

Analysis of Different Criteria for Workload Balancing on Identical Parallel Machines

Yassine Ouazene, Farouk Yalaoui, Hicham Chehade and Alice Yalaoui
Institut Charles Delaunay - Laboratoire d'optimisation des Systèmes Industriels
UMR - STMR - CNRS 6279
Université de Technologie de Troyes
12 rue Marie Curie CS 42060, 10004 Troyes, France

Abstract

This paper deals with the workload balancing problem in identical parallel machines context. It consists of assigning n different jobs to m identical parallel machines as equally as possible. First, we present a critical analysis of classical formulation based on the minimization of the maximum workload or the completion time of the bottleneck machine previously presented in the literature. We also analyze the two main performance measures, usually used in the literature, to deal with this problem. These criteria are respectively the relative percentage of imbalances (*RPI*) introduced by Rajakumar et al. (2004, 2007) and the normalized sum of square for workload deviations (*NSSWD*) proposed by Ho et al. (2009). Based on an illustrative example, we show that the classical formulation based on the minimization of the maximum completion time (or maximum workload) does not provide the optimal workload repartition. Then, we propose a mixed integer linear programming model to obtain the optimal workload allocation. This formulation is based on the minimization of the difference between the workload of the bottleneck machine and the workload of the fastest machine.

Keywords

Scheduling, parallel machines, workload balancing, mathematical programming

1. Introduction

Workload balancing is important for both service and manufacturing industries. In service industry, human resources should have a balanced workload in order to be equitable and provide a quality service. The aim of a manager is to assign jobs to each worker in such a way that their workloads are as similar as possible. In manufacturing industry, balancing the workload among the machines is important to reduce the idle times and work-in-process. It helps also to remove bottlenecks in manufacturing systems. Rajakumar et al. (2004) have addressed the workload balancing problem using different priority rules such as: random, shortest processing times and longest processing times. The authors used the relative difference of imbalance to evaluate the performances of these different strategies. In their next publication Rajakumar et al. (2007), the authors proposed a genetic algorithm that outperforms these different strategies.

Based on these classical priority rules, Raghavendra and Murthy (2010) made an effort to reduce the imbalance in random type of parallel machines addressing the loading problem in flexible manufacturing system. Later, Raghavendra et al. (2010) proposed a genetic algorithm based approach with SPT and LPT rules for reducing the imbalance between the parallel machines. The authors have concluded that their genetic algorithm provides better solutions than the strategies proposed in the first paper (2010). Raghavendra et al. (2011) applied this genetic algorithm in a case study for ten different part styles with different batch quantity on two vertical machining centers. The same genetic based heuristics algorithm was compared with other approximate approaches proposed in the literature.

In the literature, the workload balancing problem has been associated with different scheduling criteria in different ways, even by considering the workload imbalance as a constraint or as an objective. For example, Ouazene et al. (2011) addressed the identical parallel machine scheduling problem to minimize simultaneously total tardiness and workload imbalance. The authors proposed a mathematical formulation and a genetic algorithm based on the aggregation of the two objective functions. Yildirim et al. (2007) have studied the scheduling of semi-related machines with sequence dependent setups and load balancing constraints. The authors proposed a mathematical formulation of the problem and an approximate resolution based on some heuristics and a genetic algorithm.

Recently, Keskinturk et al. (2012) presented a non linear mathematical model for a parallel machine problem with sequence-dependent setups with the objective of minimizing the total relative imbalance. The authors proposed some heuristic and metaheuristic methods to solve the problem. The two metaheuristic methods consist of an ant colony optimization algorithm and a genetic algorithm. Based on various random tests, the authors have concluded that the ant colony algorithm outperforms both heuristics and genetic algorithm.

The remainder of this paper is organized as follows. In section 2, we detail and analyze the mathematical formulation proposed by Rajakumar et al. (2007). The different notations and decision variables are described. Also, some important remarks concerning this formulation are commented. Section 3 presents the modified formulation. In section 4, we present a counter example to attest of the non-optimality of the first formulation. Finally, section 5 summarizes the contribution of this paper.

