

SOLVING ITC2007 TIMETABLE SCHEDULING PROBLEM WITH LATE ACCEPTANCE

HII



0000103249

NAME: LEE KIA WEI

THESIS SUBMITTED IN FULFILMENT OF THE DEGREE OF COMPUTER
SCIENCE (SOFTWARE ENGINEERING)

FACULTY OF COMPUTER SYSTEM AND SOFTWARE ENGINEERING
UNIVERSITI MALAYSIA PAHANG

2014

Abstract

There are many timetabling problem are studied by researchers such as the Toronto dataset, Dataset of University of Melbourne, Dataset of University of Nottingham as well as the dataset from the local University such as Dataset of University Kebangsaan Malaysia(UKM) and Dataset of University Teknologi MARA(UiTM). For this thesis, we are concerning on solving the Second International Timetabling Competition (ITC2007) dataset. The ITC2007 dataset is chosen because it has additional hard constraints and software constraints compared to other datasets. Besides, ITC2007 consists of 12 datasets with different features in the terms of number of exams, number of room and number of timeslot. We are choosing the Late Acceptance Hill Climbing method to solve this dataset. The reason why using Late Acceptance Hill Climbing is due to the Late Acceptance Hill Climbing is an effective and simple method that will ease the implementation.

ABSTRAK

Terdapat banyak masalah jadual waktu yang dikaji oleh penyelidik seperti set data Toronto, dataset Universiti Melbourne, set data Universiti Nottingham dan set data dari Universiti tempatan seperti dataset Universiti Kebangsaan Malaysia (UKM) dan set data Universiti Teknologi MARA (UiTM). Bagi penyelidik ini, kami hendak menyelesaikan masalah system peperiksaan yang menggunakan set data daripada Pertandingan jadual waktu Antarabangsa Kedua (ITC2007). The set data ITC2007 dipilih kerana ia mempunyai kekangan keras dan kekangan perisian tambahan berbanding dengan set data yang lain. Selain itu, ITC2007 terdiri daripada 12 set data dengan ciri-ciri yang berbeza di segi bilangan peperiksaan, bilangan bilik dan beberapa timeslot. Kami memilih kaedah Late Acceptance Hill Climbing untuk menyelesaikan set data ini. Hal ini kerana Late Acceptance Hill Climbing merupakan kaedah yang lebih berkesan dan mudah yang akan memudahkan pelaksanaan.

TABLE OF CONTENTS

	Page
DECLARATION	iv
SUPERVISOR'S DECLARATION	v
ACKNOWLEDGMENTS	vi
ABSTRACT	vii
ABSTRAK	viii
TABLE OF CONTENTS	ix
LIST OF TABLES	x
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xii

Section	Content	Page
1	INTRODUCTION	1
1.1	Background of Study	1
1.2	Problem Statement	2
1.3	Research Objective	2
1.4	Research Scope	3
1.5	Thesis Organization	3
2	LITERATURE REVIEW	4
2.1	Overview of timetabling	4
2.2	University Timetabling problems	4
2.3	Examination Timetabling	6
	2.3.1 Constraints and objective investigated in examination timetabling problems	7
2.4	Capacitated and un-capacitated examination timetabling problem	7
	2.4.1 Dataset of University of Melbourne	8
	2.4.2 Dataset of University of Nottingham	9
	2.4.3 Toronto Dataset	10
	2.4.4 Dataset of University Kebangsaan Malaysia(UKM)	11
	2.4.5 Dataset of University Teknologi MARA(UiTM)	12

