JOB SHOP SCHEDULING WITH EARLINESS, TARDINESS AND INTERMEDIATE INVENTORY HOLDING COSTS

Kerem Bulbul
Manufacturing Systems and Industrial Engineering,
Sabanci University, Istanbul, Turkey
bulbul@sabanciuniv.edu

Abstract Motivated by the problem of scheduling customer orders in a make-to-order supply chain with relatively expensive intermediates, we consider a job shop scheduling problem with the objective of minimizing the sum of tardiness, earliness (finished goods inventory holding) and intermediate (work-in-process) inventory holding costs. Our goal is to develop scheduling heuristics to minimize the total cost and demonstrate the effectiveness of these heuristics computationally.

Keywords: machine scheduling; heuristic search; job shop scheduling; shifting bottleneck; earliness; tardiness; work-in-process inventory costs; Dantzig-Wolfe reformulation; column generation

1. Introduction

We address the problem of scheduling customer orders in job shops where our main objective is to study the trade-off between inventory holding (both work-in-process and finished goods) and tardiness costs. A study by the World Bank (Roberts, 1999) indicates that inventory holding costs account for 24% of the worldwide logistics costs that exceed $1 trillion in total. This huge financial impact that makes it essential to reduce inventories, coupled with the requirement for a high-level of on-time delivery performance introduces a trade-off between inventory holding and tardiness costs. Ideally, in order to reduce both work-in-process and finished goods inventory holding (earliness) costs, we would like to delay the production of customer orders as much as possible. Clearly, this approach increases the risk of tardiness, and we need a strategy that manages the trade-off between inventory holding and tardiness costs effectively.

High levels of inventory pose additional challenges in the rapidly changing high-tech industry in which steep price declines over a short period of time are common (Herrick and Verga, 2002), and obsolete inventories are a real danger (Barrett, 2001). In this context, it is clear that goods should not be manufactured far in advance of their delivery times, and that correct scheduling decisions are essential in order to avoid excessive inventory-related costs.

We are particularly interested in settings in which the total earliness cost is not a good indicator of the total inventory holding costs, i.e., when there is significant value added between processing stages. A specific example from the automotive industry is discussed by Bulbul et al., 2004a. In automobile manufacturing, steel is coated to delay corrosion and stored in a compact foil form until it is cut and stamped into different shapes when significantly more space for storage is required. Clearly, the inventory holding costs before and after coating, as well as before and after stamping are different, and these differences must be taken into account in order to make good scheduling decisions. For instance, a scheduling model that only considers the finished goods inventory holding (earliness) cost of a completed automobile body could prescribe a solution that prefers to keep the steel in uncoated condition due to some other constraint in the system. One way of avoiding this situation would be assigning a higher intermediate inventory holding cost to uncoated steel than that of coated steel. We refer the reader to Bulbul et al., 2004a and Bulbul, 2002 for further examples and a more in-depth discussion.

Motivated partly by the examples and issues described above, we consider a job shop scheduling problem with relatively expensive intermediate products. Our objective is to develop scheduling heuristics to minimize the sum of tardiness, earliness (finished goods inventory holding) and intermediate (work-in-process) inventory holding costs. In the single-machine environment, problems with earliness and tardiness costs have been studied in depth. (See Kanet and Sridharan, 2000 and references therein). Also, in the parallel machine environment many results have been obtained for common due date problems. (See Sundararaghavan and Ahmed, 1984, Kubiak et al., 1990, Federgruen and Mosheiov, 1996, Chen and Powell, 1999a and Chen and Powell, 1999b.) However, relatively little research exists on earliness/tardiness problems in flow- and job shop environments. For instance, Sarper, 1995 and Sung and Min, 2001 consider 2-machine flow shop common due date problems, and Chen and Luh, 1999 develop heuristics for a job shop weighted earliness and tardiness problem. Work-in-process inventory holding costs have been explicitly incorporated in Park and Kim, 2000, Kaskavelis and Caramanis, 1998, and Chang and Liao, 1994 for various flow- and job shop scheduling problems, although these papers consider different objective functions and approaches than ours. Further related references are given by Avci and Storer, 2004 who develop effective local search neighborhoods for a broad class of scheduling problems that includes the job shop weighted earliness and tardiness scheduling problem.