2. Mathematical formulation proposed by Rajakumar et al. (2007)

In this section we present the mixed integer programming model proposed by Rajakumar et al. (2007). For more clarity, we adopt the same notations and definitions.

So, this problem deals with assigning n jobs to m identical parallel machines. Each job has a deterministic integer processing time p_j where $j=1..n$. A machine can process only one job at a time and no preemption is allowed. The mathematical formulation is given below.

$$\text{Min } W_{\max} \quad (1)$$

$$\sum_{j=1}^n X_{jk} = 1, \forall k = 1..m \quad (2)$$

$$W_{\max} - \sum_{j=1}^n P_{jk} \times X_{jk} \geq 0, \forall k = 1..m \quad (3)$$

$$W_k \leq W_{\max}, \forall k = 1..m \quad (4)$$

$$W_k \geq 0, \forall k = 1..m \quad (5)$$

$$X_{jk} \in \{0,1\} \quad (6)$$

Such as:

n	Total number of jobs
m	Total number of machines
j	Job index
k	Machine index
W_k	Workload of machine k
$W_{\max}$	Maximum workload
P_{jk}	Processing time of job j on machine k
X_{jk}	Binary variable equal to 1 if job j is assigned to machine k and 0 otherwise

In the model described above, equation (1) is the objective function. It concerns the minimization of the maximum workload on the parallel machines. Equation (2) indicates that each job is scheduled only once on a machine. Equation (3) is used to ensure that the difference in workloads is greater than or equal to zero. Equation (4) defines the maximum workload. Equation (5) ensures that workloads are non-negative. Finally equation (6) indicates that X_{jk} are binary variables.

2.1 Analysis of the mathematical model

The first remark concerns the first constraint (equation 2). This constraint does correspond to its definition. In fact, as it is expressed by Rajakumar et al. (2007), each machine k process only one job j which is not correct. So, to insure that each job is scheduled only once on a machine. This constraint should be corrected as follows.

$$\sum_{k=1}^m X_{jk} = 1, \forall j = 1 \dots n \quad (7)$$

The second remark is that equations (4) and (5) in this formulation are trivial. In fact, these constraints are redundant with the constraint expressed by equation (2).

Since the workload of a machine k is defined as the sum of processing time of jobs assigned to it, we can write:

$$W_k = \sum_{j=1}^n P_{jk} \times X_{jk} \geq 0, \forall k = 1 \dots m \quad (8)$$

By replacing W_k in equation (4) we notice that it is the same as equation (3). This result is logical and predictable the maximum workload is logically greater than or equal to the other workloads. That means also that all the workloads are lower than or equal to the maximum workload. Concerning equation (5), it is not necessary to specify that the workloads are positive. This is ensured by their definition as the sum of processing times as indicated in equation (8).

The third remark is that this formulation is not directly utilizable because the processing time P_{jk} of job j on machine k , used in equation (3), is not mathematically defined. In fact, P_{jk} is not a parameter because we do not already know if job j is assigned to machine k . This correlation can be expressed as follows:

$$P_{jk} = \begin{cases} p_j & \text{if job } j \text{ is assigned to machine } k \\ 0 & \text{otherwise} \end{cases}$$

Based on these remarks the corrected mixed integer programming model proposed by Rajakumar et al. (2007) can be expressed as follows.

$$\text{Min } W_{\max} \quad (9)$$

$$\sum_{k=1}^m X_{jk} = 1, \forall j = 1 \dots n \quad (10)$$

$$W_{\max} - \sum_{j=1}^n P_{jk} \times X_{jk} \geq 0, \forall k = 1 \dots m \quad (11)$$

$$X_{jk} \in \{0, 1\} \quad (12)$$

2.2 Analysis of the other criteria

The first performance measure or criterion is the relative percentage of imbalances (*RPI*) in workloads of all machines defined by Rajakumar et al. (2004, 2007). This performance measure is also called percentage of deviation from upper bound defined as:

$$RPI = \frac{1}{M} \times \sum_{k=1}^m \frac{W_{\max} - W_k}{W_{\max}} \quad (13)$$

In the case of identical parallel machines, this criterion depends solely of the maximum completion time criterion which has been proven to be not optimal in workload balancing problem. So, minimizing the relative percentage of imbalances criterion is the same thing as minimizing the maximum of completion times.