2.4.6	Second International Timetabling Competition (ITC2007) dataset	13
2.5	Methodologies applied to the examination timetabling problem	15
2.5.1	Tabu Search(TS)	16
2.5.2	Great Deluge Algorithm (SA)	17
2.5.3	Simulated Annealing (SA)	19
2.5.4	Genetic Algorithm (GA)	20
2.5.5	Particle Swarm Optimization (PSO)	21
2.5.6	Late Acceptance Hill Climbing (LAHC)	22
2.6	Conclusion	23
3	METHODOLOGY	25
3.1	Introduction	25
3.2	ITC 2007 :Examination track	25
3.3	Dataset Formulation	27
3.4	Late Acceptance Hill Climbing (LAHC)	31
3.5	Solving Second International Competition (ITC 2007) Examination Problem with LAHC	33
3.6	Conclusion	34
4	DESIGN & IMPLEMENTATION	35
4.1	Project Implementation	35
4.1.1	Display the information of the ITC2007 dataset	35
4.1.2	Calculate the Basic Information	39
4.1.3	Calculate Conflict Matrix, Matrix Density, Number of Exam Conflicts and Student Conflicts	39
4.1.4	Sorting of Exams with LWD, LD and LE	40
4.2	Improving the Initial Solution (Time table) by Solving ITC2007 Datasets	43
4.2.1	Initiation of Initial Solution Using Pure Graph Heuristic Method	43
4.2.2	Improvement of Initial Solution	45
4.3	Conclusion	48
5	RESULT & DISCUSSION	49
5.1	Discussion of Results	49

5.2	Research Constraints and Suggestions	50
5.3	Conclusion	50
6	CONCLUSION	51
6.1	Introduction	51
6.2	Result Analysis	51
6.3	Future Work	53
	Reference	

LIST OF TABLES

NO	TABLE TITLE	PAGE
2.1	Examples of constraints of the course timetabling problems	5
2.2	Example of constraints for the examination timetabling problems	6
2.3	University of Melbourne datasets	9
2.4	University of Nottingham datasets	9
2.5	Toronto Dataset	10
2.6	University Kebangsaan Malaysia datasets (UKM06=01)	12
2.7	capacity of dataset UKM06-01	12
2.8	University Teknologi Malaysia (UiTM) dataset	13
2.9	International Timetabling competition dataset	13
2.10	Hard constraints of ITC2007	14
2.11	Soft constraints of ITC2007	14
2.12	Summary of datasets	14
5.1	Comparison of Results	

LIST OF FIGURES

NO	FIGURES TITLE	PAGE
2.1	GDA Procedures	17
2.2	Genetic Algorithm Procedure	20
2.3	Particle Swarm Optimization procedure	22
2.4	Late Acceptance Hill Climbing procedure	23
3.1	Algorithm of Late Acceptance Hill Climbing	31
3.2	Pseudocode of Later Acceptance Hill Climbing	32
4.1	Data of Exam	35
4.2	Data of Rooms	36
4.3	Data of Exam period	36
4.4	Data of Period Hard Constraints	37
4.5	Data of Room Hard Constrains	37
4.6	Data of Conflict Matrix	38
4.7	Data of Institutional Weight	38
4.8	Java Coding for Text File Reading	39
4.9	Basic Information	39
4.10	Sample Coding for Calculating the Variables	39
4.11	Largest Weight Degree LWD of Exam	40
4.12	Sample Coding for Calculate LWD	40
4.13	Largest Degree LD of Exam	41
4.14	Sample Codes for Calculate LD	41
4.15	Largest Enrollment LE of Exams	42
4.16	Sample Codes for Calculating LE	42
4.2	Flow chart of Initial solution	44
4.2.1	Flow chart of Initial solution (cont..)	45
4.2.2	Sample coding of moving and swapping of exams	46
4.2.3	Sample Coding to Identify Improvement	47
4.2.4	Flow chart for Late Acceptance Hill Climbing Technique	48

Chapter 1

INTRODUCTION

1.1 Background of Study

Generally, the university examination timetabling problems played an important role which involves allocating examination to timeslots subject to fulfil several constraints problem which are hard constraint and soft constraint. The hard constraint must all be met by the resulting timetable to ensure the feasibility of timetable. As an instance, a student cannot sit for two exams at the same time. Apart from that, there is impossible to meet all soft constraint because it is contradictory. However, it is importance to fulfil as even as possible to ensure the perfection of timetable. As instance, examination student is not allowed to take more than two examinations in a day and the examinations of student are well spread.

