Bilevel scheduling of uniform parallel machines in the context of coupling maintenance and scheduling decisions

Quentin Schau, Olivier Ploton, Vincent t'Kindt, Federico Della Croce

Bilevel scheduling of uniform parallel machines in the context of coupling maintenance and scheduling decisions

Schau Q.[1],Ploton O.[1],T'kindt V.[1] and Della Croce F.[2]

[1] Université de Tours, Laboratoire d'Informatique Fondamentale et Appliquée, Tours, France,
{quentin.schau,olivier.ploton,tkindt}@univ-tours.fr

[2] DIGEP, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, CNR, IEIIT, Torino, Italy.
federico.dellacroce@polito.it

Keywords: Scheduling, bilevel optimization, parallel machines

1 Introduction

Recent advancements in technology, such as cloud computing and the Internet of Things (IoT), have enabled the efficient management of data and information sharing. These technological developments have given rise to a new industrial era known as *Industry 4.0* (I4.0)(Chen et al. 2018). I4.0 introduces Cyber-Physical Systems (CPS) that monitor physical processes and enable the development of smart factories, including automated scheduling (Serranoruiz et al. 2021). In this paper, we focus on *Zero-Defect Manufacturing* (ZDM), a quality improvement approach grounded in data-driven methods. Specifically, we delve into a process-oriented variant of ZDM that incorporates the health status of production machines into scheduling decisions. For instance, ZDM can alleviate a machine's workload if it predicts an impending failure or, if maintenance is required, it can reschedule tasks accordingly. In essence, this approach involves the intertwining of maintenance issues with scheduling.

The ZDM framework can be represented as a bilevel optimization problem. In this setup, the leader manages system health and global scheduling, while the followers are responsible for multiple CPSs. The leader makes decisions first and hands over a subset of jobs to the followers, each one scheduling the tasks it has been assigned. Notably, both the leader and followers can only evaluate their objective function when both have taken their decisions. This bilevel optimization structure introduces complexities as the leader and followers may have distinct conflicting objectives (Dempe et al. 2015). Notably, the followers' problem can have multiple optimal solutions, leading to two categories of bilevel problems: *optimistic* and *pessimistic* (Dempe 2003). In the optimistic scenario, the follower selects, among its optimal solutions the one that leads to the smallest value of the leader's objective function, whereas in the pessimistic case, the follower chooses the one that is the worst for the leader.

Lately, to the best of our knowledge, the literature on bilevel scheduling is relatively limited (T'kindt et al. 2023).

This study focuses on a scenario with one leader and one follower (CPS) composed of parallel uniform machines where scheduling decisions occur periodically. The problem is introduced in Section 2, while complexity results and solution approaches are sketched in Section 3.

2 Optimistic bilevel parallel uniform machines scheduling

Assume we are given a set $\mathcal{J}$ of N jobs, each job j being characterized by its processing time p_j, weight w_j and due date d_j. We have a leader that manages the system health and global scheduling and a CPS, composed of parallel uniform machines, which schedules the jobs given by the leader. The leader aims to minimize the number of weighted tardy jobs while the follower aims to minimize the total completion time. We consider the optimistic case, i.e. the follower returns the schedule that leads to the minimum number of weighted tardy jobs among schedules optimal for the total completion time. So, the bilevel problem can be defined as:

$$\begin{aligned} &\min_{\substack{\mathcal{I}\subset\mathcal{J}\\ |\mathcal{I}|=n}} \quad \sum_{j\in\sigma^*} w_j U_j^L \\ &\text{st.} \\ &\qquad \sigma^* = \arg\min_{\sigma\in\mathfrak{S}_{\mathcal{I}}} \left(Lex\left(\sum_{j\in\sigma} C_j^F, \sum_{j\in\sigma} w_j U_j^L \right)\right) \end{aligned} \tag{1}$$

where:

- $\mathfrak{S}_{\mathcal{I}}$ is the set of schedules form of all jobs in $\mathcal{I}$.
- C_j^F is the completion times of the job j.
- $U_j^L = \begin{cases} 1 \text{ if } C_j^F > d_j \\ 0 \text{ otherwise} \end{cases}$
- $Lex\left(\sum_{j\in\sigma} C_j^F, \sum_{j\in\sigma} w_j U_j^L\right)$ stands for minimizing the first criterion $\sum_{j\in\sigma} C_j^F$ first, then the second criterion $\sum_{j\in\sigma} w_j U_j^L$ if there is a tie in the first criterion.

