Three-Phase Method for the Capacitated Lot-Sizing Problem with Sequence Dependent Setups

François Larroche[1,2(✉)], Odile Bellenguez[1], and Guillaume Massonnet[1]

[1] IMT Atlantique, LS2N, La Chantrerie, 4 rue Alfred Kastler, 44307 Nantes, France
`francois.larroche@imt-atlantique.fr`
[2] VIF, 10 Rue de Bretagne, 44240 La Chapelle-sur-Erdre, France

Abstract. This paper focuses on an industrial lot-sizing and scheduling problem that arises in the food industry and includes lost sales, overtimes and sequence-dependent setups on parallel machines. We propose a preliminary version of a three-phase iterative approach to optimize separately the affectation, the sequencing and the production of items. Our first numerical results suggest that with some additional improvements, this approach could be use in real-life by planners to reduce their costs.

Keywords: Lot-sizing · Iterative method · Scheduling

1 Introduction

This paper presents a new heuristic to solve a lot-sizing and scheduling problem that arises in practical cases from the food industry. The problem we address is large and complex, which real-life applications often involve multiple production machines and specific constraints, which makes this problem complex to solve and motivates the planners to use optimization tools in order to reduce the costs incurred during their production process. However, modeling such complex systems leads to large mathematical formulations that are (in general) intractable with commercial solvers for Mixed Integer Programs (MIP). As a consequence, we aim to develop new approaches that enable us to obtain solutions to this problem in a quick and efficient fashion. Our work focuses on a multi-item capacitated lot-sizing problem including lost sales, safety stock, overtimes and sequence dependent setups on parallel machines. The discrete lot-sizing is a classical problem in the Operations Research literature since the work of Wagner and Whitin [13] and it has since been enriched with various additional considerations in order to model more accurately practical situations. Limited production capacity is among the most popular extensions of the original problem and has been the topic of multiple surveys, see e.g. [9] and [12]. The model we consider further generalizes this setting to include lost sales and shortage costs, as in [2] where the authors introduce new classes of valid inequalities. Parallel resources production is more and

more common in the literature, with a distinction between identical [3] and non-identical [8] machines. We focus on the latter in this paper.

The integration of lot-sizing and scheduling constraints greatly increases the complexity of the problems, which often leads researchers to propose heuristic solutions. Notable ones include [7], who develop a MIP-based neighborhood search heuristics to address a lot-sizing and scheduling problem on parallel machines. [4] present an industrial problem in which setup times depend on the sequence of production and propose a solution procedure based on sub-tour elimination and patching. [6] reviews the main modeling techniques for this class of problems and compare the efficiency of several solution methods. Iterative heuristics has been studied in the literature for the capacitated lot-sizing and scheduling problem in [11]. In [5], the authors used a two-level method to deal with the lot-sizing and the scheduling phases separately. In the related field of Production Routing problem, [1] use a two-phase method to dissociate the production and the distribution part.

In this paper, we propose a three-phase iterative method for the production problem described above, that has been introduced in [10] and is referred to as CLSSD-PM for *multi-item capacitated lot-sizing problem with lost sales, safety stock, overtimes, and sequence-dependent setups on parallel machines*. Section 2 describes the notations and assumptions. In Sect. 3, we present an original three-phase iterative method, with a quick description of each phase. In Sect. 4, we succinctly present the preliminary results obtained on a set of instances generated from industrial data. Finally, we propose some ideas to improve the procedure as well as perspectives for future research in the Sect. 5.