There are two versions of examination problem exist which known as capacitated version and the un-capacitated version. The capacitated problem is much nearer to with the real world problem. This is due to the capacities of the room will be the hard constraints that should not be broke. The un-capacitated version of the problem does not take room capacities into consideration like what the capacitated version does. Researchers are focusing on the algorithmic performance to produce a quick and effective solution. Thus, many researchers are able to solve the un-capacitated examination problem.

Since the capacitated examination problem is more close to the real word problem as the room capacities constraint is one of the constraint that taken into consideration but the capacitated problem are still has the less attention from those

researchers. This situation might be due to the lack of the capacitated benchmark dataset. Furthermore, there are many constraints in capacitated problem makes it more difficult to compare to un-capacitated problem. The difficulties when solving the capacitated problem involves the consideration of constraints such as the amount and the size of the room are those constraints that increase the difficulties when solving capacitated problem. The result of survey from Burke et al.'s (1996) shown that 73% of the universities agreed that to schedule an examination timetable is very difficult.

In this thesis, we are using the International Timetabling Competition 2007, ITC2007 dataset. ITC 2007 dataset is a capacitated problem dataset that has been added several new constraints instead of those that are found in scientific literature. We decided to use Late Acceptance Hill Climbing method to solve this problem.

1.2 Problem Statement

There are many researchers concerned about the examination timetabling problems. There are a lots of studies related to un-capacitated dataset can be found in the literature. However, this does not match the real world examination timetabling problem. Toronto dataset is one of the examples of un-capacitated dataset. Nevertheless, there are also a small amount of researchers are choosing to use capacitated dataset which involves dataset from Nottingham and Melbourne. The individual room capacity is often a matter to be considered. However so this capacitated dataset are not really resemble the real world problem. Therefore, there might be some differences between research and practical.

ITC2007 examination dataset is a real word capacitated examination dataset. Our concern is to solve this ITC2007 examination dataset. ITC2007 contain numerous constraint that different from the dataset constraint that we seen in the literature.

1.3 Research Objective

In this thesis, we concentrate on three main objectives. First, to study the Second International Timetabling Competition (ITC2007) examination tracks. Then

we plan to solve this problem by using a suitable method which is Late Acceptance Hill Climbing. Ultimately, to validate and verify the implemented algorithm which is the Late Acceptance Hill Climbing satisfied all the hard constraints and fulfilled the soft constraints as even as possible.

1.4 Research scopes

Our thesis is concerning on solving the Second International Competition (ITC 2007) dataset problems. The ITC2007 is a capacitated dataset which will put hard constraints (the size of the room and the number of the rooms) under consideration. Thus, to construct a schedule for dataset, we will use the Late Acceptance Hill Climbing method.

1.5 Research Organization

This research will cover 6 chapters. The first chapter is the research introduction. The second chapter is covering the literature review. The third chapter will discuss the methodology based on what we found on Chapter 2. Chapter 4 and Chapter 5 will cover the description of design and how the system will be implemented. Chapter 6 will show the results produced and as well as the discussion. The last chapter is the research conclusion.

Chapter 2

LITERATURE REVIEW

2.1 Overview of Timetabling

Timetable is a list of event in a schedule with specific arrangement according to the times and places. Timetabling can be categorized into several categories, personal, transportation as well as educational timetabling (Qu et al.2009). Every timetabling has its own problems as well as different requirement and constrains to be fulfilled.

Based on our discussion in chapter 1, the timetabling problem constrains can be put into two main types which are soft constrains and hard constrains. The hard constrains are the basic constrains because all of the hard constrains are required to be fulfilled. If there is one hard constrain is omitted, the time table will never work. Thus, it is compulsory to satisfy all hard constraint and it should not be broke. One of the examples of hard constrains will be no student will be allowed to sit two exams on the same time. On the other hand, even though it is unnecessary to fulfil all the soft constraints. However, it is important to compliance the soft constraint as far as possible to achieve the perfection of timetable. The lower the violations of soft constraints, the higher quality of timetable will be produced. The soft constraints such as spreading the exams as even as student point of view.

