Landscape-based Performance Prediction for University Timetabling Optimization

Thomas Feutrier[1], Marie-Éléonore Kessaci[1], Nadarajen Veerapen[1]

Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRIStAL, France

{thomas.feutrier,mkessaci,nadarajen.veerapen}@univ-lille.fr

Mots-clés : *Timetabling, Landscape analysis, Performance prediction*

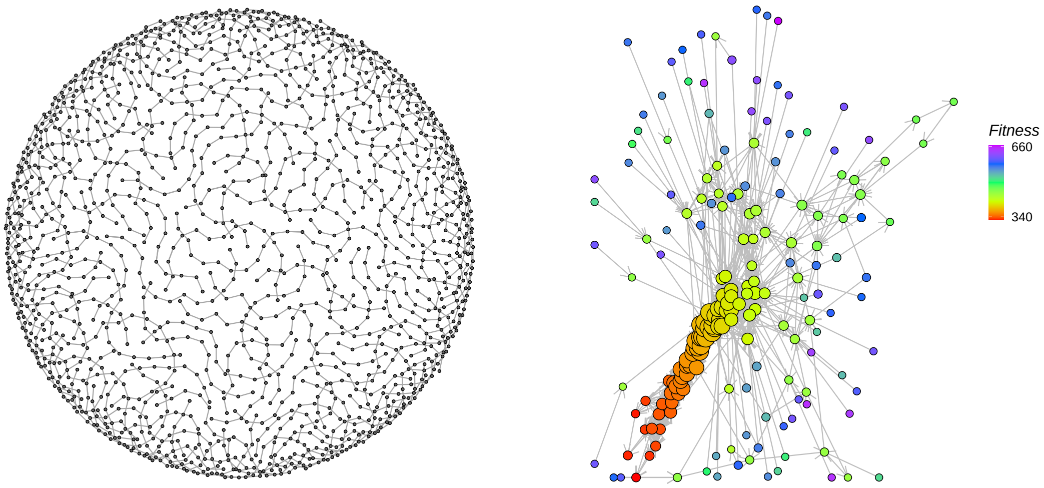
1 Context

The *Curriculum-Based Course Timetabling* (CB-CTT) is a University Timetabling problem considered as NP-hard. It is widely studied in the literature because of both scientific and practical interest. In CB-CTT, lectures have to be assigned to time slots and rooms with respect to resource constraints, i.e., the teachers' availability. The specificity of CB-CTT compared to university timetabling is the notion of the curriculum. Indeed, we only consider groups of student following predefined sets of lectures, as in French universities. The objective of the optimization problem is to minimize the number of constraint violations.

Heuristics and hyper-heuristics are known to give a good compromise between performance and rapidity to solve university timetabling problems. These methods have many parameters and can benefit from information about instance problems to improve their performance. Many features are proposed in the literature to help the parameterization of heuristics. In this work, we focus on landscape metrics like in the works of Ochoa et al. [2]. Therefore, we compute many features in order to identify the relevant ones. Then, we build a model to predict the performance of a local search based on the selected features.

2 Networks definition

A search landscape can be viewed as a spatial representation of solutions. For our work, we use a network, inspired by Local Optima Networks [2], to represent the CB-CTT search spaces. This network is built by sampling the search space using Iterated Local Search, a search method that alternates between hill climbing and perturbation phases. CB-CTT is known to be quite neutral meaning that many solutions share the same fitness value. Therefore, the exploration strategy of the hill climbing used accepts better or equal solutions to move through the search space and stops when too many neutral moves have been performed. This corresponds to a sequence of equal-fitness solutions, a plateau, that can be considered as local optima for the purposes of the search. Once the hill climbing stops, it is followed by a perturbation that jumps a certain distance in the search space and hill climbing is carried out again. This continues until a time limit is reached. The local optima plateaus (nodes) and the connections between them (arcs) define a trajectory. Repeating the ILS sampling a number of times (100 in our case) allows us to generate a network (Figure 1a). At this stage, there is very little to no connection between the trajectories because of the high level of neutrality and the large plateau size. We therefore chose to contract the network on the basis of fitness to obtain a new network where each node is a unique fitness. This new network exhibits a number of connectivity patterns (Figure 1b) and multiple graph theory metrics can be computed.



(a) Network before contraction – nodes are plateaus

(b) Network after contraction – nodes are unique fitness values

FIG. 1 – Network visualization.

3 Landscape Analysis

We consider 19 instances from the 2007 International Timetabling Competition (ITC 2007). Classical metrics of the literature have been computed and show a similar tendency between the considered instances. For each network, we notice that some nodes are very connected and inversely others are not. The general pattern we can observe is that relatively few solutions found early in the ILS trajectory share the same fitness. However, as the search progresses, trajectories often end up in the same fitness levels, as expected. Over a hundred network metrics, including ones related to node degree, edge weight, assortativity and PageRank centrality, are computed.

4 Predictive Model

Hybrid Local Search (HLS) is a successful method designed to solve CB-CTT [1]. HLS sequentially iterates a Hill Climbing, a perturbation algorithm and a simulated annealing. We propose to predict its performance using the features computed in the last section to describe the connectivity patterns in the search space. More traditional instance features, for example related to size, are also considered. First, we apply a feature selection based on correlation with the performance to select the relevant features. Then, we train a regression model. Results are encouraging since they show that the model is able to predict if a HLS run reached the best possible fitness with an R^2 equal to 0.96.

Références

[1] Tomáš Müller. ITC2007 solver description : a hybrid approach. *Annals of Operations Research*, 172(1) :429–446, November 2009.

[2] Gabriela Ochoa, Nadarajen Veerapen, Fabio Daolio, and Marco Tomassini. Understanding Phase Transitions with Local Optima Networks : Number Partitioning as a Case Study. In Bin Hu and Manuel López-Ibáñez, editors, *Evolutionary Computation in Combinatorial Optimization*, Lecture Notes in Computer Science, pages 233–248, Cham, 2017. Springer International Publishing.