2   Notations and Assumptions

The problem studied in this paper is directly related to the problem originally presented in [10]. For an extensive presentation, the reader can refer to this paper. In this problem, we consider N different items that can be produced over a discrete finite horizon of T periods and M non-identical parallel machines. We denote $\mathcal{N}$, $\mathcal{M}$ and $\mathcal{T}$ the set of items, machines and periods, respectively. For all items $i \in \mathcal{N}$, we consider a deterministic demand d_t^i in each period $t \in \mathcal{T}$, which is either satisfied from on-hand inventory or lost. We denote τ_m^i the production time of one unit of item i on machine m. Sequence-dependent setup times are denoted λ_m^{ij} for a given machine m and for each pair (i, j) of items. Note that we allow asymmetric setup times but assume they satisfy the triangle inequality. For each machine $m \in \mathcal{M}$, we consider a time-dependent capacity C_{mt} corresponding to the total time available in period t. However, in some circumstances this capacity can be exceeded up to a larger value $\overline{C}_{mt}$ which is the "true" available time capacity in period t. In that case, $\overline{C}_{mt} - C_{mt}$ corresponds to the maximum overtime allowed. Finally, we also impose that every production of item i have to be greater than a minimum production quantity denoted q_{min}^i. The objective is to minimize the costs incurred in the production problem. We denote p_{mt}^i the cost of producing one unit of item i in period t on machine m. In addition, we

denote c_{mt} the per-unit of time usage cost of machine m during period t. For each unit of overtime, an additional cost $\bar{c}_{mt}$ is applied. We define the safety stock as an "optimal" inventory level to reach in each period. Inventory-related cost are then defined with respect to its value, namely a deficit cost h_t^{i-} and an overstock cost h_t^{i+} for each unit below or above the safety stock level. Finally, for each unmet unit of demand for item i in period t, a shortage cost l_t^i is incurred.

Throughout the remainder of this paper, we use the following decision variables:

$x_{mt}^i \in \mathbb{R}_+$ Quantity of item i produced in period t on machine m
$y_{mt}^i \in \{0,1\}$ binary variables equal to 1 if item i is affected to machine m in period t
$U_{mt} \in \mathbb{R}_+$ Time usage of machine m in period t
$O_{mt} \in \mathbb{R}_+$ Time in overtime of machine m in period t
$L_t^i \in \mathbb{R}_+$ Quantity of lost sales for item i in period t
$I_t^i \in \mathbb{R}_+$ Inventory of item i on hand at the end of period t
$I_t^{i+} \in \mathbb{R}_+$ Overstock (based on safety stock value) of item i at the end of period t
$I_t^{i-} \in \mathbb{R}_+$ Safety stock deficit of item i at the end of period t
$z_t^i \in \{0,1\}$ Binary variable equals to 1 if the stock of item i is null at the end of period t

3 Three-Phase Iterative Approach

In this section we present a three-phase iterative method to solve the problem under study. The central idea is to decompose the original problem into three (smaller) subproblems we can solve sequentially, where the output of a phase serves as an input for the following one. Below is an overview of one iteration of the heuristic:

– The first phase minimizes the costs induced by production and inventory management, considering virtual *sequence-independent* setup times per item. Capacity restrictions are also relaxed based on the information provided by the other phases.
– The second phase proposes production sequences for the items affected to each machine and period.
– The last phase decides the quantity to produce and pushes its output information into the first phase of the next iteration as an additional constraint.

The procedure loops through these different phases until a stopping criteria (either the time limit or a given number of iterations without improvement) is met.

First Phase: Assignment. We describe the first phase with a mathematical model based on an aggregate formulation. In this phase, we do not consider the sequence dependent setup times, which are approximated. For that purpose, we introduce ST_{mt}^i the setup time for item i for machine m at period t (set to 0

at the first iteration). The aggregate formulation is expressed with the following MIP:

$$\min \sum_{t\in\mathcal{T}} \sum_{m\in\mathcal{M}} (c_{mt}U_{mt} + \bar{c}_{mt}O_{mt})$$

$$+ \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}} \left(h_t^{i+}I_t^{i+} + h_t^{i-}I_t^{i-} + l_t^i L_t^i + \sum_{m\in\mathcal{M}} p_{mt}^i x_{mt}^i \right) \quad (1)$$

$$\text{s.t.} \quad U_{mt} = \sum_{i\in\mathcal{N}} (\tau_m^i x_{mt}^i + ST_{mt}^i y_{mt}^i) \quad \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \quad (2)$$

$$U_{mt} \le C_{mt} + O_{mt} \quad \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \quad (3)$$