2.2 University timetabling problems

The University timetabling problems are classified into two main types either are examination or course timetabling problems. Both problems are sharing the same core problem with same characteristic. The main concern is to prevent students who attend two examinations concurrently. Nevertheless, this examination timetabling problems and course timetabling problems are also having the significant differences which can be easily differentiate. One of the examples is the hard and soft constraints

that are required to be compliant. Table 2.1 shows the examples of the hard and soft constraints.

Beside both hard and soft constraints, another different is the style to construct the course and examination. The construction could be categorized into modelling, process environment and scheduling instances. Based on the study of course timetabling by McCollum (2007) and Burke et al (1996), the timetable is produced differently by each school independently. Furthermore, examination timetabling problems is produced according to the number of courses that have been offered and the student who registered the particular course. For the construction of course timetable, the number of student that will register the course will take into main consideration (McCollum, 2007). For the scheduling instances, the examination timetables are generated according to the courses offered. Apart from that, for the course timetable, we have to schedule more than only the offered courses, but we have to consider the lab and lecturer tutorial session (McCollum, 2007).

Although the examination and course timetable has distinct different, however the degree of complexity directly depend on the degree of freedom of a student in choosing their course timetable which known as open registration (Laporte and Desroches, 1984). The higher the degree of freedom is given to student, the higher the difficulties to produce a feasible solution.

Table 2.1 Example of hard and soft constraints for timetabling problems (Abdullah, 2006)

Hard constraints
1. Students or lecturers are not allowed to attend more than one examination simultaneously.
2. Only one exam in a timeslots assigned in a classroom.
3. The participant number should not exceed the capacity of rooms.
4. The assigned room should fulfil the features required by the exam.

Soft constraints

5. Student should taking more than one exam in a day.
6. The course should be scheduled until the last timeslots of a day.

2.3 Examination timetabling

According to Schaerf(1999), examination timetable is an essential element for an educational institution. It brings large effect to the student, lecturer and administrators. The aim of examination timetabling as assigning course into a number of rooms as well as timeslot in order to satisfied the hard and soft constraints. Table 2.2 displays the examination timetabling constraints.

Table2.2 Example of hard and soft constraints (Qu et al.2009)

Hard constraints

1. The similar resources (students) are not allowed to be assigned to the exam simultaneously.
2. The resources (size of the room, number of room) of the examinations should be sufficient.

Soft constraints

3. Exams should be spread evenly.
4. Every exam should be scheduled and the heaviest exam should be scheduled as soon as possible.
5. Limiting the student number in each timeslot.
6. Assign the exams which have the same duration in the same room.
7. The time requirement (examinations should not assigned in certain timeslot).
8. Exams should be consecutively allocated.
9. Sequence of exams is required to be fulfilled.

10. Exams which have conflicts in same day should be located close to each other.

2.3.1 Constraints and objectives investigated in examination timetabling problems.

Different institution has their own difference constrains needed to achieve. There are many studies of examination timetabling are able to be found in the literature. In addition, the parties which are influenced by the timetable might have different preference to produce high quality timetable. For the instance, the further the examination spread, the more time will have for students to have enough time to do revision. Thus, as an administrator, he or she takes the responsibility to construct an examination schedule and should ensure that there is no student is allowed to take two exams in one day. The discussion below will discuss about the common constraints that usually found in timetabling problems.

There are several common used datasets in the examination timetabling research community which are the Nottingham, Melbourne dataset and Toronto dataset. Toronto dataset is the dataset that the most popular for researchers among these three types of dataset. In the year 2008, McCollum et al (2008) introduced the Second International Competition (ITC2007) dataset. This dataset are much closer to real world problem compare to benchmark problems. Besides, the other examination dataset are the Universiti Kebangsaan Malaysia, UKM dataset, and Universiti Teknologi Mara, UiTM dataset.