Remark 1.

In the case of identical parallel machine, the relative percentage of imbalances (*RPI*) depends solely of the maximum of workloads $W_{\max}$.

Proof.

Considering the definition of the *RPI* criterion given by equation (13), we have:

$$\begin{aligned}
RPI &= \frac{1}{M} \times \sum_{k=1}^m \frac{W_{\max} - W_k}{W_{\max}} \\
&= \frac{1}{m \times W_{\max}} \times \sum_{k=1}^m W_{\max} - W_k \\
&= \frac{\sum_{k=1}^m W_{\max}}{m \times W_{\max}} - \frac{\sum_{k=1}^m W_k}{m \times W_{\max}} \\
&= 1 - \frac{\sum_{k=1}^m W_k}{m} \times \frac{1}{W_{\max}}
\end{aligned}$$

Knowing that,

$$\frac{\sum_{k=1}^m W_k}{m} = \frac{\sum_{j=1}^n p_j}{m} = \mu, \mu \text{ is a constant}$$

So, the relative percentage of imbalances RPI can be written as a function of $W_{\max}$ as follows.

$$RPI = 1 - \frac{\mu}{W_{\max}}$$

The second criterion called normalized sum of square for workload deviations ($NSSWD$) has been introduced by Ho et al. (2009). This criterion is based on the sum of squares principle known in measuring variability in statistics and serves as a precise measurement criterion and it is defined as follows.

$$NSSWD = \frac{\sqrt{\sum_{k=1}^m (W_k - \mu)^2}}{\mu}, \text{ such as : } \mu = \frac{\sum_{k=1}^m W_k}{m} = \frac{\sum_{j=1}^n p_j}{m}$$

For each machine k the square error is given by $(W_k - \mu)^2$. It is easy to establish that the sum of square for workloads deviations $\sum_{k=1}^m (W_k - \mu)^2$ depends directly on the sum of square completion times or workloads.

$$\begin{aligned}
\sum_{k=1}^m (W_k - \mu)^2 &= \sum_{k=1}^m (W_k^2 - 2 \times \mu \times W_k + \mu^2) \\
&= \sum_{k=1}^m W_k^2 - 2 \times \sum_{k=1}^m W_k + \sum_{k=1}^m \mu^2 \\
&= \sum_{k=1}^m W_k^2 - 2 \times \sum_{k=1}^m W_k + m \times \mu^2
\end{aligned}$$

By considering that $\sum_{k=1}^m W_k = m \times \mu$, the sum of square for workloads deviations can be written as follows:

$$\sum_{k=1}^m (W_k - \mu)^2 = \sum_{k=1}^m W_k^2 - m \times \mu^2$$

Since $m \times \mu^2$ is a constant, then minimizing the sum of square for workloads deviations is the same as minimizing the sum of squares of completion times W_k^2 . Recently, Cossari et al. (2012) have considered this criterion together with other criteria, turns out to be closely related to some standard dispersion measures. The authors have proposed a novel algorithm that can effectively solve the minimal ($NSSWD$) scheduling problem in a limited computing time.

3. A New Mixed Integer Linear Programming Model

The model proposed does not obtain the optimal utilization of machines as mentioned in Rajakumar et al. (2007). This will be illustrated by the counter example presented in the next section. To obtain the optimal repartition of the workload among the parallel machines minimizing the maximum workload is not sufficient.