$$I_t^i = S_t^i + I_t^{i+} - I_t^{i-} \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (4)$$

$$I_t^i = I_{t-1}^i + \sum_{m\in\mathcal{M}} x_{mt}^i + L_t^i - d_t^i \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (5)$$

$$I_t^i \le D_{tT}^i (1 - z_t^i) \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (6)$$

$$L_t^i \le z_t^i d_t^i \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (7)$$

$$x_{mt}^i \le D_{t-1,T}^i y_{mt}^i \quad \forall i \in \mathcal{N}, \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \quad (8)$$

$$x_{mt}^i \ge q_{\min}^i y_{mt}^i \quad \forall i \in \mathcal{N}, \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \quad (9)$$

$$L_t^i, I_t^i, I_t^{i+}, I_t^{i-} \ge 0 \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (10)$$

$$U_{mt}, O_{mt} \ge 0 \quad \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \quad (11)$$

$$y_{mt}^i \in \{0,1\} \quad \forall i \in \mathcal{N}, \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \quad (12)$$

$$z_t^i \in \{0,1\} \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \quad (13)$$

$$x_{mt}^i \ge 0 \quad \forall i \in \mathcal{N}, \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \quad (14)$$

The objective (1) is to minimize the cost of production including line usage costs, lost sales and inventory stocks. Constraints (2) define the total working time of each machine which is equal to the production time and the approximated setup times. Constraints (3) define the time usage and overtime variables. Constraints (4) define the inventory stock from the safety stock, the overstock and the deficit. Constraints (5) are the inventory flow conservation equations through the planning time horizon. Constraints (6) and (7) ensure that lost sales for item i occur only when the corresponding stock is null. Constraints (8) use the cumulative demand over the remainder of the horizon as an upper bound on the quantity of each item produced on each machine in a given period. Constraints (9) force the production to be higher than its minimum value. Finally constraints (10) and (14) define the domain of specific variables.

An additional constraint is necessary to link the information from the following phases after the first iteration. We denote $\widetilde{y}_{mt}^i$ and $\widetilde{x}_{mt}^i$ as the value of the corresponding variables y_{mt}^i and x_{mt}^i decided in the previous iteration. In addition, we also define for each machine m and period t the residual capacity Cap_{mt}^{res} as the remaining available machine time after the previous iteration of the algorithm. Before the first iteration, we initialize these values as follows:

$\tilde{y}_{mt}^i = 0$, $\tilde{x}_{mt}^i = 0$ and $Cap_{mt}^{res} = \overline{C}_{mt}$. The following constraints (15) then indicate that for each item, the production time is bounded by the last solution but can be increased or decreased depending on items that are added to/removed from the assignment derived in the previous iteration and the residual capacity. Parameters σ_{mt}^i then correspond to the proportion of the production time available that is allocated to item i.

$$\tau_m^i x_{mt}^i + (1 - \tilde{y}_{mt}^i) ST_{mt}^i y_{mt}^i \leq \tau_m^i \tilde{x}_{mt}^i + \Big(\sum_{\substack{j \neq i \\ \tilde{x}_{mt}^j \neq 0}} (1 - y_{mt}^j)(\tau_m^j \tilde{x}_{mt}^j + ST_{mt}^j)$$

$$+ Cap_{mt}^{res} - \sum_{\substack{j \neq i \\ \tilde{x}_{mt}^j = 0}} (\tau_m^j q_{\min}^j + ST_{mt}^j) y_{mt}^j \Big) \sigma_{mt}^i \quad (15)$$

$$\forall i \in \mathcal{N}, \forall m \in \mathcal{M}, \forall t \in \mathcal{T}$$

Second Phase: Sequencing. From the first phase, we obtain the assignment of items to each pair machine-period (m, t). The goal of this phase is then to solve $M \times T$ sequencing problems with a straightforward mathematical formulation adapted from the TSP problem. For conciseness we do not detail it in this paper.