2.4 The un-capacitated and capacitated examination timetabling problems

The studies of un-capacitated problems can be found in literature. The un-capacitated problems are usually concentrated on the algorithmic performance in the purpose of producing effective and quick solution. Most of the researchers are more likely to work on a simplified version of examination problems instead of dealing with all the aspects of examination problems. There are not many common hard constraints that are addressed by the survey paper. As the instance, the exams size

should not exceed the capacity of the room; students are not allowed to take two exams concurrently and etc. For soft constraints, the common constraints that used are spreading of exams as even as possible, but not in the consecutive number of timeslot or day.

For the capacitated problem, the capacity of the room is the constraints that are included and this makes capacitated problem more realistic. However, the lack of benchmark dataset causes the capacitated problem has got less attention from the researchers. Moreover, the capacitated problems are evidently proved that is more difficult to solve. This statement has been proven by the survey paper done by Burke et al., (1996a). Capacitated problem requires more thorough data as it had to include the capacity of room and also the data of less complexity problem (eg. student and list of exams). Based on the research of McCollum (2007), to gather the extra information are difficult. In 1996, Burke et al. said the difficult of allocating examinations are due to lack of available examination hall and the splitting exams between more than one rooms. This might affect other constraints (ie. exam is split into different site or taking into account between room).

The dataset more realistic due to its concern in capacitated problem. Burke, Newall and Weare, (1996) has modified the dataset. (i.e Toronto dataset), this includes the complete capacity as if all examinations are held in one big room. This shows the simplicity of the timetabling problem as we would have to take the individual room capacity into consideration.

2.4.1 Datasets of University of Melbourne

This dataset of University of Melbourne consists of two different dataset with two timeslots on weekday and it was first introduced by Merlot et al., (2003). For every dataset, it has different capacities on difference timeslot. Furthermore, the period exclusive constrains of this dataset which the exams are addressed to a particular timeslot or can only takes place in a limited set of timeslot. This dataset has the aim to minimise the second-order conflicts on the same day or overnight. According to the investigation of Merlot et al., (2003), Cote, Wong and Saboun, (2005), the bi-objective evolutionary algorithm dataset is where tabu search (TS) and

variable neighbourhood decent (VND) were utilized. The examination dataset can be retrieved from their university website. The table below is displaying the Melbourne dataset.

Table 2.3 Melbourne dataset

Problem Instance	Exams	Students	Enrollments	Timeslots
I	521	20656	62248	23
II	526	19816	60637	31

2.4.2 Dataset of University of Nottingham

In the year 1996, Burke, Newall and Weare introduced the Nottingham dataset. This dataset includes 3 timeslot per day on weekend out of the total of 23 timeslots. The main constrains of this dataset is the capacity and it is not clashing with each other. Similar to the Melbourne dataset, it helps to prevent the second-order conflict on the same day. In the year of 1999, Burke and Newall were using the method of graph heuristic to minimize the number of second conflict on the same day. Apart from that, Merlot was also applied the same method that had been used by Burke and Newall on the Nottingham dataset in year 2003. Another year later, Burke et al. was using the great deluge algorithm (GDA) to solve the same problem. The table below display the University of Nottingham dataset.

Table 2.4 University of Nottingham datasets

Exams	Student	Enrollment	Conflict Density	Timeslots capacity
-------	---------	------------	------------------	--------------------

800	7896	34265	0.03(3%)	23	1550
-----	------	-------	----------	----	------

2.4.3 Toronto dataset

In the year of 1996, Carte, Laporte and Lee introduced the Toronto dataset which have totally 13 tracks of examination dataset which is mimicking the real-world problem. Among this 13 dataset, 3 are originated from Canadian high schools, 5 are originated from Canadian institution, 1 each from the University of King Fahd, University of Dhahram, Purdue, Indiana (Carter, Laporte and Lee, 1996) and London School of Economics. In 2001, Di Gaspero and Schaerf were applying the tabu search method in solving this dataset. They focus on the conflict. The hard constraint such as if any student register two examinations at the same timeslot and the soft constraints such as a student is required to register two exams consecutively. Burke, Newall and Weare, (1996) started to concern about the optimization of room capacity per timeslot and second-order conflict of the same day constraints. 2 year after Di Gaspero and Schaerf, Merlot et al. (2003) was using several methods to solve the Toronto dataset by using hybridize of constraint programming, simulated annealing and hill Climbing. The purpose is to prevent the required timeslots, spreading the conflict examinations within limited number of timeslots as well as to minimize second-order conflict of the same day. The Toronto datasets. Qu et al. (2009) classified the problem into I and II to help genuine comparison between scientific communities. Table 2.5 displays the Toronto dataset.