We assume that the leader periodically makes scheduling decisions every T units of time. Prior to each period, the leader receives data from the shop floor. Using this data, the leader faces three cases and makes related decisions:

- If one or more machines are not defective, the leader sets the machine speed to the maximum, i.e. V_{max}.
- If one or more machines are predicted to be defective, the leader reduces the machine speeds to the minimal speed V_0 to minimize the risk of breakdowns.
- If one or more machines are confirmed to be defective, the leader removes those machines from the follower for maintenance. The follower can no longer assign jobs to these machines for the scheduling period.

Next, the leader selects a subset $\mathcal{I} \subset \mathcal{J}$ of n jobs and assigns them to the follower. In the follower's problem, there is a lexicographical objective function. Initially, the follower minimizes the total completion time and subsequently minimizes the number of weighted tardy jobs.

Correspondingly, following the three-field classification (Graham et al. 1979), we denote this bilevel problem as $Q|V_i \in \{V_0, V_{max}\}, OPT-n|\sum_j C_j^F, \sum_j w_j U_j^L$, while the related follower problem is denoted as $Q|V_i \in \{V_0, V_{max}\}|Lex\left(\sum_j C_j^F, \sum_j w_j U_j^L\right)$.

3 Complexity and solution approaches

We show in the following section several complexity results (proofs are omitted and will be presented at the time of the conference). With respect to the complexity of problem $Q|V_i \in \{V_0, V_{max}\}|Lex\left(\sum_j C_j^F, \sum_j w_j U_j^L\right)$, the following theorem holds.

Theorem 1. *The problem $Q|V_i \in \{V_0, V_{max}\}|Lex\left(\sum_j C_j^F, \sum_j w_j U_j^L\right)$ (and correspondingly of problem $Q|V_i \in \{V_0, V_{max}\}, OPT - n|\sum_j C_j^F, \sum_j w_j U_j^L$) is $\mathcal{NP}$-hard in strong sense by reduction from the numerical 3-dimensional matching ($NUM - 3DM$).*

The answer to the $NUM - 3DM$ problem is true iff all jobs can be scheduled on time ($\sum_j w_j U_j = 0$), as a consequence, the problem is not approximable in polynomial or pseudo-polynomial time unless $\mathcal{P} = \mathcal{NP}$.

When the number of machines m is not part of the input, we can already show that problem $P2||Lex\left(\sum_j C_j^F, \sum_j U_j^L\right)$ is $\mathcal{NP}$-hard by reduction from the well known even-odd partition problem. To complete the complexity analysis, we can even prove the following theorem by providing a dynamic programming pseudo-polynomial algorithm.

Theorem 2. *$Qm|V_i \in \{V_0, V_{max}\}|Lex\left(\sum_j C_j^F, \sum_j w_j U_j^L\right)$ with m constant are $\mathcal{NP}$-hard in the ordinary sense.*

We remark, first, that problem $Q||\sum_j C_j$ is polynomial (Conway, Maxwell, Miller 1967) and that the algorithm we propose is based on a characterization of the set of solutions of the problem $Q|V_i \in \{V_{max}, V_0\}|\sum_j C_j$.

Assume we have a set $\mathcal{J}$ of n jobs, let N_{max} (respectively N_0) be the number of machines with high-speed V_{max} (respectively, with low speed V_0). Assume that V_0 divide V_{max}. We can show that an optimal schedule of the problem $Q|V_i \in \{V_{max}, V_0\}|\sum_j C_j$ is given by the repetition of patterns. In Figure 3, an illustration depicts a pattern consisting of $\alpha - 1$ blocks on high-speed machines and one block on low-speed and high-speed machines. Jobs within the blue-dashed group signify their membership to a block and can subsequently undergo permutation without change the value of $\sum_j C_j$. Let us define $\mathcal{B}_h$ a block that can be assigned to high-speed machines or low-speed machines.

