' availabilities rather than taking educators' preferences and expertise into considerations. This paper proposed a conceptual model of STP and introduced a novel VSS method to reduce the searching space. A modified ABC algorithm is applied to the STP model.

The proposed approach is simulated with a random-generated large dataset. The experimental results demonstrate that the proposed approach is able to solve the STP and handle a large dataset in an ordinary computer hardware environment thereby significantly reducing computational costs. Compared to the traditional constraint programming method, the proposed approach is more effective and can provide more satisfactory solutions in considering educators' availabilities, preferences and expertise levels.

Author Contributions: Conceptualization, Kaixiang Zhu; Data curation, Kaixiang Zhu; Formal analysis, Kaixiang Zhu and Lily D Li; Investigation, Kaixiang Zhu and Lily D Li; Methodology, Kaixiang Zhu and Michael Li; Project administration, Lily D Li; Software, Kaixiang Zhu; Supervision, Lily D Li and Michael Li; Validation, Lily D Li; Writing – original draft, Kaixiang Zhu; Writing – review & editing, Lily D Li and Michael Li.

References

1. Wren A. Scheduling, timetabling and rostering—a special relationship? International conference on the practice and theory of automated timetabling: Springer; 1995. p. 46-75.
2. Schaerf A. A survey of automated timetabling. Artificial intelligence review. 1999;13(2):87-127.
3. Valouxis C, Gogos C, Alefragis P, Housos E. Decomposing the high school timetable problem. Practice and Theory of Automated Timetabling (PATAT 2012), Son, Norway. 2012;61.
4. Bashab A, Ibrahim AO, AbedElgabar EE, Ismail MA, Elsafi A, Ahmed A, et al. A systematic mapping study on solving university timetabling problems using meta-heuristic algorithms. Neural Computing and Applications. 2020;32(23):17397-432.
5. Odeniyi O, Omidiora E, Olabiyisi S, Oyeleye C. A Mathematical Programming Model and Enhanced Simulated Annealing Algorithm for the School Timetabling Problem. Asian Journal of Research in Computer Science. 2020:21-38.
6. Pillay N. A survey of school timetabling research. Annals of Operations Research. 2014;218(1):261-93. doi: 10.1007/s10479-013-1321-8.
7. Beligiannis GN, Moschopoulos CN, Kaperonis GP, Likothanassis SD. Applying evolutionary computation to the school timetabling problem: The Greek case. Computers & Operations Research. 2008;35(4):1265-80.
8. Post G, Ahmadi S, Daskalaki S, Kingston JH, Kyngas J, Nurmi C, et al. An XML format for benchmarks in high school timetabling. Annals of Operations Research. 2012;194(1):385-97.
9. Wilke P, Ostler J. Solving the school time tabling problem using tabu search, simulated annealing, genetic and branch & bound algorithms. the proceedings of the 7th International Conference on the Practice and Theory of Automated Timetabling (PATAT 2008), Montreal: Citeseer; 2008.
10. Babaei H, Karimpour J, Hadidi A. A survey of approaches for university course timetabling problem. Computers & Industrial Engineering. 2015;86:43-59.
11. Feizi-Derakhshi M-R, Babaei H, Heidarzadeh J. A survey of approaches for university course timetabling problem. Proceedings of 8th international symposium on intelligent and manufacturing systems, Sakarya University Department of Industrial Engineering, Adrasan, Antalya, Turkey2012. p. 307-21.
12. Salhi S. Heuristic search: The emerging science of problem solving. Springer; 2017.
13. Fonseca GH, Santos HG, Carrano EG, Stidsen TJ. Integer programming techniques for educational timetabling. European Journal of Operational Research. 2017;262(1):28-39.
14. Pillay N, Qu R. Hyper-Heuristics: Theory and Applications. Springer; 2018.
15. Fong CW, Asmuni H, McCollum B. A hybrid swarm-based approach to university timetabling. IEEE Transactions on Evolutionary Computation. 2015;19(6):870-84.
16. Ishak S, Lee LS, Ibragimov G. Hybrid Genetic Algorithm for University Examination Timetabling Problem. MALAYSIAN JOURNAL OF MATHEMATICAL SCIENCES. 2016;10(2):145-78.