2. Problem Description

Consider a non-preemptive job shop with m machines and n jobs where J_i is the set of jobs to be processed on machine i. The operation sequence of job j is denoted by M_j where the i^{th} operation in M_j is represented by $M_{ij},\ i = 1, \ldots, m_j = \mid M_j \mid$. Associated with each job $j,\ j = 1, \ldots, n$, are several parameters: p_{ij}, the processing time for job j on machine i; r_j, the ready time for job j; d_j, the due date for job j; h_{ij}, the holding cost per unit time for job j while it is waiting in the queue before machine i; ϵ_j, the earliness cost per unit time if job j completes its final operation before time d_j; and π_j, the tardiness cost per unit time if job j completes its final operation after time d_j. All ready times, processing times and due dates are assumed to be integer. For clarity of notation, in the definitions above and the model below, assume that machine i for job j corresponds to the machine which performs operation M_{ij}. However, note that our proposed solution approaches are general and do not rely on such an assumption.

For a given schedule S, let C_{ij} be the time at which job j finishes processing on machine i, and w_{ij} be the time job j spends in the queue before machine i. The sum of costs for all jobs in schedule S, $C(S)$, can be expressed as follows:

$$C(S) = \sum_{j=1}^{n} h_{1j}(C_{1j} - p_{1j} - r_j) + \sum_{j=1}^{n}\sum_{i=2}^{m_j} h_{ij}w_{ij} + \sum_{j=1}^{n} \epsilon_j E_j + \pi_j T_j,$$

where $E_j = \max(0, d_j - C_{m_j,j})$ and $T_j = \max(0, C_{m_j,j} - d_j)$. Thus, the m-machine job shop scheduling problem Jm with earliness, tardiness and intermediate inventory holding costs can be formulated as:

$$\min_S \quad C(S) \tag{1}$$

s.t.

$$C_{1j} \geq r_j + p_{1j} \qquad \forall j \tag{2}$$

$$C_{i-1j} - C_{ij} + w_{ij} = -p_{ij} \qquad i = 2, \ldots, m_j,\ \forall j \tag{3}$$

$$C_{ik} - C_{ij} \geq p_{ik} \textbf{ or } C_{ij} - C_{ik} \geq p_{ij} \qquad \forall i, \forall j, k \in J_i \tag{4}$$

$$C_{m_j,j} + E_j - T_j = d_j \qquad \forall j \tag{5}$$

$$w_{ij} \geq 0 \qquad i = 2, \ldots, m_j, \forall j \tag{6}$$

$$E_j, T_j \geq 0 \qquad \forall j \tag{7}$$

The constraints (2) prevent processing of jobs before their respective ready times. Constraints (3), referred to as operation precedence constraints, ensure that any job j follows its processing sequence $M_{1j}, \ldots,$ $M_{m_j,j}$. Machine capacity constraints (4) ensure that a machine processes only one job at a time, and a job is finished once started. Constraints (5) relate the completion times of the final operations to the earliness and tardiness values.

Following the three field notation by Graham et al., 1979, $Jm/r_j/\sum_{j=1}^{n}\sum_{i=1}^{m_j} h_{ij}w_{ij} + \sum_{j=1}^{n}(\epsilon_j E_j + \pi_j T_j)$ represents problem Jm. Jm is strongly $\mathcal{NP}$-hard because a single-machine special case with all earliness costs equal to zero, i.e., the single-machine weighted tardiness problem $1/r_j/\sum \pi_j T_j$, is known to be strongly $\mathcal{NP}$-hard (Lenstra et al., 1977).