Therefore, the optimal formulation should consider simultaneously the maximum and the minimum workloads. In fact, workload balancing problem should be defined as the minimization of the difference between the maximum and the minimum workloads. In other words, the difference between the workload of the bottleneck machine and the workload of the fastest machine. Based on this definition, we propose the following mathematical model.

$$\text{Min } (W_{\max} - W_{\min}) \quad (14)$$

$$\sum_{k=1}^m X_{jk} = 1, \forall j = 1 \dots n \quad (15)$$

$$W_{\max} - \sum_{j=1}^n P_{jk} \times X_{jk} \geq 0, \forall k = 1 \dots m \quad (16)$$

$$W_{\min} - \sum_{j=1}^n P_{jk} \times X_{jk} \leq 0, \forall k = 1 \dots m \quad (17)$$

$$X_{jk} \in \{0, 1\} \quad (18)$$

Equation (14) is the new objective function. In addition to classical constraints, equation (17) ensures that the minimum workload is smaller than or equal to the other workloads.

4. Illustrative Numerical Example

In this section, we present a counter example to show that the first formulation based on the minimization of the maximum workload does not provide the optimal utilization of the parallel machines contrary to the proposed model. We consider 30 jobs to be assigned to 10 machines. The processing times of the job are given in table 1. For performances measures, we consider both relative percentage of imbalances (*RPI*) in workloads of all machines defined by Rajakumar et al. (2004, 2007) and normalized sum of square for workload deviations (*NSSWD*) introduced by Ho et al. (2009).

Table 1: Data of numerical example

<i>Job j</i>	<i>p_j</i>	<i>Job j</i>	<i>p_j</i>	<i>Job j</i>	<i>p_j</i>
1	3	11	10	21	11
2	6	12	8	22	17
3	2	13	22	23	26
4	4	14	11	24	16
5	5	15	8	25	7
6	7	16	26	26	23
7	9	17	14	27	15
8	13	18	6	28	18
9	4	19	17	29	15
10	12	20	27	30	13

We calculate also the absolute value of workload imbalance ($W_{\max} - W_{\min}$) obtained by each model. Based on this numerical example, we can conclude that the first formulation proposed by Rajakumar et al. (2004, 2007) does not provide the optimal workload repartition. In fact, this model provides a maximum relative percentage of imbalances (*RPI*) of 7.89 % where the second formulation provides a scheduling with maximum *RPI* equal to 2.63%. The normalized sum of square for workload deviations (*NSSWD*) performance measure confirms this tendency. In fact, the first model presents a normalized sum of square for workload deviations equal to 0.29 while that of the second model is equal to 0.04. In terms of absolute value of the workload imbalance the first model presents a workload imbalance equal to 3 (38-35) and the second one a workload imbalance equal to 1 (38-37). The solutions obtained by the two formulations are illustrated by figure 1.

Table 2: Obtained results for the illustrative example

Machine	W_k (MIP 1)	RPI_k (MIP 1) %	$NSSWD_k$ (MIP 1)	W_k (MIP 2)	RPI_k (MIP 2)	$NSSWD_k$ (MIP 2)
1	38	0.00	0.25	37	2.63	0.25
2	35	7.89	6.25	37	2.63	0.25
3	38	0.00	0.25	37	2.63	0.25
4	38	0.00	0.25	38	0.00	0.25
5	38	0.00	0.25	38	0.00	0.25
6	38	0.00	0.25	38	0.00	0.25
7	37	2.63	0.25	37	2.63	0.25
8	37	2.63	0.25	38	0.00	0.25
9	38	0.00	0.25	38	0.00	0.25
10	38	0.00	0.25	37	2.63	0.25
-	$W_{max} - W_{min}$ = 3	RPI_{max} = 7.89	$NSSWD$ = 0.29	$W_{max} - W_{min}$ = 1	RPI_{max} = 2.63	$NSSWD$ = 0.04

