A study of hyper-heuristics for examination timetabling

Ender Özcan · Anas Elhag · Viral Shah

Received: date / Accepted: date

1 Introduction

Examination timetabling is both a difficult and time consuming task, faced by many educational institutions worldwide [5]. The main objective is to assign periods within a specified examination timeframe and rooms to exams, whilst satisfying a range of constraints. There are two common constraint categories: *hard* and *soft*. It is imperative that all hard constraints are satisfied in a given solution, which is then referred to as a *feasible* solution. For example, students must not sit two or more exams simultaneously in the same period. Soft constraints, on the other hand, represent preferences that are not essential but should be satisfied as much as possible. For example, a student should not sit two exams in two consecutive periods on the same day. Once a feasible solution is obtained, the degree to which the soft constraints are violated is used to evaluate the quality of a timetable.

Most of the solutions to examination timetabling problems have been developed due to a need at an educational institution. Hence, different types of examination timetabling problems can be found in the literature which are solved using different types of methodologies. This could be considered as richness, but there is a downside that is comparison of the approaches becomes extremely difficult. The state-of-the-art for any problem is of interest to both practitioners and researcher. A recent competition on examination timetabling was arranged as a part of ITC2007[1]. The instances used in this competition reflects the real world examination timetabling complexities. The winner of the

E. Özcan, A. Elhag and V. Shah
University of Nottingham, School of Computer Science
Jubilee Campus, Nottingham NG8 1BB UK
Tel.: +44 (115) 95 15544
Fax: +44 (115) 846 7877
E-mail: {exo, axe}@cs.nott.ac.uk and viralshahkach@hotmail.com

[1] http://www.cs.qub.ac.uk/itc2007/

examination timetabling competition is a hybrid multistage approach which is described in [4].

Hyper-heuristics are methodologies that perform search via generation or selection of heuristics in problem solving [3]. A goal in hyper-heuristic research is designing methodologies which are capable of solving problem instances having diverse properties automatically without requiring any parameter tuning. There are a few benchmarks for examination timetabling. The most commonly used one is Toronto benchmark. The performance of hyper-heuristics have been investigated on Toronto and Yeditepe problem instances [1, 2] as well as ITC2007 instances, which includes instances for Examination Timetabling and Course Timetabling. In this study, we present performance analysis of some selection hyper-heuristics on the Examination Timetabling instances of ITC2007.

2 Experimental Results

A subset of ITC2007 instances are used during the experiments. The characteristics of each benchmark instance are summarised in Table 1. An initial timetable is constructed, firstly assigning examinations with room hard constraints to rooms and periods with just enough capacities and lengths, followed by assigning examinations with period hard constraints in a similar fashion. Finally, a weighted graph is used to determine the order in which to timetable the remaining exams. If the resulting timetable is infeasible, it is reset and the entire process of timetabling starts again, first considering exams with room hard constraints and so on. Then the solution in hand is iteratively improved using a selection hyper-heuristic which perturbs this solution generating a new one using a chosen low level heuristic and then decides whether to accept or reject the new solution. Different combinations of heuristic selection {Simple Random (SR), Greedy (GR), Reinforcement Learning (RL)} and acceptance {Improving Only (IO), Improving and Equal (IE), Great Deluge (GD)} methods are used as hyper-heuristics during the experiments. Six different perturbative low level heuristics were implemented. The main objective of these low level heuristics is to make slight modifications on the current timetable, in an attempt to lower the soft constraint violations, such as rescheduling of rooms, or swapping exams.

Each experiment is repeated 50 times and a run is terminated after 500 seconds complying with the ITC2007 competition rules. The experiments are carried out on a 2.83GHz Intel Core 2 Duo E8300 XP machine with a memory of 3.23GB. The results are provided in Figure 2. Feasible solutions are obtained for almost all problem instances, except for Exam4. In the overall, Reinforcement Learning performs better than the rest of the heuristic selection methods, while as an acceptance method, Great Deluge is better than the others. Table 3 shows a comparison between our approach and the approaches of the winners of the competition.

Table 1 The characteristics of the ITC2007 examination timetabling problem instances.

Problem	No. of exams	No. of students	No. of rooms	No. of time-slots	Conflict density
Exam1	607	7891	7	54	5.05
Exam2	870	12743	49	40	1.17
Exam3	934	16439	48	36	2.62
Exam4	273	5045	1	21	15.00
Exam5	1018	9253	3	42	0.87
Exam6	242	7909	8	16	6.16
Exam7	1096	14676	15	80	1.93
Exam8	598	7718	8	80	4.55

The details of the hyper-heuristic approach and more results using additional hyper-heuristics will be provided at the conference.

References

1. Bilgin, B., Özcan, E., Korkmaz, E.E.: An experimental study on hyper-heuristics and exam timetabling. In: Practice and Theory of Automated Timetabling VI, *Lecture Notes in Computer Science*, vol. 3867, pp. 394–412. Springer (2007)
2. Burke, E., Kendall, G., Misir, M., Özcan, E.: Monte carlo hyper-heuristics for examination timetabling. Annals of Operations Research pp. 1–18 (2010)
3. Chakhlevitch, K., Cowling, P.: Hyperheuristics: Recent developments. In: C. Cotta, M. Sevaux, K. Sirensen (eds.) Adaptive and Multilevel Metaheuristics, *Studies in Computational Intelligence*, vol. 136, pp. 3–29. Springer Berlin Heidelberg (2008)
4. Muller, T.: Itc2007 solver description: A hybrid approach. Annals of Operations Research **172**(1), 429–446 (2009)
5. Qu, R., Burke, E.K., McCollum, B., Merlot, L., Lee, S.: A survey of search methodologies and automated system development for examination timetabling. Journal of Scheduling **12**(1), 55–89 (2009)

Table 2 Soft constraints score for the datasets.