3. Solution Approach

We propose two heuristics in order to solve this problem. Both of these heuristics rely on machine-based decomposition approaches. First, we generalize the Dantzig-Wolfe decomposition and column generation technique that was developed and successfully applied to the $m-$machine flow shop version of this problem ($Fm/r_j/\sum_{j=1}^{n}\sum_{i=1}^{m} h_{ij}w_{ij} + \sum_{j=1}^{n}(\epsilon_j E_j + \pi_j T_j)$) by Bulbul et al., 2004a. Second, we propose a shifting bottleneck heuristic for this problem. The shifting bottleneck heuristic was originally developed for the problem $Jm//C_{max}$ by Adams et al., 1988 and then was applied to job shop scheduling problems in order to minimize the maximum lateness or the makespan under various constraints. In addition, Pinedo and Singer, 1999, Singer, 2001 and Mason et al., 2002 designed shifting bottleneck heuristics for various job shop weighted tardiness scheduling problems.

Observe that Jm would be a linear program (LP) if we knew the sequence of jobs on each machine. Using this property, we often develop two-phase heuristics for earliness/tardiness problems. First, a good job processing sequence is found for each machine, and then idle time is inserted either by a specialized algorithm or by solving a linear program. This latter step is commonly referred to as *timetabling*. For our problem, once jobs are sequenced, and assuming jobs are renumbered in sequence order on each machine, the optimal schedule is found by solving the linear program TTJm below.

$$\min_{C_{ij}} \quad C(S) \tag{8}$$

s.t.

$$(2)-(3) \text{ and } (5)-(7)$$

$$C_{ij} - C_{ij-1} \geq p_{ij} \qquad \forall i, j \in J_i, j \neq 1 \tag{9}$$

In our proposed solution approaches, we construct job processing sequences either by reformulating Jm using Dantzig-Wolfe decomposition and solving its linear programming relaxation approximately, or by solving the subproblems of a shifting bottleneck heuristic. In both cases, once the job processing sequences are obtained, we find the optimal schedule given these sequences by solving TTJm. The linear program TTJm has $O(nm)$ variables and constraints and can be solved effectively for typical job shop problem sizes, e.g., $m = 10$ and $n = 10, 15$, considered in the literature. Furthermore, partial job processing sequences may be accommodated easily in TTJm by specifying only a subset of the constraints (9) that correspond to partial sequences. Thus, TTJm could also be potentially embedded in a branch-and-bound algorithm.

4. Work-in-process and Future Research

We are currently working on the development of the heuristics mentioned in the previous section. The column generation algorithm presented by Bulbul et al., 2004a for the flow shop version of our problem exploits the duality between Dantzig-Wolfe decomposition and Lagrangian relaxation in order to solve the linear programming relaxation of the reformulated problem effectively. The single-machine subproblems in the column generation algorithm are either weighted completion time or earliness/tardiness problems with unequal ready times. These subproblems are solved using the approaches developed by Bulbul et al., 2004b. All of these algorithms can be generalized to the job shop scheduling problem under consideration without much difficulty.

The shifting bottleneck heuristic is an iterative machine-based decomposition algorithm with four basic steps: first, a disjunctive graph representation of the problem is constructed. Second, appropriate single-machine subproblems are formulated and solved for each unscheduled machine. Third, one of the unscheduled machines is selected as the "bottleneck" machine based on some machine criticality measure, and the schedule of this machine is fixed. Finally, the previous scheduling decisions are re-evaluated, and some machines are re-scheduled if necessary. For our problem, the most significant challenge is formulating the single-machine subproblems and designing good heuristics to solve them. Our initial results indicate that these subproblems are generalizations of the single-machine earliness/tardiness scheduling problem $1/r_j/\sum(\epsilon_j E_j + \pi_j T_j)$ which is strongly $\mathcal{NP}$-hard.

Ultimately, we plan to compare these two machine-based decomposition approaches described above in extensive computational tests and hope to obtain insights that could lead to simpler scheduling rules for the job shop scheduling problem with tardiness, earliness and intermediate inventory holding costs.

References

Adams, J., Balas, E., and Zawack, D. (1988). The shifting bottleneck procedure for job shop scheduling. *Management Science*, 34(3):391–401.

Avci, S. and Storer, R. (2004). Compact local search neighborhoods for generalized scheduling. Submitted for publication.

Barrett, L. (2001). CNET news.com. http://news.com.com/2100-1033-257278.html.

Bulbul, K. (2002). *Just-In-Time Scheduling with Inventory Holding Costs*. PhD thesis, University of California at Berkeley.