17. Kouhbanani Nejad SF, Farid D, Sadeghi H. Selection of Optimal Portfolio Using Expert System in Mamdani Fuzzy Environment. Industrial Management Studies. 2018;16(48):131-51.
18. Bělohlávek R, Dauben JW, Klir GJ. Fuzzy logic and mathematics: a historical perspective. Oxford University Press; 2017.
19. Junn KY, Obit JH, Alfred R, Bolongkikit J. A Formal Model of Multi-agent System for University Course Timetabling Problems. Singapore: Springer Singapore; 2019. p. 215-25.
20. Odeniyi O, Omidiora E, Olabiyisi S, Aluko J. Development of a modified simulated annealing to school timetabling problem. International Journal of Applied Information Systems. 2015;8(2):16-24.
21. Sutar SR, Bichkar RS. Genetic algorithms based timetabling using knowledge augmented operators. International Journal of Computer Science and Information Security. 2016;14(11):570.
22. Ahmed LN, Özcan E, Kheiri A. (2015) Solving high school timetabling problems worldwide using selection hyper-heuristics. Expert Systems with Applications, 42 (13). pp. 5463-5471. ISSN 0957-4174.
23. Domrös J, Homberger J. An evolutionary algorithm for high school timetabling. Proceedings of the ninth international conference on the practice and theory of automated timetabling (PATAT 2012): Citeseer; 2012. p. 485-8.
24. Sørensen M, Stidsen TR. High School Timetabling: Modeling and solving a large number of cases in Denmark. Proceedings of the ninth international conference on the practice and theory of automated timetabling (PATAT 2012)2012. p. 359-64.
25. Fonseca GH, Santos HG, Carrano EG. Late acceptance hill-climbing for high school timetabling. Journal of Scheduling. 2016;19(4):453-65.
26. Skoullis VI, Tassopoulos IX, Beligiannis GN. Solving the high school timetabling problem using a hybrid cat swarm optimization based algorithm. Applied Soft Computing. 2017;52:277-89.
27. Katsaragakis IV, Tassopoulos IX, Beligiannis GN. A Comparative Study of Modern Heuristics on the School Timetabling Problem. Algorithms. 2015;8(3):723-42.
28. Babaei H, Karimpour J, Oroji H. Using fuzzy c-means clustering algorithm for common lecturers timetabling among departments. 2016 6th International Conference on Computer and Knowledge Engineering (ICCKE): IEEE; 2016. p. 243-50.
29. Oprea M. MAS_UP-UCT: A multi-agent system for university course timetable scheduling. International Journal of Computers Communications & Control. 2007;2(1):94-102.
30. Tkaczyk R, Ganzha M, Paprzycki M. AgentPlanner-agent-based timetabling system. Informatica. 2016;40(1).
31. Tan JS, Goh SL, Kendall G, Sabar NR. A survey of the state-of-the-art of optimisation methodologies in school timetabling problems. Expert Systems with Applications.165:113943.
32. Karaboga D, Gorkemli B, Ozturk C, Karaboga N. A comprehensive survey: artificial bee colony (ABC) algorithm and applications. Artificial Intelligence Review. 2014;42(1):21-57.
33. Bolaji ALa, Khader AT, Al-Betar MA, Awadallah MA. University course timetabling using hybridized artificial bee colony with hill climbing optimizer. Journal of Computational Science. 2014;5(5):809-18.
34. Bolaji ALa, Khader AT, Al-Betar MA, Awadallah MA. A hybrid nature-inspired artificial bee colony algorithm for uncapacitated examination timetabling problems. Journal of Intelligent Systems. 2015;24(1):37-54.
35. Alzaqebah M, Abdullah S. Hybrid bee colony optimization for examination timetabling problems. Computers & Operations Research. 2015;54:142-54.
36. Karaboga D. An idea based on honey bee swarm for numerical optimization. Technical report-tr06, Erciyes university, engineering faculty, computer …; 2005.
37. Bukchin Y, Raviv T. Constraint programming for solving various assembly line balancing problems. Omega. 2018;78:57-68.