Table 2.5 Toronto Dataset (Qu et al., 2009)

Problem Instance	Exams	Students	Enrollments	Conflict Density	TimeSlots
ear83 I	190	1125	8109	0.27	24
ear83 II	189	1108	8014	0.27	24
yor83 I	181	941	6034	0.29	21
yor83 II	180	919	6012	0.29	21

sta83 I	139	611	5751	0.14	13
sta 83 II	138	549	5689	0.14	13
car91 I	682	16925	56877	0.13	35
car91 II	682	16925	56242/56877	0.13	35
lse91	381	2726	10918	0.06	18
car92 I	543	18419	55522	0.14	32
car92 II	543	18149	55189/55522	0.14	32
rye92	486	11483	45051	0.07	23
hec92 I	81	2823	10632	0.42	18
hec92 II	80	2823	10625	0.42	18
tre92	261	4360	14901	0.18	23
uta92 I	622	21266	58979	0.13	35
uta 92 II	638	21329	59144	0.13	35
ute92	184	2749	11793	0.08	10
kfu93	461	5349	25113	0.03	42
pur93 I	2419	30029	120681	0.03	42
pur 93 II	2419	30029	120686/120681	0.03	42

2.4.4 Dataset of Universiti Kebangsaan Malaysia (UKM)

UKM datasets are one of the capacitated dataset (Ayob et al., (2007)). This datasets required to schedule all the examinations. The constrains are not allow student to take more than one exam simultaneously and taking three exams in a low on a same day. The room exclusive constraints which are the specific room for exams must be satisfied and the students must be assigned in the same room for those who sit for consecutive exams. The purpose of this dataset is spreading the exams evenly and prevent the students take the exams consecutively at the same day. Table below displays the dataset of UKM.

Table 2.6 University Kebangsaan Malaysia datasets (UKM06-01)

Exams	Student	Enrollment	Timeslots	capacity
818	14047	75857	42	1550

Table 2.7 Room capacity of dataset UKM06-01

Room	Room Capacity
LobiA(DECTAR)	70
LobiB(DECTAR)	70
Pseni(DECTAR)	152
LobiUtama(DECTAR)	270
Dewan(DECTAR)	610
DGemilang	610
DPBestari	850

2.4.5 Datasets of Universiti Teknologi Mara (UiTM)

In the year 2004, Kendall and Hussin introduced the dataset of UiTM. These dataset are similar to dataset of UKM which requires scheduling all the exams. The constraints involved are the coincidence and the first-order conflict (ie. Examinations that are necessary to carry out together should be addressed in the same timeslot.) The purpose of this dataset is spreading the exam as even as possible. The spreading of dataset is based on the calculation of the proximity value (Carter, Laporte and Lee, (1996)) and the exams that are assigned at the weekend will get a penalty. Table 2.8 displays the examination dataset of UiTM.

Table 2.8 University Teknologi Malaysia (UiTM) dataset

Exams	Students	Enrolments	Timeslots
2063	84675	357761	40

2.4.6 The dataset of Second International Timetabling Competition (ITC2007)

ITC2007 dataset are divided into examination and course timetabling. However, our research will only concentrate on the exams dataset only. The aim of ITC2007 is for researchers to perform their searching technique on the real-world timetabling problem. The purpose of those constraints is to minimize second-order conflict at the same day, minimize mixed duration of exams within a timeslot, minimize the usage of a particular timeslots or rooms and schedule heavy exams as early as possible. The examination tracks are able to obtain in McCollum et al., (2008). There are several researchers did their investigation on this dataset, they are McCollum et al.,(2008) and Gogos AleFragis and Housos, (2008). McCollum was using iterated forward search, great deluge algorithm and hill climbing while AleFragis and Housos were using a multistage approach that which are mathematical programming, simulated annealing and GRASP. In 2009, McCollum et al. used two-phases approach with adaptive heuristics ordering as the constructive phase and used extended great deluge algorithm (GDA) to improve the solution. Table below display the dataset of ITC2007 examination tracks, the constraints and the comparison between ITC2007 dataset with others.