Bulbul, K., Kaminsky, P., and Yano, C. (2001, revised 2004b). Preemption in single machine earliness/tardiness scheduling. Submitted for publication to Journal of Scheduling.

Bulbul, K., Kaminsky, P., and Yano, C. (2004a). Flow shop scheduling with earliness, tardiness and intermediate inventory holding costs. *Naval Research Logistics*, 51(3):407–445.

Chang, Shi-Chung and Liao, Da-Yin (1994). Scheduling flexible flow shops with no setup effects. *IEEE Transactions on Robotics and Automation*, 10(2):112–122.

Chen, Haoxun and Luh, P.B. (1999). An alternative framework to Lagrangian relaxation approach for job shop scheduling. In *Proceedings of the 38th IEEE Conference on Decision and Control*, volume 1, pages 913–918.

Chen, Zhi-Long and Powell, W.B. (1999a). A column generation based decomposition algorithm for a parallel machine just-in-time scheduling problem. *European Journal of Operations Research*, 116(1):220–232.

Chen, Zhi-Long and Powell, W.B. (1999b). Solving parallel machine scheduling problems by column generation. *INFORMS Journal on Computing*, 11(1):78–94.

Federgruen, A. and Mosheiov, G. (1996). Heuristics for multimachine scheduling problems with earliness and tardiness costs. *Management Science*, 42(11):1544–1555.

Graham, R.L., Lawler, E.L., Lenstra, J.K., and Rinnooy Kan, A.H.G. (1979). Optimization and approximation in deterministic sequencing and scheduling: a survey. *Annals of Discrete Mathematics*, 5:287–326.

Herrick, S.L. and Verga, G. (2002). True inventory costs: A "back of the envelope" method to calculate true inventory savings. http://www.gvcsolutions.com/Trueinventorycosts.html.

Kanet, J.J. and Sridharan, V. (2000). Scheduling with inserted idle time: problem taxonomy and literature review. *Operations Research*, 48(1):99–110.

Kaskavelis, C.A. and Caramanis, M.C. (1998). Efficient Lagrangian relaxation algorithms for industry size job-shop scheduling problems. *IIE Transactions*, 30(11):1085–1097.

Kubiak, W., Lou, S., and Sethi, S. (1990). Equivalence of mean flow time problems and mean absolute deviation problems. *Operations Research Letters*, 9(6):371–374.

Lenstra, J.K., Rinnooy Kan, A.H.G., and Brucker, P. (1977). Complexity of machine scheduling problems. *Annals of Discrete Mathematics*, 1:343–362.

Mason, S.J., Fowler, J.W., and Carlyle, W.M. (2002). A modified shifting bottleneck heuristic for minimizing total weighted tardiness in complex job shops. *Journal of Scheduling*, 5(3):247–262.

Park, Moon-Won and Kim, Yeong-Dae (2000). A branch and bound algorithm for a production scheduling problem in an assembly system under due date constraints. *European Journal of Operations Research*, 123(3):504–518.

Pinedo, M. and Singer, M. (1999). A shifting bottleneck heuristic for minimizing the total weighted tardiness in a job shop. *Naval Research Logistics*, 46(1):1–17.

Roberts, P.O (1999). Supply chain management: New directions for developing economies. http://www.worldbank.org/html/fpd/transport/ports/trf_docs/new_dir.pdf.

Sarper, H. (1995). Minimizing the sum of absolute deviations about a common due date for the two-machine flow shop problem. *Applied Mathematical Modelling*, 19(3):163–161.

Singer, M. (2001). Decomposition methods for large job shops. *Computers and Operations Research*, 28(3):193–207.

Sundararaghavan, P.S. and Ahmed, M.U. (1984). Minimizing the sum of absolute lateness in single-machine and multimachine scheduling. *Naval Research Logistics Quarterly*, 31(2):325–333.

Sung, C.S. and Min, J.I. (2001). Scheduling in a two-machine flowshop with batch processing machine(s) for earliness/tardiness measure under a common due date. *European Journal of Operations Research*, 131(1):95–106.