A Statistical Analysis of the Features of a Dynamic Tabu Search Algorithm For Course Timetabling Problems

Ruggero Bellio · Luca Di Gaspero · Andrea Schaerf

1 Introduction

In this work we propose a dynamic tabu search algorithm for the solution of timetabling problems, and we undertake a systematic statistical study of the relative influence of the relevant features on the performances of the algorithm. In particular, we apply statistical methods for the design and analysis of experiments. The ultimate objective is to develop a procedure for obtaining the best combination of parameters for the algorithm for a given instance and predicting them for the unseen ones.

The study focuses on a basic timetabling problems, namely the course timetabling problem formulations used for the International Timetabling Competition (ITC-2007) as track 3 (see [2] for details). The instances upon which the algorithm is experimented are also the official ones of the competition.

The analysis is still ongoing, and it includes screening of important factors, building of response surfaces based on suitable experimental designs, classification of new instances based on input features.

2 Dynamic Tabu Search

Tabu Search (TS) is a well-known metaheuristic technique, and we do not describe it here for the sake of brevity, but we refer to [3,4] for details.

Our TS is dynamic in the sense that it changes continuously the shape of the cost function in an adaptive way, thus causing the search trajectory to pass through infeasible states and visit states that have a different structure than the previously visited ones.

R. Bellio
Dipartimento di Scienze Statistiche – Università di Udine
via Treppo 18, I-33100, Udine, Italy
E-mail: bellio@dss.uniud.it

L. Di Gaspero and A. Schaerf
Dipartimento di Ingegneria Elettrica, Gestionale e Meccanica – Università di Udine
via delle Scienze 208, I-33100, Udine, Italy
E-mail: l.digaspero@uniud.it, schaerf@uniud.it

We briefly describe here our instantiation of the features of TS and its customisation to our problem:

Neighborhood function: Move one lecture to a different room and/or a different timeslot.

Neighborhood sampling: The *full* neighborhood is traversed and all non-tabu neighbors are evaluated.

Tabu length: The tabu list is of variable size, so that each performed move remains in the list for a number of iterations randomly selected between t_{min} and t_{max}.

Prohibition: A move is tabu if a move in the list involves the same lecture.

Aspiration criterion: The aspiration criterion is the standard one: a tabu move is accepted if it leads to a state that is better than the current best one.

Stop criteria: The stop criterion is based on the number of iterations since the last improvement (parameter *max idle iterations*). However, when the number of idle iterations reaches its maximum, the search is not stopped but it restarts from the best state found that far. This procedure is finally stopped when it could not find any improvement for a given number of rounds (parameter *max idle rounds*).

Cost function dynamics: Our tabu search makes use of an adaptive modification of the weights for the violations of the hard constraints. Namely, the weight of each hard component is let to vary according to the so-called *shifting penalty* mechanism: if for a number k of consecutive iterations all constraints of that component are satisfied (resp. not satisfied), then the weight is divided (multiplied) by a factor α.

The main parameters of the algorithm are the following:

Tabu length: in order to reduce the number of parameters we consider here only t_{min}, and we fix t_{max} to $t_{min} + 5$ throughout the analysis.

Cost dynamics: The cost dynamics are incorporated in the factor α. Besides α, we also consider the also the minimum (w_{min}), the maximum (w_{max}), and the initial weight (w_{start}) of the hard constraints.

Idle rounds: The number of maximum idle rounds i_r is also a parameter of the algorithm, and it is included in the analysis.

3 Statistical Analyses

Statistical analyses are performed following the approach of sequential experimental design [7], using the R statistical software [5].

First, a screening experiment was run to eliminate unimportant algorithm parameters. After this step, we identified t_{min} and α as the most crucial parameters for minimizing the total cost.

The second step consists in a series of experiments developed according to the principles of response surface methodology. The aim of this step is to approximate the relation between mean cost and algorithm parameters with a second-order linear model. A scaled version of cost is taken as the response variable, in order to stabilize the coefficient of variation of the response across different instances. Following [6], random seed was taken as a (random) blocking factor. It was verified that second-order surfaces were providing an acceptable fit in most of the available instances, and then the results from all the instances were merged together by means of suitable mixed models [1]. In particular, all the parameters describing the response surface were taken as random

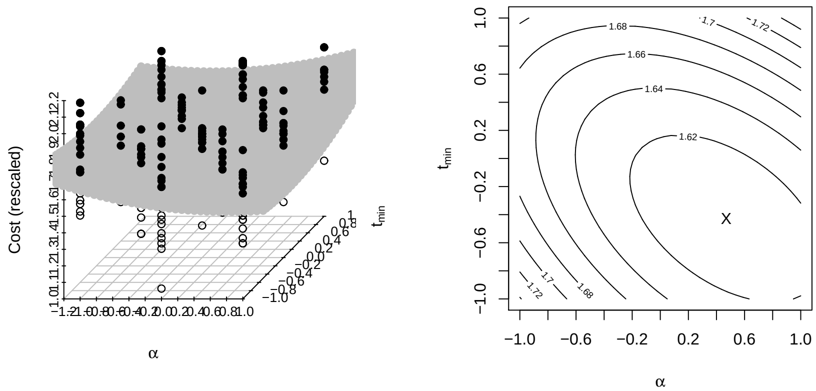

Fig. 1 Fitted second-order response surface (left) and contours of constant cost with the point of miminum cost (right).

coefficients. This provided a more realistic model, and had the result to further smooth the fitted responses

Figure 1 shows the fitted surface and the contours of constant cost for one of the instances, for which it was possible to locate a point of minimum mean cost. It is apparent that the contours are orientated along a negative-slope line, suggesting that the tabu search algorithm would make more progress for those combinations of α and t_{min} having negative correlation.

The final step of the analysis comprises the use of instance features to classify new instances and obtain the combination of algorithm parameters providing the optimal mean cost. This amounts to use the information extracted from the fitted models to study the relation between the optimal tuning of algorithm parameters and instance features. This part of the research is still ongoing.

References

1. Jørgen Bang-Jensen, Marco Chiarandini, Yuri Goegebeur, and Bent Jørgensen. Mixed models for the analysis of local search components. In Thomas Stützle, Mauro Birattari, and Holger Hoos, editors, *Engineering Stochastic Local Search Algorithms (SLS-2007)*, number 4638 in Lecture Notes in Computer Science, pages 91–105. Springer-Verlag, 2007.
2. Luca Di Gaspero, Barry McCollum, and Andrea Schaerf. The second international timetabling competition (ITC-2007): Curriculum-based course timetabling (track 3). Technical Report QUB/IEEE/Tech/ITC2007/CurriculumCTT/v1.0/1, School of Electronics, Electrical Engineering and Computer Science, Queens University, Belfast (UK), August 2007. ITC-2007 site: `http://www.cs.qub.ac.uk/itc2007/`.
3. Fred Glover and Manuel Laguna. *Tabu search*. Kluwer Academic Publishers, 1997.
4. Holger H. Hoos and Thomas Stützle. *Stochastic Local Search – Foundations and Applications*. Morgan Kaufmann Publishers, San Francisco, CA (USA), 2005.
5. R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008.
6. Enda Ridge and Daniel Kudenko. Screening the parameters affecting heuristic performance. In Hod Lipson, editor, *GECCO*, page 180. ACM, 2007.
7. Thomas P. Ryan. *Modern Experimental Design*. John Wiley & Sons, 2007.