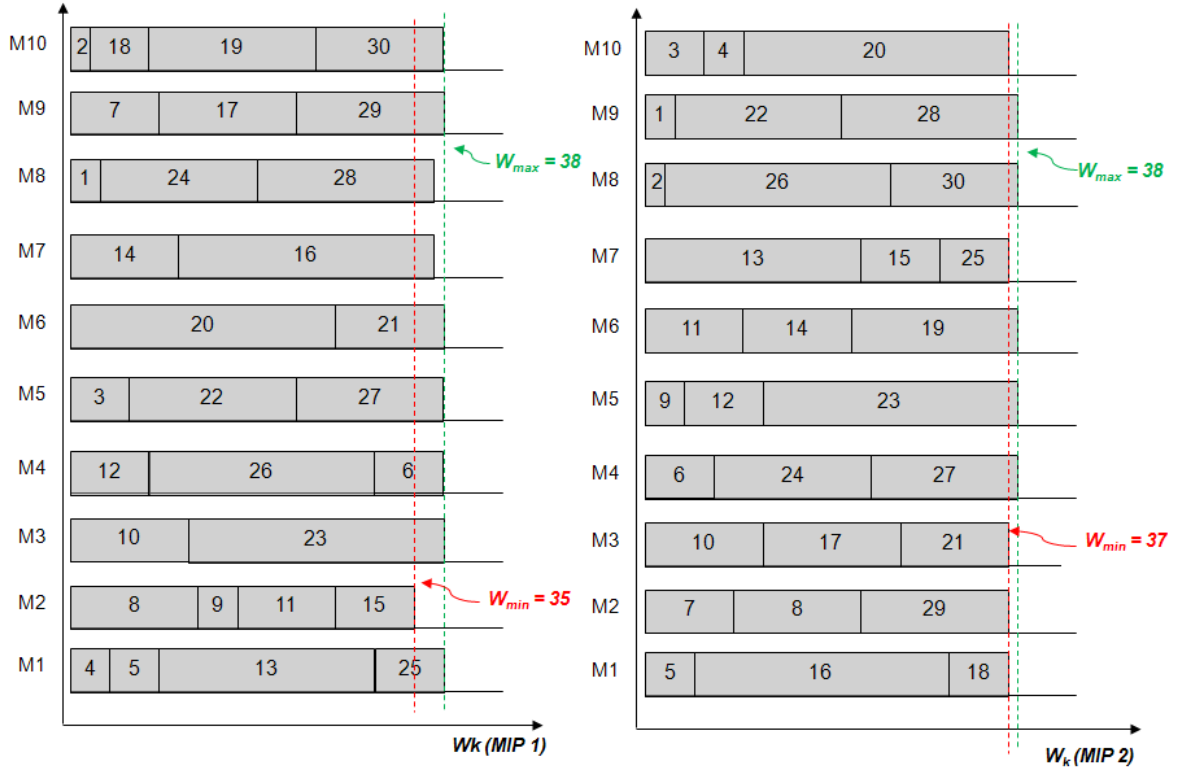


Figure 1: The two solutions obtained by model 1 and model 2

A larger experimental study has been proposed by Ouazene et al. (2013). In fact, based on this formulation, we have presented a comparative study and provided optimal solutions for different test instances issued from the literature.

5. Conclusion

In many organizations, distributing workload as equally as possible among a group of persons or machines is a real challenge and sometimes it is considered as a strong constraint especially when it concerns human resources. That is why workload balancing is an important optimization criterion. In this paper, we study the different criteria proposed in the literature to deal with this problem in the case of identical parallel resources. Based on this study, we have established that the workload balancing problem can be formulated as a problem of minimizing the difference

between the workload of the bottleneck machine and the workload of the fastest machine. A possible extension of this research is to consider this new formulation of the workload imbalance as new fitness functions of the approximate algorithms proposed in the literature.

