Case Study: University Lecture Timetabling Without Pre-registration Data

San-Nah Sze, Chia-Lih Bong, Kang-Leng Chiew, Wei-King Tiong, Noor Alamshah Bolhassan

Department of Computational Science and Mathematics
Department of Computing and Software Engineering
Universiti Malaysia Sarawak
94300 Kota Samarahan, Sarawak, Malaysia
snsze@unimas.my, lihlih1206@gmail.com, klchiew@unimas.my, wktiong@unimas.my, bnalamshah@unimas.my

Abstract

This paper focuses on university lecture timetabling at Faculty of Computer Science and Information Technology (FCSIT), Universiti Malaysia Sarawak (UNIMAS). In this case study, course pre-registration is not a practice. Therefore, there is no precise estimation on course registration and causes faculty's experienced planner to arrange the timetable by *curriculum-based*. However, *curriculum-based* timetable will create a lot of changes after the semester has started. Besides, students are increasing consistently from semester to semester although the number of venue resources remains the same. Due to all these issues, the objective of this study is to develop a computerised algorithm to minimise the clashes issue and increase venue utilisation. Data pre-processing algorithm was carried out to predict course registration. Then, a two-stage heuristic method is proposed to solve the faculty course timetabling problem by *student-based*. The simulator was tested with three real semesters' data from FCSIT. All the timetable solutions generated by the simulator are no-clash solution with minimum unallocated courses. In term of venue utilisation, two-stage heuristic solution manages to allocate exactly with the demand up to 98% but real solution can perform best at only 75%.

Key words: lecture timetabling; heuristic; post-enrolment

Introduction

University course timetabling is an important practical problem that is regularly needed to be solved in institutions [1]. Degree of difficulty for course timetabling increases enormously with an increasing number of students and courses [2]. It is tough enough to produce a feasible solution, not to mention a solution that satisfies people who rely on it [3].

Course timetabling is complex with uncertainties as course requests of students are unknown at the point in time when the timetable is created [4]. The number of students who participate in each course which indicates the course size for suitable venue allocation remains unknown for timetabling. Moreover, the problem becomes much more complex when it involves large number of students because the individual freedom of the students to choose courses grows along [5].

During the second International Timetabling Competition (ITC) in the year 2007, the competition proposed to split university course timetabling problem into two formulations, namely *curriculum-based* [6] and *post-enrolment* [7]. Since then, these two formulations have received more attention than others among academic researchers. The difference between two formulations is that the constraints and objectives in *curriculum-based* course timetabling are based on the concept of *curriculum*, which is a set of courses particularly for a group of students (usually grouped by intake and programme). On the other hand, the constraints and objectives are based on the course registration for *post-enrolment* course timetabling (also known as *student-based* course timetabling).

Case Study

The timetabling problem as described in this study is a real problem identified at Faculty of Computer Science and Information Technology (FCSIT), Universiti Malaysia Sarawak (UNIMAS). FCSIT offers a total of five undergraduate programmes. All five programmes are four-year programmes with a total of seven semesters on campus and one semester of industrial training. FCSIT has designed different course plans for each programme with a list of courses to enrol for each semester.

Similar to the policy implemented in most universities, student has to pass all courses in order to graduate. Some advanced courses require students to pass certain prerequisite courses before they can enrol them. These situations cause that particular group of students to have a different list of course registrations than other students from the same batch and programme. Students who failed a course in the previous semester and repeat the course in current semester are called "repeaters". For FCSIT course timetabling problem, "repeaters" are defined as students with course registrations which are not tally with the course plan. Thus in this case, "repeaters" include student who fail and repeat a course or student who enrol courses without following the course plan.

Generally, universities or institutions solve their course timetable either through curriculum-based timetabling formulation or student-based post-enrolment timetabling formulation. Previous solving method in FCSIT is a manual method developed by experienced timetable planners. Starting from semester 1 (2014/2015), the faculty utilised a commercial timetabling software to assist the timetabling process. Both timetabling means are based on list of courses in the course plan to design a curriculum-based timetabling.

There are several issues with the current FCSIT course timetabling problem:

(a) Inconsideration of "repeaters" in curriculum-based timetabling
(b) Inconsideration of "critical students/graduating student"
(c) No course registration data for post-enrolment timetabling
(d) Insufficient large capacity venues

The number of students' intake for FCSIT has increased tremendously in the last 5 years (as shown in Figure 3.1). Existing venue resources at the faculty are more on smaller size

ISBN 978-1-5090-4897-7

venue. The trend of increasing student intake is predicted to be continued and keep growing in the future. Therefore, large size venue insufficiency is getting more and more serious unless some planning and actions are taken.