Third Phase: Production. The last phase solves the production problem. We denote Λ_{mt} the sum of setup times on machine m in period t obtained from the previous phase. If $\Lambda_{mt} + \sum_{i \in \mathcal{N}} q_{\min}^i \tilde{y}_{mt}^i > \overline{C}_{mt}$, i.e. the capacity is not sufficient to carry out the setup times and the minimum production for the current affectation, we set $\Lambda_{mt} = 0$ and $\tilde{y}_{mt}^i = 0$ for all $i \in \mathcal{N}$. Otherwise, the values Λ_{mt} and $\tilde{y}_{mt}^i$ remain the same. Note that since the assignment decisions $\tilde{y}_{mt}^i$ are fixed in the first phase, they are considered as parameter in this step and thus the only binary variables considered are z_t^i to ensure a First-Come First-Serve discipline, which make the problem easy to solve.

$$\min \sum_{t \in \mathcal{T}} \sum_{m \in \mathcal{M}} (c_{mt} U_{mt} + \bar{c}_{mt} O_{mt})$$

$$+ \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{N}} \left(h_t^{i+} I_t^{i+} + h_t^{i-} I_t^{i-} + l_t^i L_t^i + \sum_{m \in \mathcal{M}} p_{mt}^i x_{mt}^i \right) \quad (16)$$

$$\text{s.t. } U_{mt} = \sum_{i \in \mathcal{N}} \tau_m^i x_{mt}^i + \Lambda_{mt} \qquad \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \quad (17)$$

$$U_{mt} \leq C_{mt} + O_{mt} \qquad \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \quad (18)$$

$$O_{mt} \leq \overline{C}_{mt} - C_{mt} \qquad \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \quad (19)$$

$$\text{Constraints (4)–(7)}$$

$$x_{mt}^i \leq D_{t-1,T}^i \tilde{y}_{mt}^i \qquad \forall i \in \mathcal{N}, \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \quad (20)$$

$$x_{mt}^i \geq q_{\min}^i \tilde{y}_{mt}^i \qquad \forall i \in \mathcal{N}, \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \quad (21)$$

$$\text{Constraints (10), (11), (13), (14)}$$

Update the Parameters of the First Phase. At the end of the third phase, there is a crucial update step on the input parameters of the first phase. The algorithm updates the individual setup times based on their marginal contribution to the production sequence computed in the second step. That is, for each item i in the sequence of production on machine m in period t, the setup time ST^i_{mt} corresponds to the additional time needed to insert this item is the sequence based on its direct predecessor and successor. On the other hand, we use the cheapest insertion time for all items that are not already included in the current sequence. This idea is described more in details in [1]. Current affectation decisions are stored in the variables $\tilde{y}^i_{mt}$, while *a priori* production quantities $\tilde{x}^i_{mt}$ are directly extracted from the production decisions x^i_{mt} in the third phase. The three-phase algorithm is presented in flowchart 1.

4   Experimental Results

In this section, we present the experimental results obtained with our three-phase method. Our instances are designed from actual data from industrial cases. We derived 48 different instances combining the following parameters: $N \in \{20, 30, 40\}$, $T \in \{15, 30\}$, $M \in \{1, 2\}$. Note that these preliminary tests are not representative of the industrial reality, where the number of different items can be significantly larger than this benchmark. We use IBM Ilog CPLEX v12.10.0 to solve each phase and set a time limit of 900 s as a stopping criterion. If UB corresponds to the best feasible solution found by the method tested, the gap is defined as follows:

$$\text{Gap} = 100 \cdot \frac{UB - LB}{LB}, \tag{22}$$

where LB is the best lower bound obtained after 4 h with CPLEX. Table 1 presents a Gap comparison between the 3-phase method and a straightforward MIP resolution in 900 s by CPLEX. We set $\sigma^i_{mt} = 1/N$ in constraints (15). Although the results suggest that our method needs refinement before we consider applying it on larger instances, it shows some promises as a decision support tools for practitioners. In particular for larger instances, the 3-phase method seems to be more robust to obtain adequate solution. We observe that the computational time is mainly decided by the first and the second phase while the third one is solved quickly. For instances with $N \geq 30$, the second phase take a significant amount of the total computational time (Fig. 1).