References

- Cossari, A., Ho, J.C., Paletta G. and Ruiz-Torres, A.J., a new heuristic for workload balancing on identical parallel machines and a statistical perspective on the workload balancing criteria. *Computers & Operations Research*, 39(7):1382 – 1393, 2012.
- Ho, J.C., Tseng, T.L., Ruiz-Torres, A.J. and L'opez, F. J., minimizing the normalized sum of square for workload deviations on m parallel processors. *Computers & Industrial Engineering*, 56(1):186–192, 2009.
- Keskinturk, T., Yildirim, M. B., and Barut, M., an ant colony optimization algorithm for load balancing in parallel machines with sequence-dependent setup times. *Computers & Operations Research*, 39(6):1225 – 1235, 2012.
- Ouazene, Y., Hnaïen, F., Yalaoui F. and Amodeo, L., the joint load balancing and parallel machine scheduling problem. In Bo Hu, Karl Morasch, Stefan Pickl, and Markus Siegle, editors, *Operations Research Proceedings 2010*, Operations Research Proceedings, pages 497–502. Springer Berlin Heidelberg, 2011.
- Ouazene Y., Yalaoui F., Chehade H., and Yalaoui A., workload balancing in identical parallel machine scheduling using a mathematical programming method. *International Journal of Computational Intelligence Systems*, 6(Supplement 2):58 – 67, 2013.
- Raghavendra, B. V. and Murthy, A. N. N., some solution approaches to reduce the imbalance of workload in parallel machines while planning in flexible manufacturing system. *International Journal of Engineering Science and Technology*, 2(5): 724 – 730, 2010.
- Raghavendra, B. V., Murthy, A. N. N. and Jayaram, M., job sequence to minimize the workload imbalance on parallel machines through genetic algorithm. *International Journal of Engineering Science and Technology*, 3(1): 204 – 210, 2011.
- Raghavendra, B. V., Murthy, A. N. N. and Rao, N. B., some solution approaches to reduce the imbalance of workload in parallel machines while planning in flexible manufacturing system through genetic algorithm. *International Journal of Engineering Science and Technology*, 2(1): 313 – 324, 2010.
- Rajakumar, S., Arunachalam, V.P. and Selladurai, V., workflow balancing strategies in parallel machine scheduling. *International Journal of Advanced Manufacturing*, 23:366 –374, 2004.
- Rajakumar, S., Arunachalam, V.P. and Selladurai, V., workflow balancing in parallel machines through genetic algorithm. *International Journal of Advanced Manufacturing Technology*, 33:1212 – 1221, 2007.
- Yildirim, M.B., Duman E., Krishna, K. and Senniappan K., Parallel machine scheduling with load balancing and sequence dependent setups. *International Journal of Operations Research*, 4(1):42– 49, 2007.

Biography

Yassine Ouazene is currently a PhD-student in the Industrial Systems Optimization Laboratory at the University of Technology of Troyes in France. His PhD thesis concerns the optimization of production lines design. He received his engineering degree in Industrial Engineering in 2009 from the Polytechnical School of Algiers (Algeria), followed by his master's degree in Systems Optimization and safety in 2010 from the University of Technology of Troyes (UTT, France).

Farouk Yalaoui obtained his degree in Industrial Engineering from the Polytechnical School of Algiers in 1995, his master's degree in Industrial System Engineering from Polytechnics Institute of Lorraine in 1997, his Ph.D. degree in Production Management from the Troyes University of Technology (UTT) in 2000 and followed by a Habilitation à Diriger les Recherches from Compiegne University of Technology in 2006. He is currently a full Professor and head of Industrial Systems Laboratory. His research topic focuses on scheduling problems, system design, operations research, modeling, analysis and optimization of logistic and production systems, reliability and maintenance optimization. He is author of more than 180 publications or communications with one book, 7 book chapters and 38 articles in journals. He also published more than 110 papers in conference proceedings. He is a member of several international conference committees and International journal Boards.

Hicham Chehade received his engineering degree in Industrial Systems Engineering and his master's degree in Systems Optimization and safety in 2005, followed by his Ph.D. degree in Systems Optimization and safety from the

University of Technology of Troyes (UTT, France) in 2009. He is currently an associate professor in the Industrial Systems Optimization Laboratory at the University of Technology of Troyes in France. His research fields concerns mainly the optimization of production, assembly and robotic lines design. He has also other research topic such as scheduling problems, operations research, modeling and simulation.

Alice Yalaoui received her engineering degree in Industrial Systems Engineering in 2001, followed by her Ph.D. degree in Systems Optimization and safety from the University of Technology of Troyes in 2004. She is currently an associate professor at the Industrial Systems Optimization Laboratory at the University of Technology of Troyes in France. Her research fields concerns mainly the analysis and optimization of logistic and production systems, reliability and maintenance optimization.