New benchmark datasets for the RCMPSP

Rob Van Eynde[1] and Mario Vanhoucke[1,2,3]

[1] Faculty of Economics and Business Administration, Ghent University, Belgium
rob.vaneynde@ugent.be
mario.vanhoucke@ugent.be
[2] Operations and Technology Management Centre, Vlerick Business School, Belgium
[3] UCL School of Management, University College London, United Kingdom

Keywords: Multiproject scheduling, benchmark datasets, decoupled scheduling.

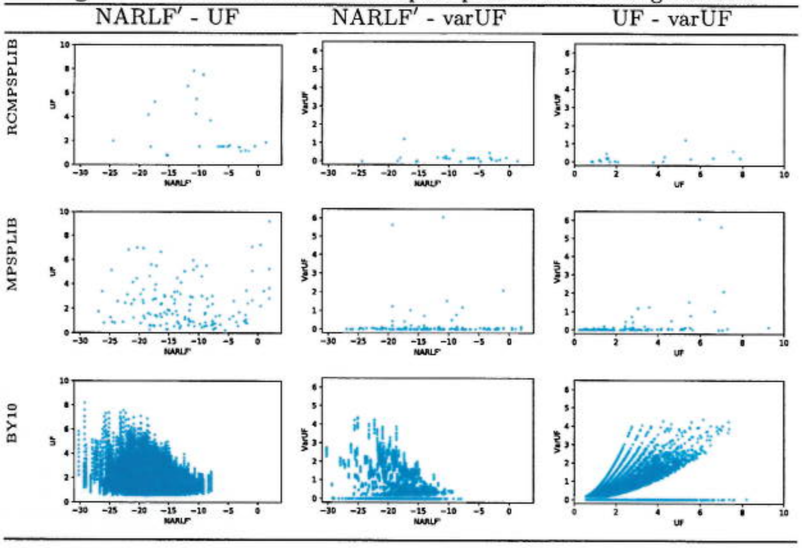
1 Introduction

In the resource-constrained project scheduling problem (RCPSP), a set of activities subject to precedence and resource constraints needs to be scheduled. Many extensions to the base problem have been developed and analysed (Hartmann and Briskorn 2010). One of the extensions is the resource-constrained multi-project scheduling problem (RCMPSP) in which a portfolio of different projects needs to be scheduled. All projects require (a subset of) the available renewable resource types, but there are no interproject precedence relations. The problem is to construct a resource feasible schedule in which a (time-related) objective is minimised. In recent years, the multi-project scheduling problem and its extensions have received more attention from researchers. However, the developed solution procedures are not always tested on the same datasets, inhibiting an unbiased comparison between the algorithms. The objectives of our research are threefold. First, we evaluate the existing datasets based on the feature space that they occupy. Second, we augment the generation procedure of Browning and Yassine (2010b) and use it to generate new datasets that cover a wider range of parameter values. Third, we set up a computational experiment with two scheduling procedures to analyse the differences between the new and existing datasets.

2 Existing datasets

In literature there are three publicly available datasets for the RCMPSP: MPSPLIB (Homberger 2007, Homberger 2012), BY10 (Browning and Yassine 2010b) and RCMPSPLIB (Vázquez *et. al.* 2015). Furthermore, there is the dataset of Wauters *et. al.* (2016) for the multi-mode extension of the RCMPSP. Because we address the basic variant, we will omit the last one from our analysis. To evaluate the datasets, we base ourselves on three resource-related measures that were designed for the multi-project context (Browning and Yassine 2010a). The first measure is the Normalised Average Resource Loading Factor (NARLF), which indicates whether the main part of the total resource demand occurs in the first or second half of the portfolio duration. We propose a small adaptation to the way the measure is calculated and name this alternative NARLF′. Second, the Modified Average Utilisation Factor (MAUF) compares the total resource demand to the total resource supply in the critical path schedule and takes the maximum ratio over all resource types. A higher MAUF means that a resource type is more constrained. The third measure (σ^2_{MAUF}) calculates the variance of the MAUF values of the different resource types from the maximum value and exhibits how equally constrained the different resource types are. Figure 1 shows the two-dimensional feature space plots for all instances of the three existing datasets.