5   Perspectives

We study a problem encountered in the food industry. Since the problem is too complex to obtain good solution using a straightforward MIP resolution, we base our approach on an iterative resolution from a decomposition into three smaller subproblems. Our first preliminary results suggest that some enhancement must
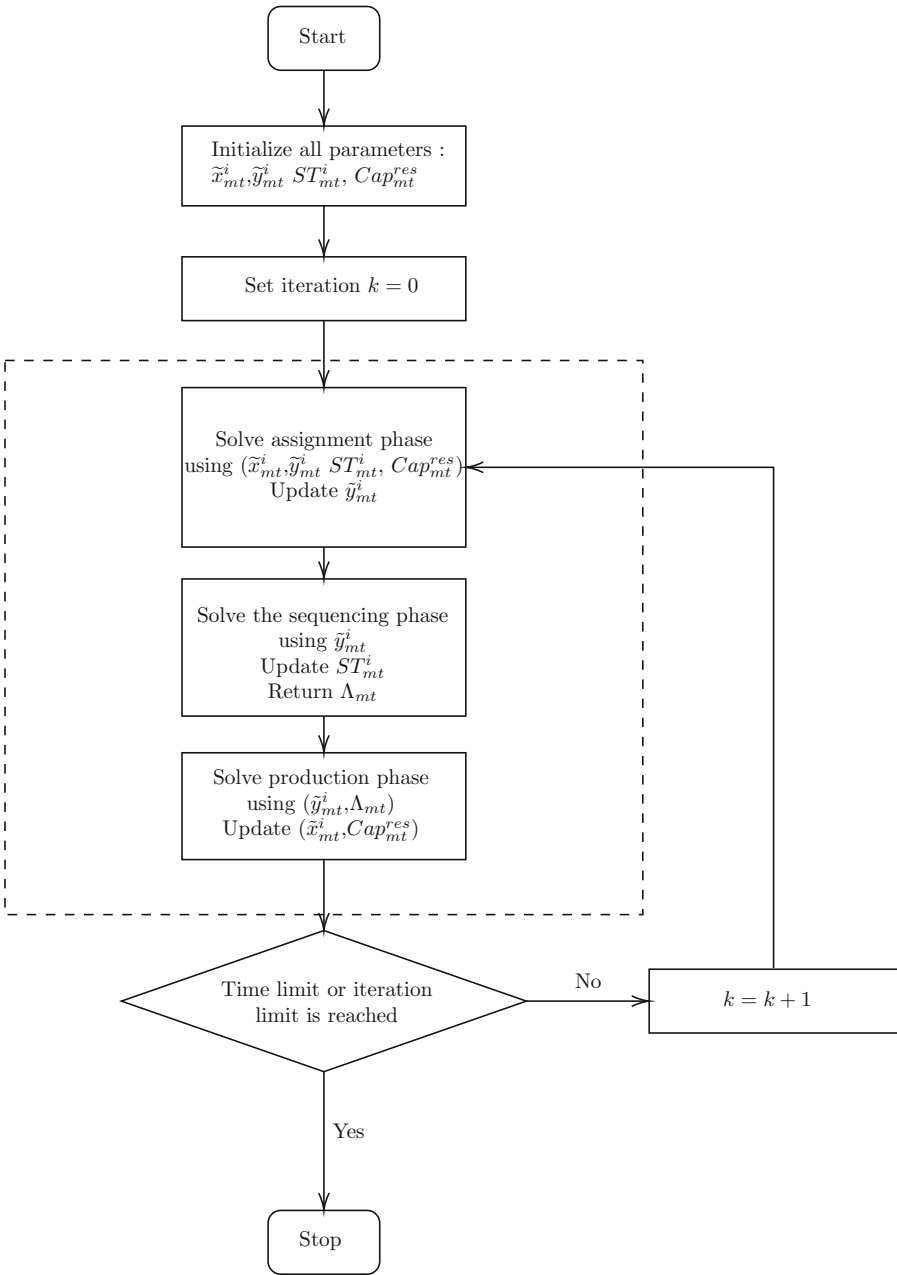
Fig. 1. Flowchart of the 3-phase approach

