

Statistical assessments to methods for minimizing tardiness in parallel machines scheduling problems

Ozlem Senvar, Andres Felipe Bernate Lara, Farouk Yalaoui, Frédéric Dugardin

*Institut Charles Delaunay,
Laboratoire d'Optimisation des Systèmes Industriels (LOSI),
Université de Technologie de Troyes,
12, rue Marie Curie – CS 42060- 10004 Troyes cedex - France*

Keywords: Scheduling parallel machines, Statistical assessments, Metaheuristics

1 Introduction

This study proposes a set of solutions methods for the problem of scheduling in which a set of n jobs in m identical machines minimizing the total tardiness, when jobs have different release dates. According to Lawler's representation, this problem is denoted as $P_m / r_i / \sum T_i$. The tardiness is known as the time passed over the due date to finish each job. The described problem is NP-hard since the relaxation without release date was proved to be NP-hard by Koulamas [1].

To the best of our knowledge regarding this scheduling problem, this problem is less studied than one machine scheduling problem. For the current problem, Yalaoui [2] studied the problem selecting an outperforming heuristic method. In 2012, Kacem *et al.* [3] proposed a branch and bound algorithm to minimize the weighted tardiness minimization. Other publications refer to similar problems including different constraints of jobs or different objective functions.

In this study, we assume that all machines are identical and preemption is not allowed. Performance analyses on the solution methods are examined in order to extract alternative conclusions on the hardness of instances and their potential links with parameters of instances (n , m , release dates, processing times, and due dates).

Actually, in practice for industrial applications of the problem, decision makers need to extract further information about instances. Regarding the problem and the theoretical results, in order to have some sort of statistical guarantee on high performances of the choice and/or the guarantee of rationality, a statistical decision theoretic framework is useful for accurate, precise, and consistent decision making. From this standpoint, this study executes the solution methods of described scheduling problem and handles results with statistical assessments. This study provides statistical assessments to the specified methods along with the proposition of performance indicators enhancing relationships in terms of predetermined inputs.

2 Background Information

The parallel machines scheduling problem describes an industrial context where a workshops has multiple resources to process incoming jobs. Raw materials restrictions or conditionals production schedules impose minimal dates to launch a job processing. Solving this NP-hard problem for big size instances is a priority, much more reasonably the tardiness criterion becomes à high priority in quality service evaluations [1].

To solve the described scheduling problem, we propose a set of heuristic and metaheuristic methods, including priority list methods, local search methods, iterated local search methods, Tabu search methods and a multi start partial local search method. A total of 171 different methods are developed, for a permutation of initials solutions and iterative search procedures.

Methods are tested in a set of 4500 instances considering the instances generator model proposed by Yalaoui [2]. Instances are generated for $n=12, 40$ and 100 , and $m=2, 3$ and 5 . Release dates, processing times and due dates are generated using uniform distributions in range described on the framework given in [2].

Results are handled in order to link instance parameters (n and m) with the performance of the set of tested methods. Following section describes the performed analyses.

3 Statistical Assessments

Regression analysis is widely used for prediction and forecasting, where its use has substantial overlap with the field of machine learning. Regression analysis is also used to understand which independent variables are related to the dependent variable, and to explore the forms of these relationships. In restricted circumstances, regression analysis can be used to infer causal relationships between the independent and dependent variables.

Traditional linear regression model can acquire knowledge through the least squares method and store that knowledge in the regression coefficients. However, linear regression has a rigid model structure and set of assumptions that are imposed before learning from the data Turanoglu *et al.* [4].

In order to adapt statistical assessments to the described scheduling context, we are using experimental results of solving 4500 different instances with 171 methods.

Results are exposed in the final version of the article.

4 Conclusion

In this study, we propose a set of solution methods for the parallel machines scheduling problem with total tardiness minimization. Results are handled in order to link instance's predetermined inputs (n and m) with the performance of the set of tested methods and recommendations are given for real case applications emphasizing the difficulties from theoretical and practical perspectives. At the end, further directions regarding the issues for theoreticians and practitioners are included.

References

- [1] Koulamas C. The total tardiness problem: review and extensions. *Optimal research*, 42:1025-1041, 1994.
- [2] Yalaoui F. Minimizing total tardiness in parallel-machine scheduling with release dates. *Journal of Applied Evolutionary Computation*, 3:21–46, 2012.
- [3] Kacem I., Souayah N., and Haouari M. Branch-and-bound algorithm for total weighted tardiness minimization on parallel machines under release dates assumptions. *RAIRO - Operations Research*, 46:125–147, 4 2012
- [4] Turanoglu E, Senvar O., Kahraman C. Oil Consumption Forecasting in Turkey using Artificial Neural Network. *International Journal of Energy Optimization and Engineering*, 1(4):89-105, 2012.