Hyper-heuristics	Exam 1			Exam 2		
	min	μ	σ	min	μ	σ
RL-EAI	8685	8879.14	774.74	**778**	**800.3**	**32.9547**
RL-I	9608	9858.5	685.562	790	811.44	30.7639
RL-GD	9460	9636.12	770.975	**778**	**800.3**	**32.9547**
SRP-EAI	**8584**	**8825.16**	**851.734**	814	845.82	42.0096
SRP-I	9017	9279.02	902.688	807	845.98	26.6906
SRP-GD	9142	9477.22	842.323	789	826.56	46.194
G-EAI	9178	9387.08	1088.85	799	827.96	44.3133
G-I	9179	9428.28	1150.74	783	812.54	42.8858
G-GD	9178	9387.08	1088.85	787	813.4	50.3648
	Exam 3			Exam 4		
	min	μ	σ	min	μ	σ
RL-EAI	32662	35409.5	3574.89	infeasible	n/a	n/a
RL-I	33240	35431.6	2780.72	infeasible	n/a	n/a
RL-GD	31260	34259.4	3064.11	infeasible	n/a	n/a
SRP-EAI	35210	37026.9	2941.63	infeasible	n/a	n/a
SRP-I	34386	39293.4	3101.42	infeasible	n/a	n/a
SRP-GD	31493	34052.3	3657.36	infeasible	n/a	n/a
G-EAI	34071	36090.1	1269.97	infeasible	n/a	n/a
G-I	33574	35340.7	1094.84	infeasible	n/a	n/a
G-GD	**31227**	**32974.6**	**1901.22**	infeasible	n/a	n/a
	Exam 5			Exam 6		
	min	μ	σ	min	μ	σ
RL-EAI	**7541**	**7588.2**	**101.002**	30415	31634.1	2464.28
RL-I	**7541**	**7588.2**	**101.002**	30625	30640.1	57.1829
RL-GD	**7541**	**7588.2**	**101.002**	**29695**	**30126.9**	**1255.71**
SRP-EAI	7677	7726.98	100.82	33775	34086.7	723.078
SRP-I	7677	7726.98	100.82	37485	37503.1	69.5532
SRP-GD	7772	7815.56	108.451	38175	38283.9	88.3193
G-EAI	7658	7662.6	24.6792	37900	38855.9	152.445
G-I	7658	7662.6	24.6792	37900	38855.9	152.445
G-GD	7658	7662.6	24.6792	37900	38855.9	152.445
	Exam 7			Exam 8		
	min	μ	σ	min	μ	σ
RL-EAI	**15116**	**15539**	**701.098**	21678	27446.3	4199.8
RL-I	16722	16912	579.296	20978	21178.4	253.238
RL-GD	15178	15549.1	631.938	23389	23583.4	405.198
SRP-EAI	15291	15492.3	372.928	21812	22207.3	660.057
SRP-I	16941	17185.1	363.882	21522	22056.6	809.114
SRP-GD	15660	16010.3	476.773	22657	23273.3	810.811
G-EAI	16622	16739.8	322.287	**20168**	**20562.6**	**710.712**
G-I	16796	16863	216.641	20236	20990.9	806.43
G-GD	16622	16739.8	322.287	**20168**	**20562.6**	**710.712**

Table 3 Comparison between our results and the results of the top winners of the competition.

	Exam1		Exam2		Exam3	
Ranking	Winner	Score	Winner	Score	Winner	Score
1^{st}	Muller	4370	Muller	400	Muller	10049
2^{nd}	Gogos	5905	De Smet	623	Gogos	13771
3^{rd}	De Smet	6670	**Özcan**	**778**	Pillay	15917
4^{th}	Atsuta	8006	Gogos	1008	Atsuta	17669
5^{th}	**Özcan**	**8584**	Pillay	2886	**Özcan**	**31227**
6^{th}	Pillay	12035	Atsuta	3470	De Smet	Infeasible
	Exam4		Exam5		Exam6	
	Winner	Score	Winner	Score	Winner	Score
1^{st}	Muller	18141	Muller	2988	Muller	26585
2^{nd}	Gogos	18674	De Smet	3847	Gogos	27640
3^{rd}	Atsuta	22559	Gogos	4139	De Smet	27815
4^{th}	pillay	23582	Atsuta	4638	Atsuta	29155
5^{th}	**Özcan**	**Infeasible**	Pillay	6860	**Özcan**	**29695**
6^{th}	De Smet	Infeasible	**Özcan**	**7541**	Pillay	32250
	Exam7		Exam8			
	Winner	Score	Winner	Score		
1^{st}	Muller	4213	Muller	7742		
2^{nd}	De Smet	5420	Gogos	10521		
3^{rd}	Gogos	6572	Atsuta	14317		
4^{th}	Atsuta	10473	Pillay	15592		
5^{th}	**Özcan**	**15116**	**Özcan**	**20168**		
6^{th}	Pillay	17666	De Smet	Infeasible		