Table 1. Average gaps obtained with the 3-phase method

Problem size	Avg. gap (%)	
N-M-T	3-phase	CPLEX
20-1-15	77.57	0.57
20-1-30	47.51	3.15
20-2-15	37.4	1.28
20-2-30	73.9	2.01
30-1-15	47.72	3.1
30-1-30	97.23	17.68
30-2-15	60.64	6.95
30-2-30	118.42	27.29
40-1-15	289.74	585.16
40-1-30	372.61	2051.43
40-2-15	112.14	16.76
40-2-30	154.04	1053.87

be implemented to improve the method. To that end, we could consider the addition of diversification mechanisms in the first phase to speed up the convergence. Another direction under investigation rely on the clustering approach introduced in [10] to simplify the second phase and compute efficiently good production sequences. Another research direction is to use heuristics or MIP-heuristics in the first and second phases to speed up the resolution.

References

1. Absi, N., Archetti, C., Dauzère-Pérès, S., Feillet, D.: A two-phase iterative heuristic approach for the production routing problem. Transp. Sci. **49**(4), 784–795 (2015)
2. Absi, N., Kedad-Sidhoum, S.: The multi-item capacitated lot-sizing problem with setup times and shortage costs. Eur. J. Oper. Res. **185**(3), 1351–1374 (2008)
3. Beraldi, P., Ghiani, G., Grieco, A., Guerriero, E.: Rolling-horizon and fix-and-relax heuristics for the parallel machine lot-sizing and scheduling problem with sequence-dependent set-up costs. Comput. Oper. Res. **35**(11), 3644–3656 (2008)
4. Clark, A.R., Morabito, R., Toso, E.A.V.: Production setup-sequencing and lot-sizing at an animal nutrition plant through ATSP subtour elimination and patching. J. Sched. **13**(2), 111–121 (2010). https://doi.org/10.1007/s10951-009-0135-7
5. Dauzère-Péres, S., Lasserre, J.B.: Integration of lotsizing and scheduling decisions in a job-shop. Eur. J. Oper. Res. **75**(2), 413–426 (1994)
6. Guimarães, L., Klabjan, D., Almada-Lobo, B.: Modeling lotsizing and scheduling problems with sequence dependent setups. Eur. J. Oper. Res. **239**(3), 644–662 (2014)
7. James, R.J.W., Almada-Lobo, B.: Single and parallel machine capacitated lotsizing and scheduling: new iterative MIP-based neighborhood search heuristics. Comput. Oper. Res. **38**, 1816–1825 (2011)

8. Kaczmarczyk, W.: Modelling set-up times overlapping two periods in the proportional lot-sizing problem with identical parallel machines. DMMS **7**(1–2), 43 (2013)
9. Karimi, B., Fatemi Ghomi, S., Wilson, J.: The capacitated lot sizing problem: a review of models and algorithms. Omega **31**(5), 365–378 (2003)
10. Larroche, F., Bellenguez, O., Massonnet, G.: Clustering-based solution approach for a capacitated lot-sizing problem on parallel machines with sequence-dependent setups. Int. J. Prod. Res. (2021, Submitted)
11. Mateus, G.R., Ravetti, M.G., de Souza, M.C., Valeriano, T.M.: Capacitated lot sizing and sequence dependent setup scheduling: an iterative approach for integration. J. Sched. **13**(3), 245–259 (2010). https://doi.org/10.1007/s10951-009-0156-2
12. Quadt, D., Kuhn, H.: Capacitated lot-sizing with extensions: a review. 4OR **6**(1), 61–83 (2008). https://doi.org/10.1007/s10288-007-0057-1
13. Wagner, H.M., Whitin, T.M.: Dynamic version of the economic lot size model. Manag. Sci. **5**(1), 89–96 (1958)