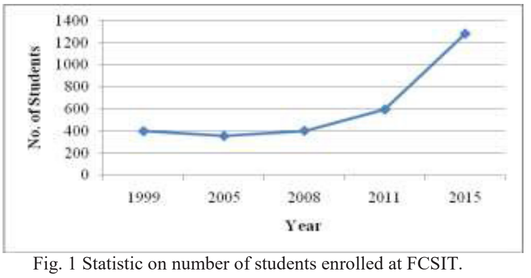

Fig. 1 Statistic on number of students enrolled at FCSIT.

Data Pre-Processing Algorithm

In this research, there is a need to conduct data pre-processing in order to get more comprehensive timetable. To be specific, the purposes of data pre-processing are identified as follow:

(a) Prediction on course registration
(b) Identify critical students and courses

A. Prediction on Course Registration

As there do not have course registration data, data pre-processing algorithm is designed to predict "who" and the number of potential students for all offered course for that semester. Few types of data are required in order to generate the potential list. Student data and their previous semesters' result data are compared with list of courses on course plan to predict course registrations including the potential of repeating the courses.

Each student has a list of all the courses to be enrolled based on his/her course plan. Excluding freshman student, each student has a list of previous semesters' passed courses. Student result data is important in course registration prediction; the algorithm eliminates the potential registration for courses that the student has passed. A list of data input needed for the prediction is defined in TABLE I. The course registration prediction algorithm is defined as in Algorithm 1.

Prepare Camera-Ready paper in full size format, on A4 size or 8 1/2" x 11" (215.9 mm x 279.4 mm) paper.

TABLE I
DATA INPUT FOR PRE_PROCESSING ALGORITHM

Data		Description
Student data	S	Active student for current semester
Student course result	C_PASS_S	Student result from previous semesters, list of al passed courses
Course offer	C_OFFER	List of courses to be allocated for current semester
Course guidebook	C_ALL_S	Students' course schedule plan for each year of study and program.
	C_PRE_C	Pre-requisite courses for advance courses

Algorithm 1: Potential course registration

FOR each active student S as s_i
    Potential courses for s_i, $C\_POTENTIAL_{s_i} = C\_ALL_{s_i} - C\_PASS_{s_i}$
    FOR each $C\_POTENTIAL_{s_i}$ as C_j
        p = true
        IF $C_j \notin C\_OFFER$
            p = false
        ELSE
            $C\_PRE_{c_j}$ = prerequisite course of C_j
            IF $C\_PRE_{c_j} \cap C\_POTENTIAL_{s_i} \neq \emptyset$
                p = false
            ENDIF
        ENDIF
        IF p = true
            Set $s_i \in S_{c_j}$
        ENDIF
    ENDFOR
ENDFOR

The output of data pre-processing on course registration prediction is a list of potential student for all offering courses. Consequently, the size for each course is predicted based on the number of potential students. The accuracy on size of course will ease the venue allocation for each course to maximise the venue utilisation.

B. Identify Critical Students and Courses

Data pre-processing also identifies the lists of critical students and critical courses. There are few steps in extracting these lists:

Step 1 Getting list of potential students for all courses
Step 2 Filter and select critical students who left a few number of potential course registrations which can be taken in one semester
Step 3 List of critical courses can be compiled by the union of potential courses for all critical students

The analysis on critical student data helps to make better decision and planning for timetabling. Originally, different sets of courses are offered for different semesters, either semester 1 or semester 2. Through this analysis, it helps to identify special course to offer so that critical students able to complete their study in shortest time. Better consideration on these critical students helps them graduate on time which saves them a lot of time and money. Thus, it takes care of university reputations.

Two-Stage Heuristic

A two-stage heuristic is proposed to solve this lecture timetabling case study. This is due to its simplicity to cope with different hard and soft constraints in two stages.

In the first stage, *course grouping stage* clusters all the courses into a number of course groups with a simple rule that is courses in a course group must contain different set of potential students without any intersection. This rule is to ensure that there would not be any timetable conflict existed during timeslot-venue allocation stage.

In the second stage, *timeslot-venue allocation stage* allocates timeslot for each course group generated at prior stage. After that, each course in a course group is assigned to a particular venue at the allocated timeslot for its course group. It means that all courses in a course group would have lecture during the allocated timeslot simultaneously at different venues.