Table 2.9 International Timetabling competition dataset.

Instance	Conflict Density (%)	Exams	Students	Periods	Rooms	Period HC	Room HC
Exam 1	5.05	607	7891	54	7	12	0
Exam 2	1.17	870	12743	40	49	12	2
Exam 3	2.62	934	16439	36	48	170	15
Exam 4	15	273	5045	21	1	40	0

Exam 5	0.87	1018	9253	42	3	27	0
Exam 6	6.16	242	7909	16	8	23	0
Exam 7	1.93	1096	14676	80	15	28	0
Exam 8	4.55	598	7718	80	8	20	1
Exam 9	7.48	169	655	25	3	10	0
Exam 10	4.97	214	1577	32	48	58	0
Exam 11	2.62	934	16439	26	40	170	15
Exam 12	18.45	78	1653	12	50	9	7

Table 2.10 ITC 2007 Hard Constraints

Hard Constraints	
H1	Students are not allowed to register more than one examination at concurrently
H2	The length of exam should not violate the length of timeslot.
H3	The capacity of exams should not exceed the capacity of room.
H4	Schedule exam according to specified room .e.g. Exam I should address to Room 10.
H5	The examination must be respect to the sequence or ordering e.g. schedule Exam 1 followed by Exam 2

Table 2.11 ITC2007 Soft Constraints.

Soft Constraints	
S1	<i>2 examinations per day</i>
S2	<i>Two consecutive examinations: minimize students taking examinations consecutively at the same day.</i>
S3	<i>Mixed duration: minimize number of examinations with different duration that are addressed into the same room .</i>
S4	<i>Spreading of exams: every examination should be spread as evenly as possible throughout the exam period.</i>
S5	<i>Larger examination schedule late in the timetable: minimise the number of large exams assigned into the late timeslot of the timetable.</i>
S6	<i>Room penalty: minimise the number of exams scheduled in room with penalty.</i>
S7	<i>Period penalty: minimise the number of exams addressed in timeslot with penalty.</i>

Table 2.12 Summary of datasets

Constraints		Ump	UKM	Toronto	Melbourne	Nottingham	IT0
Ex	Clash free	H	H	H	H	H	
	Scheduled all exams	H	H	-	S	S	

	Exam preference - Specified arrangement: <i>sa</i> - Specified room: <i>sr</i> - Large exam schedule first: <i>lf</i> - Restriction on exam in particular timeslot: <i>rt</i> - Scheduled combined exam in the same timeslots: <i>ct</i>	-	-	-	-	-	H (se & S (lf)
	Consecutive exam - Two exam in a row: <i>2r</i> - Two exam in a day: <i>2d</i> - Two exam in a row overnight: <i>2n</i> - Three exam in a day: <i>3d</i>	-	-	-	-	S (2d & 2n)	S (2 & 2c)
Timeslot Related	Timeslot preference - Minimise/avoid usage : <i>tu</i>	=	=	=	=	=	S (t
	Timeslot length - Mixed duration of exams in one timeslot: <i>mt</i>	-	-	-	-	-	H & S (mt)
	Spreading - Specified spread: <i>ss</i>	S	S	S	S	H(ss)	S(ss)
Rooms related	Room distance	S	-	-	-	-	-
	No sharing of room with other exams - For specified exam only: <i>se</i>	H	H (se)	-	-	-	-
	Room Preference - Consecutive exam scheduled in the same room: <i>cr</i> - Minimise/ avoid usage: <i>ru</i> - Specified room: <i>sr</i>	-	H	-	-	-	H (c & (ru)
	Split exam into different rooms - Same building only: <i>sb</i> - As close as possible: <i>cp</i>	H(sb) & S(cp)	-	-	-	-	-
	Capacity - Total seats: <i>ts</i> - Individual room: <i>ir</i>	H (ir)	H(ts & ir)	H(ts)	H(ts)	H(ir)	H(t

H =hard constraint; S =Soft constraint; shaded cell = constraint not considered.