Fig. 1. The 2-dimensional feature space plots of the existing datasets

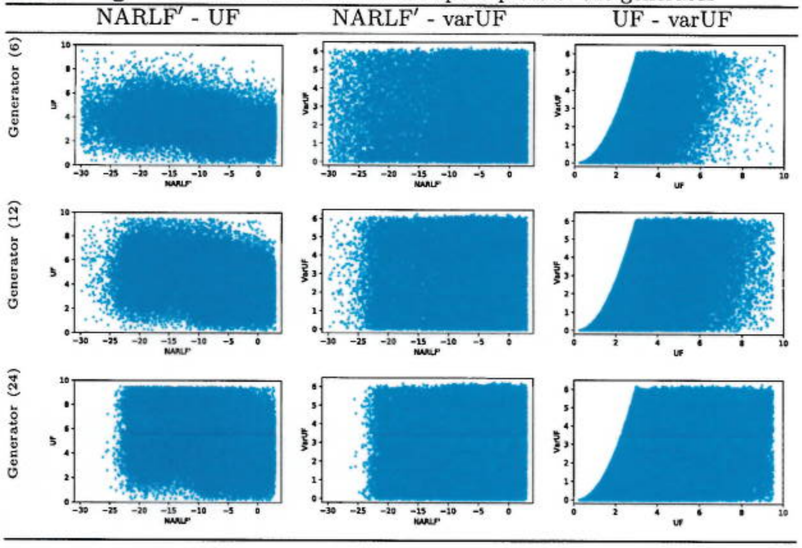


These plots show us that most of the instances in RCMPSPLIB and MPSPLIB have a low variance of resource utilisation. Furthermore, we see that BY10 contains a lot of instances but most of them have a NARLF′ lower than -10. Furthermore, the analysis of the network topology shows that most of the networks of the constituent projects are rather parallel. Based on these insights we adapt the generation procedure of Browning and Yassine (2010b) such that it is able to generate instances over a wider range of parameter combinations. The major adaptation is that RanGen2 (Vanhoucke *et. al.* 2008) is used to generate single-project networks, as it is capable of generating instances over the whole spectrum from parallel to serial networks. In order to compare the procedure with the existing sets, we generated instances of 6, 12 and 24 projects over the whole range of parameter combinations. The plots in Table 2 show that the generator is able to obtain instances over a broader feature space than the existing datasets. Although the resulting datasets cover a wider range, we observe dependencies between NARLF′ and the network structure of the single-projects. As a consequence, the feasible range of NARLF′ values for an instance depends on the network structure of its constituent projects. We propose the new datasets 6_60, 12_60 and 24_60, where each instance consists of 6, 12 and 24 projects respectively and each project contains 60 activities.

3 Evaluation using heuristics

In a second study we analyse the differences between the datasets in a computational experiment using decoupled schedule generation schemes (SGS) and a genetic algorithm (GA). Decoupled SGS's are similar to the priority rule based scheduling schemes from single-project literature. The difference lies in the fact that a single-project SGS selects one activity from the eligible set at each decision moment, while a decoupled SGS first selects a project and then selects an eligible activity from that project. In our computational experiment we tested all combinations of 16 project priority rules and 16 activity priority rules. Based on the results we come to the following conclusions. First, decoupled SGS's outperform single-project SGS's on all datasets when the objective is the minimisation of

Fig. 2. The 2-dimensional feature space plots of the generator



average project delay. Second, the choice of project priority rule has a larger impact on the performance than the choice of activity priority rule. Third, project rules related to the total resource demand perform better on the existing datasets, while critical path based rules perform better on the datasets that we propose. This indicates that the new datasets cover a part of the feature space that is not present in the existing datasets.