Result and Analysis

Three real semester datasets were collected from FCSIT for testing with the simulator. The details are shown in TABLE II.

TABLE II
DATASET GENERAL INFORMATION

Dataset	**2011/12-1**	**2014/15-1**	**2014/15-2**
Students	520	1279	1236
Courses	49	56	56
Predicted Course Registration	2989	6128	4878
Student Clusters	25	35	35

A. Overall Solutions Comparison

TABLE III
OVERALL SOLUTIONS COMPARISON

Dataset	**Real Solution**	**Two-Stage Heuristic Solution**
2011/12-1	Feasible clash-free solution	Feasible clash-free solution
2014/15-1	Timetable clashes	Clash-free solution (4 unallocated courses 7%)
2014/15-2		Feasible clash-free solution

As shown in TABLE III, two-stage heuristic method produced clash-free timetable solutions for all datasets but there are four unallocated courses for 2014/15-1. Dataset 2014/15-2 has same amount of course offer and almost same amount of student as compared to dataset 2014/15-1. However, two-stage heuristic solution has 4 unallocated courses for 2014/15-1 but manage to allocate all courses in dataset 2014/15-2. This is because the number of predicted course registrations is 20.4% lesser than in dataset 2014/15-1 (refer to TABLE II). Meanwhile, same sets of timetabling problem solved by FCSIT produced timetable solutions with clashes by the commercial timetabling software. Besides, there are issues with venue capacity problem where the number of students for some courses exceeded the maximum capacity of allocated venue. Perhaps the actual registration for some courses outnumbered the estimation by curriculum-based with inconsideration of "repeaters". Usually, the first draft of timetable solution design by FCSIT timetable planners will be used as reference for course registration. Timetable re-schedule is needed if receiving any timetable clashes report. However, first draft manual method timetable of dataset 2011/12-1 is feasible without the necessity of re-schedule. The reason behind is course grouping concept has been applied in timetabling for that semester to divide courses with no student clashes into course groups before proceed with manual timeslot and venue allocation.

B. Data Pre-processing: Course Registration Prediction

TABLE IV
PREDICTED VS. ACTUAL REGISTRATION
BY COURSE SIZE VENUE DEMAND

Dataset 2014/15-2	**Actual**	**Prediction**
DK (540 capacity)	10	10
BS (150 capacity)	0	1
TMM (120 capacity)	12	14
TR (30 capacity)	34	31
TOTAL	**56**	**56**

TABLE IV compares difference in varies venue demand based on actual course registration and our predicted course registration. There is slightly higher demand for larger size venue for the prediction course size as the prediction is based on the assumption that each student is enrolling the maximum number of potential courses. Generally, the predicted course registration is close to the actual course registration.

C. Venue Utilisation

Summary of timeslot-venue allocation for small dataset 2011/12-1 is shown in TABLE V. Venue demand for most courses is either TMM or TR. 98% of venue demand are fulfilled with their matching type. There are 3 timeslots of TMM demand are allocated with BS venue due to course group structure. In this case, two courses with size demand for TMM are allocated into the same course group. There is only one TMM available at a time and thus, one course will be allocated to the next available size venue which is BS.

TABLE V
TIMESLOT-VENUE ALLOCATION FOR 2011/12-1

Venue Demand	**Timeslot Available**	**Two-Stage Heuristic Solution**		**Differ**	
		Predicted	Allocated		
DK (540👤)	34	20	20	0	0%
BS (150👤)	37	3	6	+3	100%
TMM (120👤)	48	25	22	-3	12%
TR (30👤)	480	93	93	0	0%
TOTAL Demand		**141 timeslots**			

TABLE VI and TABLE VII show the comparison of venue allocation between real solution and two-stage heuristic for dataset 2014/15-1.

TABLE VI
TIMESLOT-VENUE ALLOCATION FOR 2014/15-1 (REAL SOLUTION)

Venue Demand	**Timeslot Available**	**Real Solution**		**Differ**	
		Predicted	Allocated		
DK (540👤)	22	33	18	-15	46%
BS (150👤)	68	12	67	+55	458%
TMM (120👤)	48	47	37	-10	21%
TR (30👤)	480	130	100	-30	23%
TOTAL Demand		**222 timeslots**			

ISBN 978-1-5090-4897-7

TABLE VII
TIMESLOT-VENUE ALLOCATION FOR 2014/15-1 (PROPOSED HEURISTIC)

Venue Demand	**Timeslot Available**	**Two-Stage Heuristic Solution**		**Differ**	
		Predicted	Allocated		
DK (540)	22	18	18	0	0%
BS (150)	68	21	42	+21	100%
TMM (120)	48	53	36	-17	32%
TR (30)	480	101	97	-4	4%
TOTAL Demand		**193 timeslots** (reduced 13.1%)			

From TABLE VI, it is proven that the real solution has violated the hard constraint, which is the venue capacity must fit the allocated course. The predicted DK demand was 33 timeslots but only 18 timeslots being allocated. The other 15 DK timeslots demand has been allocated with BS venue which violated the capacity demand of DK.

