Adaptive Robust Parallel Machine Scheduling

Izack Cohen[1], Krzysztof Postek[2] and Shimrit Shtern[3]

[1] Faculty of Engineering, Bar-Ilan University, Ramat-Gan, Israel
izack.cohen@biu.ac.il
[2] Faculty of Electrical Engineering, Mathematics and Computer Science, University of Technology, Delft, The Netherlands
k.s.postek@tudelft.nl
[3] Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology, Haifa, Israel
shimrits@technion.ac.il

Keywords: robust optimization, machine scheduling, adaptive scheduling.

1 Problem definition:

Parallel machine scheduling (PMS) problems are multi-stage scheduling problems, which are widely researched owing to their theoretical importance and multiple applications in manufacturing, cloud computing, and project management, among others. Real-life PMS settings involve uncertainty about task duration, which may be characterized by the randomness of each task duration and, possibly, a dependence between task durations.

An ideal scheduling approach should accommodate uncertainty to ensure realistic guarantees on the objective function value and permit adjustments of later-stage decisions based on different observed task lengths (e.g., different duration realizations of the task scheduled first may result in different allocation decisions of the next tasks).

Real-life parallel machine scheduling problems can be characterized by: (i) limited information about the exact task duration at scheduling time, and (ii) an opportunity to reschedule the remaining tasks each time a task has completed processing and a machine becomes idle. Robust scheduling has been used to deal with the first characteristic. However, the existing literature on robust scheduling does not explicitly consider the second characteristic – the possibility to adjust decisions as more information about the tasks' duration becomes available, despite the fact that re-optimizing the schedule every time new information emerges is a standard practice.

2 Methodology/results:

In this paper, we develop a robust optimization based scheduling approach that takes into account, at the beginning of the planning horizon, the possibility that scheduling decisions can be adjusted. We demonstrate that this adaptive approach can lead to better here-and-now decisions. To that end, we develop the first mixed integer linear programming model for *adjustable robust scheduling*, where we minimize the worst-case makespan. Using this model, we show via a numerical study that adjustable scheduling leads to solutions with better and more stable makespan realizations compared to static approaches.

We focus on makespan minimization, which is a standard performance measure for PMS. Indeed, makespan minimization is used for load balancing, an important issue for many scheduling applications. When deciding whether to use the expected value or worst-case value, several factors should be considered. Optimizing over an expectation requires specifying the full probability distribution of task duration, information that is often not readily available or is costly to acquire. Moreover, the makespan of a single realization can significantly differ from the expected value; thus, if the exact scheduling problem is not repeated multiple times, optimizing over the expected value may not be translated into good performance in practice. In contrast, much less information is needed when specifying a set that includes all the reasonable duration realizations, and a worst-case optimization approach provides a guarantee on the performance of any realization in such a set. Therefore, we choose a setting where the scheduler minimizes the worst-possible makespan of a set of tasks over some uncertainty set, which captures all reasonable scenarios within the support of the distribution. This is in line with the paradigm of Robust Optimization (RO), where the best solution is sought under the assumption that the problem's parameters are initially unknown and that, given the decisions, nature picks their worst-possible values from an *uncertainty set* consisting of outcomes that include the true realization with high probability.

We consider the classical version of PMS, where m identical machines process $n \geqslant m$ tasks that are available at the start of the scheduling horizon. For this problem, we construct a mixed integer linear optimization problem for minimizing the worst-case makespan, which includes all possible later-stage (re-)scheduling decisions. We compare the adaptive formulation's optimal scheduling decisions and optimal worst-case makespan to those of the optimal static allocation (SA) and static list (SL) policies.

In contrast to the majority of previous works, which compare naive implementations of the SA and SL policies without re-optimization (i.e., re-scheduling) as more information is revealed, we consider the more realistic rolling horizon implementation of these policies. Under this implementation, whenever one of the machines becomes idle, the scheduler can alter the initial order of tasks by re-solving an optimization problem with the extra information included.

3 Managerial implications:

We outline our main managerial insights for the studied setting. The insights are relevant to schedulers within multiple domains that can be modeled via PMS such as production lines in which machines process a set of tasks, computer multiprocessors ("cloud processing") for processing jobs, shipyards and ports in which ships are loaded and unloaded, doctors who treat patients in a walk-in clinic or triage setting, and teachers who educate student groups, just to name a portion of the potential use-cases.

First, our study shows that capturing the uncertainty and the relations between the durations of different tasks is vital to a realistic assessment of the makespan. Indeed, there are many settings in which the probabilistic knowledge about task durations is limited or costly to attain. In such circumstances, it is rather easy to design a polyhedral or ellipsoidal uncertainty set that frames the involved uncertainty. Ben Tal et al.(2009) provide guidance and probabilistic guarantees in favor of designing uncertainty sets that balance the level of conservatism and the probability that a constraint is violated by a scenario. Ideally, we would like to design the smallest uncertainty set that still captures the meaningful scenarios (e.g., the probability that a scenario is not included within the uncertainty set is lower than a pre-specified threshold).

Secondly, whenever the optimal wait-and-see decisions can be taken into account in the planning stage, this should be done as it lowers the maximum possible project makespan that the scheduler can promise. In other words, a bid prepared by a decision-maker who accommodates wait-and-see decisions and thus can commit to a lower makespan (and cost) would be more competitive than a bidder that does not explicitly take into account the possibility that decisions can be adapted. In particular, our experiments point out that the average advantage of adaptive-based bids is estimated to be $5-9\%$ over its non-adaptive (i.e., 'regular' RO) counterpart. We note that an adaptive policy need not necessarily be achieved by solving our mixed integer linear formulation. Indeed, it is likely that heuristic methods can be of help as well, and should be explored as an alternative to static policies.

While the previous point dealt with the superiority of adaptive robust policies over their static counterparts in the planning and contract stage, they are also preferable in the implementation stage. Specifically, policies that take the later-stage adaptivity of the decisions into account remain preferable even when the static policies are re-optimized every time new information becomes available (rolling-horizon). A hint into the reason for this is provided by the $42 - 59\%$ of the problem instances in which an adaptive policy yielded different first-stage decisions compared to a SA policy. That means that the adaptive policies not only offer better project makespan guarantees, but also select decisions that lead to better realized duration.

A very attractive feature of the adaptive policies, as revealed through our experiments, is that their performance is comparable to the perfect hindsight policy (e.g., the average difference between the promised and max perfect hindsight makespans was $0.0 - 0.1\%$ for the optimal adaptive policy compared to $5.7 - 9.8\%$ for the static robust policy). This suggests that the adaptive robust policy not only protect the decision-maker against adversarial realizations of reality but it also performs close to the perfect hindsight policy. Thus, the typical criticism about the conservatism of static robust policies (i.e., the high price paid for robustness) does not apply to the adaptive scheduling policy.

In conclusion, while robust SA policies are widely investigated and used in risk averse settings, they may achieve inferior performance in practice compared to adaptive alternatives. Since the performance gap between an optimal adaptive policy and a static one is quite significant, we recommend allocating resources for finding good adaptive policies, even if those policies are not necessarily optimal. We believe that these adaptive policies will grant their users competitive advantages both in the proposal bidding stage and in the implementation stage.

Acknowledgements

We are grateful to Michael Pinedo for consultations at the early stage of this work. Also, we thank Esther Julien for her help.

References

Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). Robust optimization. Princeton University Press.