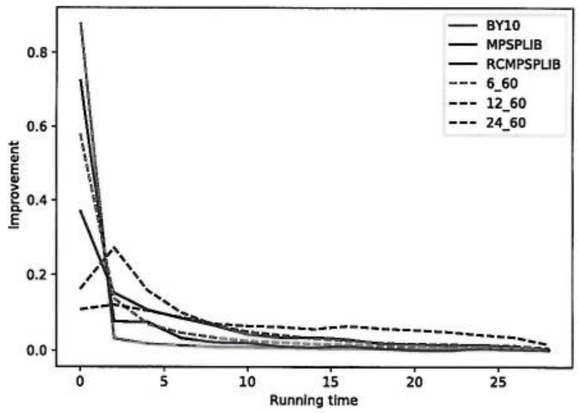
As solution procedures for the RCMPSP have advanced beyond priority rules, we execute a second computational experiment with a genetic algorithm. We implement four crossover operators and several mutation operators of Asta *et. al.* (2016). We let the GA solve each instance with a time limit of 30 seconds. Figure 3 shows the improvement found over time, relative to the total improvement after 30 seconds. The graph shows that on average, the GA quickly finds the largest improvement in the first few seconds for the datasets BY10, RCMPSPLIB and 6_60. The other datasets prove to be more challenging for the metaheuristic, as it keeps finding sizeable improvements for a longer time. This effect is the strongest for the dataset 24_60, where the figure suggests that the algorithm will still find relatively large improvements after 30 seconds. This experiment shows that MPSPLIB, 12_60 and 24_60 are more challenging for more advanced heuristics.

4 Conclusions

We propose new datasets for the RCMPSP that complement the existing sets in two ways. First, the new datasets contain a wider variety of parameter combinations, leading to different conclusions about which priority rules perform best. Second, the sets with 12 and 24 projects per instance (in addition to the set MPSPLIB) proved to be more challenging than the others for the genetic algorithm.

Furthermore, the experiments show that for the RCMPSP, decoupled schedule generations schemes outperform the traditional single-project scheduling schemes and that project prioritisation has the largest impact on performnce.

Fig. 3. Improvement profile of GA



An opportunity for future research is the design of new summary measures for the RCMPSP that are less dependent upon each other, which would facilitate generation of instances over the complete feature space.

Acknowledgements

We acknowledge the support provided by the Special Research Fund (BOF grant no. DOC014-18 Van Eynde) and the National Bank of Belgium for providing the first author with a pre-doctoral fellowship. The computational resources (Stevin Supercomputer Infrastructure) and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by Ghent University, FWO and the Flemish Government department EWI.

References

Asta, S., Karapetyan, D., Kheiri, A., Özcan, E. and Parkes, A. J., 2016, "Combining monte-carlo and hyper-heuristic methods for the multi-mode resource-constrained multi-project scheduling problem", *Information Sciences*, Vol. 373, pp. 476-498.

Browning, T. R., Yassine, A. A., 2010a, "Resource-constrained multi-project scheduling: Priority rule performance revisited", *International Journal of Production Economics*, Vol. 126(2), pp. 212-228.

Browning, T. R., Yassine, A. A., 2010b, "A random generator of resource-constrained multi-project network problems", *Journal of Scheduling*, Vol. 13(2), pp. 143-161.

Hartmann, S., Dirk, B., 2010, "A survey of variants and extensions of the resource-constrained project scheduling problem", *European Journal of Operational Research*, Vol. 207(1), pp. 1-14.

Homberger, J., 2007, "A multi-agent system for the decentralized resource-constrained multi-project scheduling problem", *International Transactions in Operational Research*, Vol. 14(6), pp. 565-589.

Homberger, J., 2012, "A (μ, λ)-coordination mechanism for agent-based multi-project scheduling", *OR spectrum*, Vol. 34(1), pp. 107-132.

Vanhoucke, M., Coelho, J., Debels, D., Maenhout, B. and Tavares, L. V., 2008, "An evaluation of the adequacy of project network generators with systematically sampled networks", *European Journal of Operational Research*, Vol. 187(2), pp. 511-524.

Vázquez, E. P., Calvo, M. P. and Ordóñez, P. M., 2015, "Learning process on priority rules to solve the RCMPSP", *Journal of Intelligent Manufacturing*, Vol. 26(1), pp. 123-138.

Wauters, T., Kinable, J., Smet, P., Vancroonenburg, W., Berghe, G. V. and Verstichel, J., 2016, "The multi-mode resource-constrained multi-project scheduling problem", *Journal of Scheduling*, Vol. 19(3), pp. 271-283.