Ideally, the predicted demand and the allocated demand should be matched 100%, which the differ percentage is equal to 0%. However, there do exist differ due to unavoidable constraints on other considerations. By comparing the differ percentage in both solutions, two-stage heuristic solution allocate better utilisation on venue resources. Two-stage heuristic even manages to achieve 0% differ for DK demand. For real solution, it can only perform best at 21% differ and this figure can up to 458%. In real solution, some lectures are allocated with outsize venue than demand which contribute to wastage.

Conclusion

A real-life university lecture at Faculty of Computer Science and Information Technology (FCSIT), Universiti Malaysia Sarawak (UNIMAS) investigated, where pre-registration is not a practice. Due to that, course timetable is "forced" to schedule by curriculum-based which lead to a lot of clashes. Therefore, a data pre-processing algorithm is proposed in this study to predict the course registration of each student. After that, a two-stage heuristic is developed to schedule the course timetable.

A comparison study is conducted between the two-stage heuristic timetable solutions generated by simulator and real timetable solutions, which created using either manual method or with commercial timetabling software assistance. Simulator timetable solution outperformed current practices with fast computational time of max 11 minutes with clash-free solution. Real solution with manual method can take one to two weeks while real solutions with software assistance can produce solution in about 1 minute. However, real solutions with software assistance have a lot of issues on venue resources clashes and violated venue maximum capacity constraint. Few round of adjustment are required which took even longer time (about 1 month) and human resources for it.

In term of venue utilisation, simulator timetable solutions illustrate that the utilisation of venue resources been greatly increased. Simulator solution manages to allocate the matched venue demand up to 98% for dataset 2011/2012-1. Besides that, simulator solution also can fulfil the demand for DK differ. On the other hand, the differ figures is much bigger in real solution. This indicates that higher percentages of courses are allocated into different venue as they demand. The differ percentage for real solutions can range from 21% to 458%; meanwhile simulator solution range from 0% to 100%.

Acknowledgement

This work was supported by Zamalah Penyelidikan Naib Canselor (ZPNC) scholarship under Universiti Malaysia Sarawak (UNIMAS) and research funding of RACE/b(3)/1247/2015(03) from Ministry of Education Malaysia.

References

[1] Badoni, R. P., & Gupta, D. K. (2015). A hybrid algorithm for university course timetabling problem. *Innovative Systems Design and Engineering*, 6(2), 6066.

[2] Gora, W., Lach, G., Lübbe, J., Pfeiffer, O., Zorn, E., & Jeschke, S. (2010). Management and Optimal Distribution of Large Student Numbers. *IEEE EDUCON 2010 Conference* (pp. 1891-1896). IEEE.

[3] Carter, M. W., & Laporte, G. (1995). Recent development in practical examination timetabling. *International Conference on the Practice and Theory of Automated Timetabling* (pp. 1-21). Springer Berlin Heidelberg.

[4] Murray, K., & Müller, T. (2007). Real-time Student Sectioning. Proceedings of the 3rd Multidisciplinary International Conference on Scheduling: Theory and Application (MISTA 2007), (pp. 598-600). Paris, France.

[5] Kristiansen, S., Sørensen, M., & Stidsen, T. R. (2011). Elective course planning. *European Journal of Operational Research*, 215(3), 713-720.

[6] Di Gaspero, L., Schaerf, A., & McCollum, B. (2007). The Second International Timetabling Competition (ITC-2007): Curriculum-based Course Timetaling (Track 3). *Technical Report QUB/IEEE/Tech/ITC2007/CurriculumCTT/v1. 0, Queen's University, Belfast, United Kingdom*.

[7] Lewis, R., Paechter, B., & McCollum, B. (2007). Post Enrolment based Course Timetabling: A Description of the Problem Model used for Track Two of the Second International Timetabling Competition. United Kingdom: Cardiff Business School.