2.5 Methodologies applied to the examination timetabling problem

The past 20 years, there are plenty amount of algorithms had been introduced, such as the meta heuristic, graph heuristic, constraints, hybridizations based methods and many else in constructing a timetable.

2.5.1 Tabu Search (TS)

In 1986, Glover introduced the Tabu search method. It is a metaheuristic search method. The way of working is similar with hill climbing but it fixed the weakness of hill climbing which Tabu search can overcome local optima which is the weakness of hill climbing.

Tabu Search explores a subnet $N'(s)$ of the neighbourhood $N(s)$ repeatedly also known as current solution. During the comparison between $N'(s)$ and $N(s)$, the lower value of the neighbourhood solution will be accepted although its value maybe worst than current solution. The basic concept of Tabu Search is pursuing the search for a local optimum is occurred due to allowing non-improving moves. However, choosing a solution from a subnet, $N'(s)$ will causes a cycling. A memory know as tabu list will be used to hold recent history of the search which is the prevention of the search stuck in a local optima. The moves are an iterative procedure. The aspiration criterion which is a mechanism that used to make a solution of tabu free if the return in good quality solution as the results.

In 2001, Di Gaspero implemented a adaptive of tabu search into hill climbing methods. The implementation helps to optimize the objective function (recolour), perturbing the current solution (shake) or achieving higher improvement (kick). The algorithms are terminated when there is a convergence, and will continue with a kick. Wilke and Ostler solved the school timetabling problem by applying the tabu search in 2008. The comparison between Genetic Algorithm (GA), Simulated annealing (SA), and brach & bound showed that the tabu search use the shortest time to solve the solution while simulated annealing produce the best results.

Abdullah, Turabieh and McCollum (2009) were the other researchers that used other methods to hybridise tabu search. They used memetic algorithm. The neighbourhood structure are holded in tabu list because it is unable to produce better solution while for the neighbourhood structure that shows improvement are used until it reaches a convergence. The method could give a better quality of solution for four of the Toronto dataset.

2.5.2 Great Deluge Algorithm (GDA)

Dueck was the founder of GDA. It is almost similar to the simulated annealing and hill climbing algorithm. Instead of using a temperature, GDA is using an upper boundary as the level of acceptance. The initial boundary of the algorithm and initial solution quality are equivalence. The worse solution will only be accepted if the cost value is lower than the boundary. The boundary is lowered in each iterations according to the predetermined rate. GDA is using only one parameter setting which is an advantages compared to SA (among others), since the effectiveness of the meta-heuristic technique is depend on parameter tuning (Petrovic and Burke, 2004). Figure 2.1 displays the great deluge algorithm procedure.

Figure 2.1 shows the GDA procedures

Choose an initial configuration

Choose the “rain speed” $UP > 0$

Choose the initial WATER-LEVEL > 0

Opt: choose a new configuration which is a stochastic small perturbation of the old configuration

 Compute $E := \text{quality}(\text{new configuration})$

 If $E > \text{WATER_LEVEL}$ then

 Old configuration := new configuration

 WATER_LEVEL := water_level + up

 If a long time no increase in quality or too many iteration

 Then stop

 Goto Opt

In the year of 1993, Dueck solved the travelling salesman problem with GDA. The length of the current tour is divided by 5000 or fixed decay rate of 0.01. Therefore, the decay rate will be difference according to the boundary and the length. Based on the investigation did by Burke and Newall (2003), GDA able to produce better quality solutions. The decay rate was computerised as: