Agent-based approach for cooperative scheduling

J. Behnamian

Department of Industrial Engineering, Faculty of Engineering, Bu-Ali Sina University, Hamedan, Iran
Behnamian@basu.ac.ir

Abstract

This paper studies the multi-factory production (MFP) network scheduling problem where a number of different individual factories join together to form a MFP network, in which these factories can operate more economically than operating individually. However, in such network that is known as virtual production network with self-interested factories with transportation times, each individual factory usually focuses on self-benefits and does not care much about the others within the network. We first describe the realistic features which incorporate in problem definition. Then two different variants of the problem are considered. In the first case, we propose approximation algorithms with the best achievable theoretical guarantee in three cases: (i) all factories are interested in the makespan, (ii) all factories are interested in the sum of completion times, and (iii) the case in which among all factories, some factories are interested in the sum of completion times and the others are interested in the makespan. Furthermore, with considering the transportation, we model the problem as a mixed integer linear programming to minimize the makespan and total completion time and solve it by CPLEX solver to obtain Pareto solutions by applying modified $\varepsilon$-constraint approach. Experiments show that this procedure is capable of producing good results to approximate the efficient set.

Keywords: Distributed scheduling, approximation algorithms, $\varepsilon$-constraint approach, pareto solutions, virtual corporation

1- Introduction

Scheduling problems are applied in various systems, since it is necessary to distribute the work among several entities. Formally, the scheduling problem as quoted in Pinedo (2008) is defined as ”Scheduling concerns the allocation of limited resources to tasks over time. It is a decision-making process that has as a goal the optimization of one or more objectives.” Since the first scheduling paper appeared in 1954, many variants of the basic scheduling problem have been formulated by differentiating between machine environments, side-constraints, and objective functions. Figure 1 shows the comparison between the single factory and virtual network production’s system with respect to three important factors to examine the production system’s capabilities.
Most of the researches in production scheduling are only concerned with the optimization of a single factory. However, in the real world factories are interested to reduce their production costs and improve their production quality via participation in production activities with other professional factories distributed in different geographical places. As we see in Figure 1, many variations of collaboration are possible and factories will choose the specific type of partnership according to their individual needs.

**Fig. 1.** Comparison between the single factory and virtual network production capabilities (Bartlett and Ghoshal (1998))

**Fig. 2.** Classification of cooperation concepts (Wiendahl et al. 1999)
Generally there are two ways to structure a collaborative manufacturing network which are defined as follows (Johansen et al. 2005):

- Vertical relations (or series relations), where the factory has a relationship with its suppliers or customers.
- Horizontal relations (or parallel relations), where the factory as a competitor produces similar goods on the same level.

Nowadays, many factories move away from vertical integration and are interested to perform horizontal cooperation with other factories. A further development in the continuing cooperative trend amongst manufacturing factories is the formation of production networks (Wiendahl et al. 1999). Lin et al. (2004) stated that in the today’s industries, manufacturing generally involves the cooperation between several companies either formed as a virtual corporation or extended company. In fact, these networks establish a new type of relations between independent factories to manage and share the network production resources in parallel structure. Virtual production network (VPN) is a kind of the distributed system in which the smaller parts of the scheduling problem are solved by local decision makers who possibly have different objectives. This network establishes a new type of the horizontal collaboration and relations between independent factories and even competitors that establish occasional collaborations on the production order they cannot take on individually. After the order has been finished, factories as partners may go to their own ways again. VPN has four advantages (Vilana and Monroy 2009):

(i) improvement in operations, (ii) access to new markets in different geographical places, (iii) diversification of the financial risks and (iv) access to new technologies.

Generally, one can say, the distributed activities, if managed properly, can reduce costs, increase flexibility and efficiency (Shi and Gregory 1998). Accordingly, the result is a highly flexible system characterized by low barriers to entry and exit, geographic flexibility, low costs, rapid technological diffusion, high diversification through contract manufacturers and exceptional economies of scale.

In such networks, some issues are important e.g. where to assign production so as to be must responsive, how to reduce transportation costs and how to reduce the complexity of network. As in Figure 1, in a production network, partners cooperate over a longer period since their integration is supposed to be more intensive but in VPNs, the relations are short-term corporations (Tuma 1998). It is worthwhile to notice that another characteristic distinguishing production network from virtual network is the availability of competence amongst partners in VPNs. Several industries such as aeronautics (Shi and Gregory 2003), electronics (Shi et al. 2005) or the automotive industry (Sturgeon and Florida 1999) encounter virtual networks.

The recent remarkable attention in distributed manufacturing management in both academia and the industry has demonstrated the significance of such system. During last years, many scholars have considered VPNs phenomenon from different perspectives (Monroy and Vilana Arto, 2010). Since in this respect the simultaneous scheduling of the factory in such production network has not been considered in any researches, the current paper considers the distributed scheduling (DS) for a virtual network in which several heterogeneous factories disperse geographically in different places and each of the network’s factories has a parallel machines environment. In other words, in this problem, production processes performed in different factories are similar but the machines in different factories may have different speeds and different objective functions.

For this complicated production network, proposing an efficient scheduling technique is critical to react quickly to market changes. The difficulties faced in this problem are not only scheduling the jobs in a most favorable way, but also assigning the best factory to the jobs. Therefore, for generating the schedule in the distributed manufacturing environment, these two sub-problems must be simultaneously considered. Due to these reasons, scheduling in a multi-factory environment has become much more complex than traditional one. In this paper, we assumed that each factory in the production network has a parallel machine environment. Since the DS reduces to the regular scheduling problem if the number of factory (F) is one, and this latter problem as a classic parallel machine with makespan objective even with two machines is a NP-hard problem, we can easily conclude that the DS is also a NP-hard problem and consequently, no optimal polynomial time algorithms can be designed unless P=NP. Although there are
many approximation algorithms which solve the parallel machine scheduling problem on single factory (Diedrich et al. 2010), there is not any approximation algorithm for multi-factory scheduling with parallel machines. In this paper, following Saule and Trystram (2009), we are interested to design the approximation algorithms with the best achievable approximation ratios. Furthermore, in this paper we propose a mathematical modeling. To the best of our knowledge, this paper presents for the first time a mathematical modeling and approximation algorithms for the multiple-factory production scheduling with self-interested heterogeneous factories. Because different interests often exist among self-interested factories, a multi-objective approach for the distributed scheduling is clearly requested. Therefore, to obtain Pareto solutions for the problem, we also make benefit of the $\varepsilon$-constraint technique.

The paper is organized as follows. Section 2 gives the literature review of some previous results. In Section 3, we present properties of the scheduling problem. Section 4 presents the framework of polynomial approximation algorithms in details. Mathematical modeling and its numerical results is given in Section 5. Finally, Section 6 concludes the paper with some remarks and recommends some future research areas.

2- Literature review

During the late 1970s’ and early 1980s’, the most of empirical researches reveal that not only the single factory needs to manage, but multi-factory scheduling should also be considered. However, the literature review indicates that the research during this period mainly was concerned with single factory scheduling.

In this respect and based on empirical research in microelectronic industry’s production networks, Flaherty (1989) concluded that production network geographically has dispersed facilities and have shared common infrastructure mechanism which led it to a synergy advantage in the production. Shi and Gregory (1998) considered a relation of entities in multi-factory network as a matrix connection, where each node (i.e. factory) affects the other nodes and hence cannot be managed in isolation. Khurana and Talbot (1999) investigated that how each factory in the production network could influence each other in a complex structure. In order to increase key capability, Loeser (1999) proposed an approach for expansion of factories via cooperation with other factories. Li et al. (2000) were the first to name global manufacturing virtual networks by proposing a strategic positioning model for these factories based on three vectors: globalization, strategic alliances and value and supply chains. Fleury and Fleury (2003) give reasons that individual efficiency in today’s new competitive market is not enough and it is necessary to be connected to the groups of firms which have efficient performance.

First studies about multi-plant scheduling started by the early 1980s. In the first study on joint scheduling of production and distribution in a complex network, Williams (1981) compared a dynamic programming based algorithm with several existing heuristics. The objective of the problem was to minimize average production and distribution costs. In that paper, the author assumes a constant demand rate and this limits the applicability of paper. Vercellis (1999) considered a capacitated master production planning and capacity allocation problem in a multi-plant chain in his/her mixed integer linear programming (MILP) model. Neiman and Lesser (1996) proposed a search algorithm namely repair process and the set of heuristics for job shop scheduling problem with cooperative distributed agents. In that research, the authors suppose that the agent has not a complete view of the resources availability, or of the jobs to be scheduled. The solving algorithm proposed in this paper considered all these issues in search. Sandholm (2000) studied the distributed scheduling problem and proposed a levelled commitment contracting protocol that helped to self-interested agents to adapt themselves with possible future events. Karatza (2001) presented a method which included a simulation-based modeling to examine machine performance under a variety of workloads in distributed scheduling. In that paper, the author scheduled a multi-factory production network in which half of the total machines had double the speed of others.

Jia et al. (2007) focused on distributed scheduling problems with the job shop environment. They used genetic algorithm (GA) to optimize different criteria. However, they solved several problems with single 6-job, 3-factory instances in which each factory has only four machines, corrected it and extended the
modeling concept to the problem which delivery times were considered for jobs. More recently, Jia et al. (2002) refined and improved the previous GA in the same problem. Chan et al. (2007) proposed GA with dominant genes approach for distributed scheduling with the job shop environment in factories with respect to the criterion minimizing the deviation of solution obtained. By simulation runs, they also explored the significance of considering maintenance in that scheduling environment.

Based on the intelligent and agile characteristics of a multi-agents system and the optimistically capacity, Zhang et al. (2008) presented a multi-agents based model for distributed job shop scheduling problem. In that research, the authors proposed a distributed scheduling mechanism integrating multi-agents GA and dynamic scheduling strategies based on multi-agent negotiation. You and Grossmann (2008) formulated the multi-plant industrial networks scheduling as a mathematical modeling. Due to the difficulty of solving the mathematical model directly, they developed a new optimization method namely factory wide optimization to solve the combined production/distribution scheduling problem. Chung et al. (2009) studied the job shop distributed production networks scheduling problem for the case where perfect and imperfect maintenances are considered during the process of distributed scheduling. This study assumes that maintaining the reliability in an acceptable level is required for each entity in the network. They adapted the well-known GA approach to minimize the makespan of jobs. To minimize the maximum completion time, Naderi and Ruiz (2010) proposed six different mixed integer programming models for the distributed permutation flowshop scheduling problem. They also proposed several heuristics and a variable neighborhood descent for the problem. Zhang and Gen (2010) proposed a multi-objective GA approach to solve process planning and scheduling problems. The authors showed the efficacy of their proposed algorithm was better than non-dominated sorting GA II (NSGA-II).

3- Problem description

In this paper we are interested in classical settings, that is, all data in all problems used are known deterministically, when scheduling is undertaken; at any time, each job can be processed by at most one machine; each machine can process only one job; there are \( n \) independent jobs that are available at time 0; machines are available at all times if are not busy, with no breakdown; and a job once started on the machine must be completed on it without interruption. By \( f = \{1, \ldots, F\} \) we denote the set of independent factories formed the production network, which have a specific local job cluster. Each factory \( f \) owns a cluster \( CL^f \). Each factory has \( m^f \) identical parallel machine that all of them have speed \( v^f \). The set of all jobs produced by factory \( f \) is denoted by \( n^f \) that can be own cluster or migrated job from other factories’ cluster, with elements \( j_{i,q}^f \). By \( j_{i,q}^f \) we denote the job belong to set cluster factory \( f \) that executed on factory \( q \). If \( f = q \), the job is executed locally, otherwise it is migrated and transportation time between factory \( f \) and \( q \) must be considered. If the factory \( f \) executed all local job, then \( CL^f = n^f \). With this definition \( n = \sum_{f=1}^{F} n^f = \sum_{f=1}^{F} CL^f \) is the number of jobs that must be scheduled on \( F \) factories. As mentioned above, factories in virtual networks plan to maximize their own profit; so we need to define two objective functions, i.e., makespan and total completion time that each factory’s job cluster belongs to one of them. For this we define two symbols R1 and R2 that R1 contains all factories and their corresponding jobs which their objective function is total completion time and in similar definition, R2 indicates the factories’ jobs which are interested in the makespan.

In this paper, the following basic factors must be taken into account when a production network is scheduled. 

**Heterogeneous factories.** In a distributed manufacturing, the factories have two types of structure: parallel structure and series structure. In a parallel structure each factory may produce the finished products, but in series structure some intermediate factories may be supplied with other factories (Chung et al. 2009).

In the parallel structure, the factories can be homogeneous or non-homogeneous. In multi factory production network, each factory can be considered as an individual entity which has different efficiency and is subject to different constraints, for example, machine advances, worker cost, tax, close to suppliers,
and transportation facilities, etc. In our problem, we assume that each factory has identical machines and different factories have different speeds (Sturgeon and Florida 1999). Systems of such multiple-factory manufacturing cells have many real-life applications, for example, it is common in the semiconductor manufacturing that the newer and modern machines have faster processing speeds in comparison with existing machines (Dessouky 1998). The majority of literature only has been dedicated to homogeneous factories. Since normally the real-world distributed systems are heterogeneous, we consider the case in which factories are non-homogeneous and each of them has specific machine’s speed.

**Self interested factories.** Basically, there are two types of the distributed scheduling (Chan and Chung 2007). In the first type, to optimize the same objective function all factories which belong to the same factory, operate cooperatively. They may lose the individual’s benefit to improve the global objective of network as a whole. In second one which is named a virtual network, some set of orders cannot be met, because there is not a factory of sufficient production potential, so some factories joint together and create a virtual production network in order to satisfy these production orders. Therefore, in the virtual network a number of different individual factories join together to form a multi-factory production network, in which these factories can operate more economically than operating individually. This network has a larger production capacity and many different resources. Therefore, it can meet many more production orders together than each of factories separately. In such a network, each individual factory usually focuses on self-benefits. In other words, it wants to maximize its own profit while not caring about the other factories’ profits within the virtual network. In this paper, the second type of self-interested factories will be considered. In such cases, a factory is willing to accommodate other factories’ jobs only for compensation.

Next two sections are devoted to two different cases. In the first case, we considered self-interested heterogenous factories and proposed several scheduling algorithms for different variants of problem with respect to the variety of self-interested factories’ objective functions. In the last one, two groups of factories without caring about the individual objective of each factory in the same group are considered. In this case, for modeling and solving the new problem we assume that it is possible the jobs migrate from their original factory to other factories but a transportation time is incurred.

4- **Case 1: Approximation results for production network without transportation**

Assume the value of $Z$ achieved in a given schedule $\sigma$ with respect of criterion A is denoted by $Z_A(\sigma)$. Consider the problem of simultaneous minimizing the set $(Z^1, ..., Z^F)$. Each solution $\sigma$ has a corresponding value in the objective space $(Z^1_A(\sigma), ..., Z^F_A(\sigma))$ with respect of criterion A. We say that a schedule $\sigma$ is nondominated with criteria A and B, if there is no schedule $\bar{\sigma}$ such that $Z_A(\bar{\sigma}) \leq Z_A(\sigma)$, $Z_B(\bar{\sigma}) \leq Z_B(\sigma)$ and at least one of the two inequalities is strict (Agnetis et al. 2004). As mentioned earlier, all local single factory scheduling problems in production network are NP-hard and no optimal polynomial time algorithms can be designed. Thus, we are interested in polynomial approximation algorithms. In this section, we are interested to design the less achievable approximation ratios. Formally, a scheduling algorithm $algo$ has a performance ratio $(\rho^1, ..., \rho^F)$ if for all valid schedules $\sigma$ and for all factories $F; Z^f_f(algo) \leq \rho Z^f(\sigma)$ where $Z^f$ is the objective of factory $f$.

Our main goal in this section is to propose algorithms with the best achievable theoretical guarantee in three cases: (i) all factories are interested in the makespan, (ii) all factories are interested in the sum of completion times, and (iii) the case in which $F'$ factories are interested in the sum of completion times and $F'' = F - F'$ factories are interested in the makespan.

4-1- **Numerical example**

This sub-section considers two random generated examples with self-interested factories. With this example, we are going to show how, even if only one objective is concerned, a non-dominated solution will be obtained. Table 1 shows the processing time of 12 jobs on 2 factories in two different regions R1 and R2. Assume that the machine’s speeds in factories F1 and F2 are 1 and 2, respectively.
Table 1. Processing time of jobs in factories 1 and 2

<table>
<thead>
<tr>
<th>Processing times</th>
<th>Jobs of R1</th>
<th>Jobs of R2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1  2  3  4  5  6</td>
<td>7  8  9  10  11  12</td>
</tr>
<tr>
<td>( p_1^1 )</td>
<td>40 50 10 40 30 50</td>
<td>50 40 10 30 40 20</td>
</tr>
<tr>
<td>( p_1^2 )</td>
<td>20 25 5 20 15 25</td>
<td>25 20 5 15 20 10</td>
</tr>
</tbody>
</table>

Figure 3 shows the resultant Gantt chart for two examples.

![Gantt chart for Solution 1](image1)

(a) Solution 1

![Gantt chart for Solution 2](image2)

(b) Solution 2

Fig. 3. The Gantt chart of two numerical examples with two selfish factories

The objective value for each \( R \) of two solutions is shown in Table 2.

Table 2. Numerical example with two competing factories

<table>
<thead>
<tr>
<th>Solution</th>
<th>( R=1 (C_{\text{max}}) )</th>
<th>( R=2 (C_{\text{max}}) )</th>
<th>( R=1 (\sum C) )</th>
<th>( R=2 (\sum C) )</th>
<th>( R=1 (\sum C) )</th>
<th>( R=2 (C_{\text{max}}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>70</td>
<td>90</td>
<td>235</td>
<td>260</td>
<td>235</td>
<td>90</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>90</td>
<td>230</td>
<td>350</td>
<td>230</td>
<td>90</td>
</tr>
</tbody>
</table>

As can be seen in this table, in the case which all factories have makespan objective, the solution 2 dominated the solution 1. We also see when two factories have total completion time (TCT) objective,
both solutions are nondominated schedules and finally in case 3 which factory 1 is interested to TCT but factory 2 is interested to minimize makespan, solution 2 is nondominated solution.

4-2- Approximation algorithm

Before the approximation bounds are introduced, two useful propositions are proposed on the single machine problem. They will serve to compute approximability bounds and basically construct the concept that all known bounds on the best approximation ratios for the single factory case are valid for multi factory production network.

Let \( M \) denote the sum of machines in all factories and \( p_i \) denote the processing time of job \( i \). We also show the completion time of job \( i \) in factory \( f \) with \( C_i^f \) and consequently, we have \( C_{\text{max}} = \max_{f \in \{1,...,F\}} \max_{i \in \{1,...,n\}} C_i^f \).

**Proposition 1.** The distributed scheduling of the self-interested entities with makespan objective function cannot be approximated with a performance ratio less than \((1,2,...,F)\) on \( M = 1 \) with speed \( v \).

**Proof.** In this case, all factories are interested in the makespan. Consider the instance that each factory has only one job and \( M = 1 \) with speed \( v \) is available for scheduling \( n = F \) jobs. All jobs have \( p_i = 1 \). Obviously, for all jobs, the best makespan of each factory that can be achieved is \( \frac{1}{v} \). For any permutation with any efficient schedule, one factory will have a makespan of \( \frac{1}{v} \), another one will have a makespan of \( \frac{2}{v} \), and so on. Thus, it is impossible to obtain an algorithm that guarantees a ratio better than \((1,2,...,F)\).

Now consider the instance where at least one factory (factory \( q \)) has a job with \( p_i > 1 \). In such an instance, if factory \( q \) is the first factory in permutation which gets a performance ratio of 1, all other factories will get a performance ratio worse than \( F \). 

**Proposition 2.** The distributed scheduling of the self-interested entities with sum of completion times of unitary jobs cannot be approximated with a ratio less than \((F,...,F)\) on \( M = 1 \) with speed \( v \).

**Proof.** Let us consider that each factory owns \( n^f = r \) jobs with \( p_i = 1 \) that must be scheduled on \( M = 1 \) with speed \( v \). The optimal value for each factory \( f \) is \( \sum C_i^f = \sum_{i=1}^{r} i / v = \frac{r(r+1)}{2v} \). The sum of completion times of all jobs is the sum of objective values of all factories which can be calculated by the following equation.

\[
\sum C_i^f = \frac{Fr(Fr+1)}{2v}
\]

All factories have the same priority for their scheduling, thus no factory should obtain a better objective value than the others. Let us consider that factory \( f \) obtains the \( F \)th of the sum of completion times over all jobs, so we have \( \frac{\sum C_i^f}{F} = \frac{Fr(Fr+1)}{2v} = \frac{r(r+1)}{2v} \). We can now compute the minimum approximation ratio obtained by ensuring equity between the factories: \( \frac{\sum C_i^f}{\sum C_i^f} = \frac{r(Fr+1)/2v}{r(r+1)/2v} \) which is equal to \( F - \frac{F-1}{r+1} \). Now if \( r \to \infty \) then ratio tends to \( F \).

4-2-1- Competitive factories with \( C_{\text{max}} \) objective

As can be seen in the previous sub-section, for the case (i) in which all factories are interested in the makespan, no algorithm can guarantee a constant performance ratio better than \((1,2,...,F)\) for scheduling jobs on the single machine. The following algorithm is proposed for multi-factory scheduling on \( M \geq 1 \) machines in which all factories are self-interested and have \( C_{\text{max}} \) objective. In Algorithm 1 we
use $\rho$-approximation algorithm such as LPT (largest processing time) rule with $\rho = \frac{4}{3}$ for the single factory case.

**Algorithm 1:** $VDN(C_{\text{max}})$ algorithm

1. Given a $\rho$-approximation algorithm for the single factory on $M$ machines,
2. For each factory $f$, compute schedule in factory $f$ ($\sigma^f$) such as $C^f_{\text{max}}(\sigma^f)$,
3. Group jobs of each factory $f$ according to $\sigma^f$,
4. Schedule jobs of factory $f$ according to $\sigma^f$ between $\sum_{f' < f} C^f_{\text{max}}(\sigma^f) \cdot \sum_{f' \leq f} C^f_{\text{max}}(\sigma^f)$.

**Theorem 1.** Algorithm $VDN(C_{\text{max}})$ is a $(\rho, 2\rho, \ldots, F\rho)$-approximation algorithm of the distributed scheduling with self-interested factories with an unknown order that are interested in the makespan.

**Proof.** For this at first the validity of the schedule must be checked. In other words, it must be verified that the jobs of each factory $f$ are scheduled in the disjoint intervals of length $C^f_{\text{max}}(\sigma^f)$ according to $\sigma^f$.

Now the theorem can proved as follows.

$$C^f_{\text{max}} = \sum_{f' \leq f} C^f_{\text{max}}(\sigma^f).$$

If the factories are ordered by increasing values of $C^f_{\text{max}}(\sigma^f)$, inequality $C^f_{\text{max}} \leq f \cdot C^f_{\text{max}}(\sigma^f)$ holds true. Moreover, if we assume that $\sigma^f$ is generated by a $\rho$-approximation algorithm, we have $C^f_{\text{max}} \leq f \cdot \rho$ and Theorem is proved. 

Note that if LPT rule is used, the performance ratio of $VDN(C_{\text{max}})$ algorithm is $\left(\frac{4}{3}, \frac{8}{3}, \ldots, \frac{4F}{3}\right)$. Following Gantt chart shows the application of $VDN(C_{\text{max}})$ algorithm with some modification for the numerical example introduced in sub-section 4.1. In modified LPT, in each stage we assign the job with the largest processing time among the unscheduled to the fastest available factory in which the machine’s speed in factories F1, F2 are 1 and 2, respectively (see Figure 4).
4-2-2- Competitive factories with TCT objective

A single factory problem with $M = 1$ can be solved in polynomial time by SPT rule. However, for more than one factory, this problem is NP-hard. As we see in Proposition 2, no algorithm can guarantee a constant ratio better than $(F, \ldots, F)$ for the single machine case. For general form with $M$ machines with TCT objective, the basic proposed steps are designed as shown in Algorithm 2.

**Algorithm 2: $VDN(\sum C)$ algorithm**

1: Let $\sigma^1, \sigma^2, \ldots, \sigma^F$ be $F$ modified SPT schedules on $M$ machines. In this modification, we assumed that in each assigning, among unscheduled jobs the job with the smallest processing time is assigned to the available machine with the smallest speed.

2: Each schedule $\sigma^f$ can be seen as $M$ independent schedules denoted $\sigma^f_1, \sigma^f_2, \ldots, \sigma^f_M$.

3: For each machine $j$, schedules $\sigma^f_j$ of each factory $f$ are mixed by the algorithm greedy algorithm into $\sigma_j$. In this step, we construct the schedule $\sigma$ namely the greedy algorithm (GR) through sorting jobs in the non-decreasing order of $C^f_\text{max}(\sigma^f)$.

4: A global schedule on $M$ machines is constructed by executing $\sigma_j$ on the machine $j$ ($1 \leq j \leq M$).

In the modified SPT algorithm, in each factory we check the assignment the job to the machine which is available and compare the obtained results for all factories and the solution with smallest objective value is chosen.
Theorem 2. VDN ($\Sigma C$) Algorithm is a $(F,...,F)$-approximation algorithm for the case that all factories are interested in the sum of completion times.

**Proof.** At first, we analyze the single machine case before the general case with $M$ machines is proved. In single machine case initially we compute the final schedule of all factories’ jobs from the schedules of each of factory’s jobs. To do this, we use $\sigma^1, \sigma^2,...,\sigma^F$ notations to show the independent schedules for the jobs of each factory. We construct the schedule $\sigma$ namely *greedy algorithm* (GR) through sorting jobs in nondecreasing order of $C_i^f(\sigma^f)$.

In schedule $\sigma$ by applying the GR algorithm on single machine, inequality $C_i^f(\sigma) \leq FC_i^f(\sigma^f)$ holds true for all $f$ and $i \leq n^f$. Let us consider job $J_i^f$. On single machine, inequality $C_i^f(\sigma) \leq C_i^f(\sigma^f)$ implies $C_i^f(\sigma^f) \leq C_i^f(\sigma^f')$. Thus, for each factory $f'$, the set of job $J_i^f$ is scheduled before $C_i^f(\sigma^f)$ in $\sigma$ such that $\sum_{i',f'} J_i^f \leq C_i^f(\sigma^f)$. For each factory, there is one such set $J_i^f$ is completed when all jobs of $\cup_{i'} J_i^f$ are completed. Thus, $C_i^f(\sigma) = \sum_{i,f} \sum_{j \in f} p_i^f \leq FC_i^f(\sigma^f)$.

Using GR, it is possible to mix $F$ schedules and ensure that the completion time of each job in each schedule is not degraded by a factor greater than $F$. In particular, if a schedule is a $\rho$-approximation on the sum of completion times, then by mixing it with the $F - 1$ other schedules, the final schedule is a $(F\rho)$-approximation for this factory. In the case that all factories are interested in the sum of completion times with $M = 1$, algorithm GR$(SPT^1,...,SPT^F)$ is a $(F,...,F)$-approximation algorithm.

With this reasoning for the single machine case and with paying attention to the independency of machines in $M$ parallel machine case, the Theorem 2 is proved.

Figure 5 shows Gantt chart of applying VDN ($\Sigma C$) algorithm for the numerical example introduced in sub-section 4.1.

![Gantt chart](image)

**Fig. 5.** The Gantt chart of the solution produced by applying VDN ($\Sigma C$) algorithm

4-2-3- Competitive factories with $C_{\text{max}}$ or TCT objective

In this case among $F$ factories, $F'$ factories are interested in the sum of completion times and $F'' = F - F'$ factories are interested in the makespan. The idea in this sub-section is to incorporate techniques presented two cases (i) and (ii) into a single scheduling algorithm called VDN ($\Sigma C:C_{\text{max}}$). Consider the sub-instance of $F'$ factories interested in the makespan. Let $\sigma^{c_{\text{max}}}$ be the schedule generated by VDN ($C_{\text{max}}$) on this sub-instance. Algorithm 3 shows VDN ($\Sigma C:C_{\text{max}}$) procedure.
Algorithm 3: VDN ($\sum C_{\text{max}}$) algorithm

1: Build a schedule $\sigma^{C_{\text{max}}}$ for all factories’ makespan applying the only VDN($C_{\text{max}}$) algorithm.
2: For each factory $f$ interested in the sum of completion times, let $\sigma^f$ be a modified SPT schedule.
3: We modify GR algorithm in which each factory is assumed to have a same priority. We can get another algorithm which the factories do not have the same priority assumption. Let $\pi^1, \pi^2, \ldots, \pi^F$ be real positive values such that $\sum \pi^f = 1$. We build schedule $\sigma^\pi$ using the modified greedy algorithm (GR$\pi$) in which the jobs are scheduled in nondecreasing order of $\frac{c_i^f(\sigma^f)}{\pi^f}$ and whenever $\forall f, \pi^f = \frac{1}{F}$ then $\sigma = \sigma^\pi$. Now in this step apply GR$\pi$ for the factories which are interested in the sum of completion times, for which $\pi_{c_{\text{max}}} = \frac{F}{F}$ and $\pi^f = \frac{1}{F}$.

Theorem 3. VDN ($\sum C_{\text{max}}$) algorithm is a $(\frac{F}{F'}, \ldots, \frac{2F}{F'}, \ldots, F)$-approximation algorithm for the case in which $F'$ factories are interested in the sum of completion times and $F'' = F - F'$ factories are interested in the makespan.

Proof. For the proof, we assumed that all factories interested in the sum of completion times get a performance ratio of $F$. The corresponding proof is similar to the one of Theorem 2 and it is omitted.

Consider now factory $f$ interested in the makespan. Without loss of generality, we have $C_{\text{max}}(\sigma^{c_{\text{max}}}) \leq f^\rho C_{\text{max}}^\sigma$ (from Theorem 1) which means that $\forall i \leq n_f$, $C_i^f(\sigma^{c_{\text{max}}}) \leq f^\rho C_{\text{max}}^\sigma$.

In GR$\pi$ algorithm the completion times in $\sigma^\pi$ of jobs belonging to factory $f$ are not degraded by a factor greater than $\frac{1}{\pi^f}$. Applying the reasoning similar to which proposed in Theorem 2, it is straightforward to prove that using GR$\pi$ algorithm, in schedule $\sigma^\pi$ with $M = 1$, inequality $C_i^f(\sigma) \leq \frac{C_i^f(\sigma^f)}{\pi^f}$ for all $f$ and $i \leq n^f$ holds true. Due to this statement, all jobs of the factories interested in $C_{\text{max}}$ are not reduced by a factor greater than $\forall i \leq n^f$, $C_i^f(\sigma^{c_{\text{max}}}) \leq \frac{F}{F'} f^\rho C_{\text{max}}^\sigma$ which concludes the proof.

5- Case 2: Mathematical modeling for production network with transportation

In this section, we add another realistic assumption to the problem. As mentioned earlier, in the classic scheduling problems set of orders must be scheduled on machines in a single factory. However, the production processes within many industrial factories are distributed over the several manufacturing factories. In non-cooperative solution, all factories are responsible for the production of their local cluster. In such system a schedule should give enough flexibility to a local scheduler. This can be attained by transporting the jobs from the overloaded machine to the machine which has fewer workloads. Nevertheless, a certain price must be paid in order to produce solutions in which all factories have an incentive to participate. Therefore, considering the transportation time from one factory to another is important in the scheduling process.

In considering the transportation time in modeling the scheduling problem some assumptions must be regarded which are as follows:

- It is assumed that there are transportation routes among the factories and number of the transporters is an infinitive with the same quantity, capacity, time and cost of the drive.
- Due to the distribution of factories and the need for coordinated scheduling, communication between the entities of network is necessary.
- The main target activity of these transporters is transferring a job form the initial factory which is ordered on it to process on another factory with the objective of solution quality improvement.

The sum of the transportation times to carry the job $i$ form the factory $f$ to factory $q$ and returning it into the primary factory’s warehouse (to delivered to local customers) is denoted by $T_{ij}^{fq}$, which is
assumed to be independent of the jobs being transported. We also suppose that loading and unloading times are not considered separately and included in transportation times.

It is important to notice that for considering the transportation among network’s entities, an information exchange infrastructure between members is required to coordinate between global scheduling and local scheduling. Obviously, the partners’ flexibility can be improved if each factory is informed about the loading and availability of manufacturing resources at various partners. Albeit, this information should be exchanged as early as possible in order to help the network’s factories to plan more accurately and to adjust their capacities more effectively (Karatza 2001).

In Figure 6 with the simple numerical example, we show how jobs’ transportation can be considered in the main body of scheduling and reduce the global objective function.

In the above figure, we can see that after the jobs are locally scheduled, at the time $t=19$ the first machine in Factory 2 is idle while the second machine in Factory 1 is overloaded. Absolutely, in ideal distributed structure, the system is preferable that a balanced use of machines can be achieved and the overloading is avoided. After detailed information is exchanged between two factories, the workload is uniformly distributed over all of the machines by transporting $J_7$ from Factory 1 to Factory 2. As seen in Figure 6, the global objective function is improved in comparison with the first scheduling.

This simple example makes sense that transportation is an integrated feature in multi-factory scheduling systems and considering this assumption makes the scheduling problems more practical.

5-1- MILP model
Mathematical models are the natural starting point for detailed problem characterization (Granot et al. 1997). Before the model is presented, we introduce the parameters and indices employed. They are defined as follows.

- $i, j, k$  index of jobs; $i, j, k \in \{1, 2, ..., n\}$
- $f, q$  index of factories; $f, q \in \{1, 2, ..., F\}$
- $p_{iq}^{f}$ modified processing time of job $i$ in factory $q$ which originally is ordered to factory $f$
- $r_{iq}^{f}$ the transportation time to carry the jobs from factory $f$ to factory $q$ where $t_{iq}^{f} = T_{IQ}^{f}/2$
- $L$  large positive number

\[ x_{ij}^{f} = \begin{cases} 1 & \text{if job } i \text{ is scheduled immediately before job } j \text{ in factory } f, \\ 0 & \text{otherwise.} \end{cases} \]

\[ y_{i}^{f} = \begin{cases} 1 & \text{if job } i \text{ is assigned to factory } f, \\ 0 & \text{otherwise.} \end{cases} \]
if job \( i \) is originally ordered to factory \( f \),
\[
\begin{cases}
1 & \text{if job } i \text{ is originally ordered to factory } f, \\
0 & \text{otherwise}.
\end{cases}
\]

We introduced the dummy jobs 0 and \( n + 1 \). The processing time of these jobs is set at 0 and for all \( f \),
\( C'_0 = 0 \).

The problem can now be formulated as:
\[
Z = \min \left( \sum_{j \in R_1} C'_j, \ C'_\max \right) \tag{1}
\]

s.t. 
\[
\sum_{j=1}^{n} y'_j = 1, \quad i = 1, 2, \ldots, n, \tag{2}
\]
\[
\sum_{i=0}^{n} \sum_{j=1}^{F} x'_{ij} = 1, \quad j = 1, 2, \ldots, n, \tag{3}
\]
\[
\sum_{j=1}^{n} x'_{ij} = m'_f, \quad f = 1, 2, \ldots, F, \tag{4}
\]
\[
\sum_{j=1}^{n} x'_{ij} \leq 1, \quad j = 1, 2, \ldots, n, \tag{5}
\]
\[
\sum_{i=1}^{n} \sum_{j=1}^{F} x'_{ij} = 2 \cdot y'_j, \quad j = 1, 2, \ldots, n, \ f = 1, 2, \ldots, F, \tag{6}
\]
\[
\sum_{j=1}^{n} \sum_{i=1}^{F} x'_{ij} \leq 1, \quad i = 1, 2, \ldots, n, \tag{7}
\]
\[
\sum_{j=1}^{n} (x'_{ij} + x'_{ij}) \leq 1, \quad i = 1, 2, \ldots, n-1, \ j > i, \tag{8}
\]
\[
p'_{ij} = w'_j \left( \frac{P_j}{v'_j} + 2r'^{\text{inf}} \right), \quad i = 1, 2, \ldots, n, \ f, q = 1, 2, \ldots, F, \tag{9}
\]
\[
C'_j - C'_i \geq p'^{\text{inf}} - L(1 - x'_{ij}), \quad i, j = 0, 1, \ldots, n, \ i \neq j, \ f = 1, 2, \ldots, F, \tag{10}
\]
\[
C'_\max \geq C'_j, \quad j \in R_2, \ j = 1, 2, \ldots, n, \ f = 1, 2, \ldots, F, \tag{11}
\]
\[
C'_\max \geq C'_f, \quad f = 1, 2, \ldots, F, \tag{12}
\]
\[
x'_{ij}, y'_j \in \{0, 1\}, \quad i, j = 0, 1, \ldots, n+1, \ i \neq j, \ f = 1, 2, \ldots, F, \tag{13}
\]
\[
C'_i \geq 0, \quad i = 1, 2, \ldots, n, \ f = 1, 2, \ldots, F. \tag{14}
\]

In the proposed mathematical model, Relation (1) is the objective function. Relation (2) indicates that job \( i \) requires only one factory. Constraint (3) states that job \( i \) in the factory \( f \) requires only one machine. Relation (4) ensures that dummy job 0 has exactly one successor on each machine in the factory. Constraint (5) guarantees that job \( j \) can be assigned to at most the first position of each factory. Constraint (6) controls that every job can be either a successor or predecessor on each machine in the factory to which it is assigned. Constraint (7) shows that every job has at most one succeeding job. Constraint set (8) states that a job cannot be at the same time both a predecessor and a successor of another job. Relation (9) modifies processing time for the job \( i \) according to the distance between original factory of job which belongs to it and factory that job finally processed on it and speed of machines in this factory. Constraint (10) establishes the relationship between the completion times of jobs \( i \) and \( j \) that are assigned to the same machine in the specific factory. Constraint sets (11) and (12) define the makespan of job which belong to R2. Constraints (13) and (14) define zero-one variables and nonnegative variables, respectively. Note that the value of \( C'_i \) is zero when the job \( i \) is not assigned to factory \( f \).
5-2- Solving MILP via $\epsilon$-constraint

5-2-1- Improved approach

We point out that, to the best knowledge of the authors, there does not exist any methodology in the literature to obtain Pareto solutions for the problem modeled in previous sub-section. In order to obtain Pareto solutions we make use of $\epsilon$-constraint approach. This approach, as shows in the following, chooses one objective as a main objective function of the model and all other objective functions put in constraints with their upper bounds.

$$
\begin{align*}
\min & \quad f_r(x) \\
\text{s.t.} & \quad X \in S \\
& \quad f_j(x) \leq \bar{U} \quad \text{for all} \quad t \neq r
\end{align*}
$$

The following results can be found in Miettinen (1999).

**Theorem 4.** A solution of the $\epsilon$-constraint problem is weakly Pareto optimal for multi-objective optimization (MOP).

**Theorem 5.** A feasible point $x^*$ of the $\epsilon$-constraint problem is Pareto optimal if and only if it is a solution of MOP for every objective function, where $\epsilon_i = f_i(x^*)$ for $i = 1, 2, \ldots, k$ and $i \neq j$.

According to the Mavrotas (2009) with some modifications the results of this approach can improve. This modification in the problem with $p$ objective function is as follows.

$$
\begin{align*}
\min & \quad f_r(x) + \epsilon \sum_{i \neq r} \epsilon_i s_i \\
\text{s.t.} & \quad X \in S \\
& \quad f_j(x) + s_i = \bar{U} \quad \text{for all} \quad t \neq r
\end{align*}
$$

where $\epsilon$ is a small number between $10^{-3}$ and $10^{-6}$ (Mavrotas 2009).

**Theorem 6.** A solution of the improved $\epsilon$-constraint is only efficient solutions and the generation of weakly efficient solutions is avoided.

**Proof.** See Mavrotas (2009).

5-2-2- Enumerating the solution

Given the job sets $J_A$ and $J_B$ of the two factories, the two objective functions $Z_A(.)$ and $Z_B(.)$, and an integer $Q$, find a schedule $\sigma^*$ such that $Z_B(\sigma^*) \leq Q$, and $Z_A(\sigma^*)$ is minimum. This methodology is known as $\epsilon$-constraint approach in bicriteria optimization (T’Kindt and Billaut 2006).

Applying this new mathematical modeling, in a new formulation, the objective function calculates the maximum completion time of only the jobs in R2. The optimal solution to the formulation minimizes the $\sum C_j$ for jobs in R2 but simultaneously ensures that the jobs in R1 do not exceed the pre-fixed makespan value of $C_{max}(Q)$.

In mathematical modeling, the objective function can now be formulated as:

$$
\text{Min } Z = \sum_{j \in R_1} C_j', \quad (17)
$$

And constraint $C_{max} \leq Q$ is added to the constraint sets.
Now if $Q$ is too small, an instance of problem may not have feasible solutions. If there is at least one feasible solution, we say that the instance is *feasible*. Note that the problem of finding, among optimal schedules, one scheduling which is also nondominated can be always addressed by *binary search* (Agnetis et al. 2004). Let $\sigma^*$ be the optimal solution to $(Z_A: Z_B \leq \bar{Q})$. We can solve $(Z_A: Z_B \leq \bar{Q}/2)$, obtaining a solution $\sigma'$. If $Z_A(\sigma') > Z_A(\sigma^*)$, $\bar{Q}/2$ is indeed too small, so we try next $3\bar{Q}/4$, or else we decrease $Q$ again to $\bar{Q}/4$. This goes on until we individuate the smallest value $Q^*$ of $Q$ such that the value of the optimal solution to $(Z_A: Z_B \leq Q)$ is still equal to $Z_A(\sigma^*)$. Following Agnetis et al. (2004) Algorithm 4 shows the scheme of algorithm for enumerating the nondominated schedules via $\epsilon$ - constraint approach.

**Algorithm 4:** Enumerating the nondominated schedules via $\epsilon$ -constraint approach

**Start**

1. **Step 1:** Initialize: $S:=\emptyset; Q:=+\infty; i:=0$;
2. **Step 2:** generate the set of non-dominated points
   
   $\{$
   
   **While** the schedule is feasible do
   
   $i:=i+1$;
   
   $\sigma^i=$Solve the scheduling problem for single objective with $Z_B(\sigma^i) \leq Q$ by CPLEX;

   $S:= S \cup \sigma^i$;

   $Q' := Z_B(\sigma^i)$;

   $Q := Q - \epsilon$;

   **End while**

$\}$

**Step 3.** Output optimization results.

**End**

This is, however, a computationally intensive procedure and is feasible only for small-sized problem instances. We test the MILP models using CPLEX solver and consider two instances: 10-job with 2-factory problem and 25-job with 4-factory. The processing times for each instance are generated using a discrete uniform distribution from 30 to 60. Each instance is run several times with considering different $Q$ in constraint $C_{\text{max}} \leq Q$, according to Algorithm 4 and the results of two objectives are reported in Table 3, separately.
Table 3. Solutions obtained using $\epsilon$-constraint approach for two instances

<table>
<thead>
<tr>
<th>Solution’s number</th>
<th>Instance 1</th>
<th></th>
<th>Instance 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$C_{\text{max}}$</td>
<td>$TCT$</td>
<td>RT (s)</td>
<td>$C_{\text{max}}$</td>
</tr>
<tr>
<td>1</td>
<td>75</td>
<td>inf</td>
<td>50</td>
<td>170</td>
</tr>
<tr>
<td>2</td>
<td>80</td>
<td>427.273</td>
<td>300</td>
<td>180</td>
</tr>
<tr>
<td>3</td>
<td>85</td>
<td>427.273</td>
<td>300</td>
<td>190</td>
</tr>
<tr>
<td>4</td>
<td>90</td>
<td>427.273</td>
<td>300</td>
<td>200</td>
</tr>
<tr>
<td>5</td>
<td>91</td>
<td>427.273</td>
<td>300</td>
<td>250</td>
</tr>
<tr>
<td>6</td>
<td>93</td>
<td>407.273</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>7</td>
<td>94</td>
<td>360.000</td>
<td>300</td>
<td>350</td>
</tr>
<tr>
<td>8</td>
<td>95</td>
<td>358.182</td>
<td>300</td>
<td>400</td>
</tr>
<tr>
<td>9</td>
<td>100</td>
<td>357.273</td>
<td>300</td>
<td>450</td>
</tr>
<tr>
<td>10</td>
<td>105</td>
<td>357.273</td>
<td>300</td>
<td>500</td>
</tr>
<tr>
<td>11</td>
<td>110</td>
<td>354.000</td>
<td>300</td>
<td>550</td>
</tr>
<tr>
<td>12</td>
<td>115</td>
<td>352.727</td>
<td>300</td>
<td>600</td>
</tr>
<tr>
<td>13</td>
<td>116</td>
<td>347.273</td>
<td>300</td>
<td>650</td>
</tr>
<tr>
<td>14</td>
<td>119</td>
<td>339.818</td>
<td>300</td>
<td>700</td>
</tr>
<tr>
<td>15</td>
<td>120</td>
<td>335.636</td>
<td>300</td>
<td>750</td>
</tr>
<tr>
<td>16</td>
<td>121</td>
<td>331.091</td>
<td>300</td>
<td>800</td>
</tr>
<tr>
<td>17</td>
<td>125</td>
<td>331.091</td>
<td>300</td>
<td>850</td>
</tr>
<tr>
<td>18</td>
<td>126</td>
<td>331.091</td>
<td>300</td>
<td>900</td>
</tr>
<tr>
<td>19</td>
<td>127</td>
<td>331.091</td>
<td>300</td>
<td>950</td>
</tr>
<tr>
<td>20</td>
<td>128</td>
<td>331.091</td>
<td>300</td>
<td>1000</td>
</tr>
<tr>
<td>21</td>
<td>130</td>
<td>331.091</td>
<td>300</td>
<td>1050</td>
</tr>
<tr>
<td>22</td>
<td>135</td>
<td>331.091</td>
<td>300</td>
<td>1100</td>
</tr>
<tr>
<td>23</td>
<td>140</td>
<td>311.152</td>
<td>300</td>
<td>1150</td>
</tr>
<tr>
<td>24</td>
<td>145</td>
<td>310.424</td>
<td>300</td>
<td>1200</td>
</tr>
<tr>
<td>25</td>
<td>170</td>
<td>310.424</td>
<td>300</td>
<td>1500</td>
</tr>
</tbody>
</table>

*RT (s): running time (second)

Figures 7 and 8 illustrate the results.

Fig. 7. The non-dominated solutions obtained by $\epsilon$-constraint method for 10-job with 2-factory instance
The preliminary results indicate considerable potential to obtain good solutions by implementation of \( \epsilon \)-constraint approach. From the results shown in Figs. 7 and 8, it can be observed that the model can find near-optimal solutions efficiently. This method is, however, a computationally intensive procedure and is feasible only for the small-size instances.

6- Conclusion and recommendation for future study

Since a multi-factory scheduling is NP-hard, no optimal polynomial time algorithms can be designed, so this paper proposed the polynomial approximation algorithms for the problem in which factories’ production network can have an own criterion to optimize. Using a deeper analysis for the case that all factories are interested in the makespan, all factories are interested in the total completion time and the case that some factories are interested in the sum of completion times and remaining factories are interested in the makespan, we designed the best achievable approximation ratios by proposing the approximation algorithms for each case. In this paper, we also proposed a mixed integer linear programming for modeling scheduling problem with transportation times among network’s factories. In this problem, we consider the problem where the jobs belong to one of two disjoint sets; the makespan criterion needs to be minimized for one of the sets, while the total completion time needs to be minimized for the other. In the modeling of problem, we use the multi-objective method namely augmented \( \epsilon \)-constraint approach in which the total completion time must be minimized while another criterion i.e. the makespan as a constraint must be less than the predetermined value \( Q \). By systematic changing in \( Q \)’s values, we solve it in small-sized instances by solver CPLEX and generate the set of non-dominated solutions. Since the problem of scheduling known as virtual production network with self-interested factories has received surprisingly little attention in the literature, despite its ability to model several real-world situations, a worthy direction of future research is to develop heuristics and metaheuristics for large-sized instances.

References