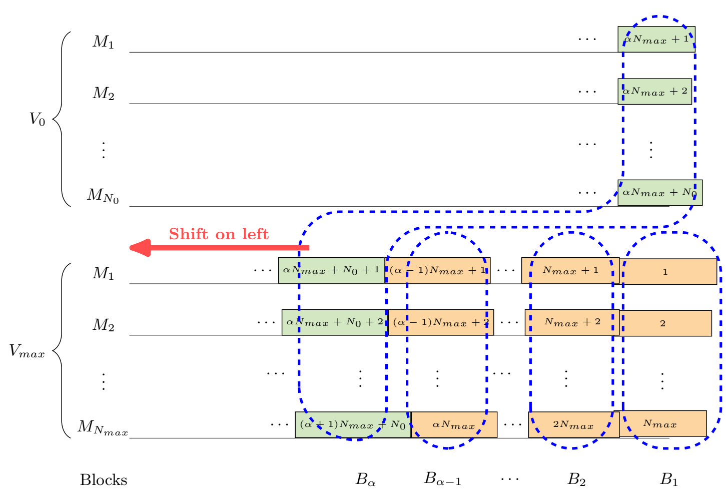

Fig. 1. Scheduling pattern with block structure for the optimal solution of $Q|V_i \in \{V_{max}, V_0\}|\sum_j C_j$

Hence, we know for all jobs j on which block $\mathcal{B}_h$ it is scheduled. Using this characterization, by extending the approach on parallel machines scheduling proposed in (Hall et al. 2001) to the considered problem, a dynamic programming recursion can be formulated to solve $Q|V_i \in \{V_0, V_{max}\}, OPT - n|\sum_j C_j^F, \sum_j w_j U_j^L$ by means of function OPT.

$$OPT(\vec{\mathcal{A}}^0, \mathcal{A}^{\vec{max}}, \vec{C}^0, C^{\vec{max}}, j, n') \quad (2)$$

where:

- j is the job to add to the schedule under construction
- n' the number of jobs that already scheduled
- $\vec{C^\alpha}$ is the vector of completion times for the machines with the speed α
- $\vec{\mathcal{A}}^\alpha$ is the vector indicating at the current recursion the available machines with speed α.

  For each machine i we have $\vec{\mathcal{A}}_i^\alpha = \begin{cases} \text{True if the job } j \text{ can be schedule on the machine } i \\ \quad\quad \text{with the speed } V_\alpha \\ \text{False otherwise} \end{cases}$

Here, OPT returns the minimal value of $\sum_{k=1}^n w_k U_k$ when jobs $\{1, \ldots, j\}$ have been scheduled and machines complete at times $\vec{C}^0$ and $C^{\vec{max}}$. By computing $OPT(T\vec{rue}, T\vec{rue}, \vec{C}^0, C^{\vec{max}}, n)$ for all relevant $\vec{C}^0$ and $C^{\vec{max}}$, the following lemma holds.

Lemma 1. *A dynamic programing algorithm find an optimal solution for the problem* $Q|V_i \in \{V_0, V_{max}\}, OPT - n|\sum_j C_j^F, \sum_j w_j U_j^L$ *with the time complexity* $\mathcal{O}\left(4^m \left(\sum_j p_j\right)^m \times N \times n^2 log(n)\right)$

In addition, a MIP can be formulated for the bilevel problem that will be presented in details at the time of the conference.

References

Conway, R. W., Maxwell, W. L., Miller, L. W. 1967, Theory of Scheduling, Addison-Wesley Publishing Company

Graham, R. L., Lawler, E. L., Lenstra, J. K., Kan, A. R. , 1979, Optimization and approximation in deterministic sequencing and scheduling: a survey, *Journal Title*, Vol. 5, pp. 287–326.

Hall, N. G. , Lesaoana, M., Potts, C. N. : Scheduling with fixed delivery dates. Operations Research. Vol 49, pp134–144 .

Dempe, S.: Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. Optimization. 52, 333–359 (2003).

Dempe, S., Kalashnikov, V. Perez-Valdes, G.A., Kalashnikova, N., 2015, “Bilevel programming problems", *Springer*.

Chen, B., Wan, J., Shu, L., Li, P., Mukherjee, M., Yin, B. 2017, Smart Factory of Industry 4.0: Key Technologies, Application Case, and Challenges,IEEE Access, Vol. 6, pp. 6505-6519

Serrano-Ruiz, J. C., Mula, J., Poler, R., 2021, Smart manufacturing scheduling: A literature review, Journal of Manufacturing Systems, Vol. 61, pp. 265-287

T'kindt, V., Della Croce, F., Agnetis, A.: Adversarial bilevel scheduling on a single machine. European Journal of Operational Research, articles in press (available on line november 2023).