ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

European Journal of Operational Research xxx (2004) xxx–xxx

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/dsw

Short Communication

On the single machine total tardiness problem

Z.J. Tian, C.T. Ng *, T.C.E. Cheng

Department of Logistics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Received 3 October 2002; accepted 10 March 2004

Abstract

In this paper we study the single machine total tardiness problem. We first identify some optimality properties based on which a special case with a given number of distinct due dates is proved polynomially solvable. The results are then extended to the case with release dates.
© 2004 Published by Elsevier B.V.

Keywords: Scheduling; Single machine; Total tardiness

1. Introduction

Among the classical scheduling problems, the single machine total tardiness problem without release dates (SMTTP), i.e., $1\|\sum T_i$, is one of the most widely studied. Since the first theoretical development by Emmons [3], an abundance of papers on SMTTP have been published. Two papers in the literature are of particular significance for the study of this well-known problem. Lawler [7] gave a pseudo-polynomial algorithm to solve SMTTP in $O(n^4P)$ time, where n is the number of jobs and P is the total processing time of all jobs. This implies that the problem cannot be strongly NP-hard. By a reduction from a restricted version of the NP-hard Even–Odd Partition problem, Du and Leung [2] resolved the complexity status of SMTTP, which had remained open for decades. Du and Leung's results, together with Lawler's pseudo-polynomial algorithm, confirm the ordinary NP-hardness of SMTTP. Before the publication of Du and Leung [2], the majority of the literature on this problem was focused on finding a polynomial algorithm or proving its NP-hardness. So, a great variety of enumerative algorithms have been proposed. Most of the algorithms rely heavily on the dominance rules developed by Emmons [3] and Lawler [7], both of which have been extended by other researchers. Following the resolution of the complexity status of this problem, a substantial body of literature was centered on heuristic algorithms. Koulamas [5] gave a comprehensive review of the total tardiness problems with an emphasis on critically evaluating heuristic algorithms for SMTTP. However, as indicated by Chen et al. [1], developing approximation algorithms with good performance guarantees for this problem is

* Corresponding author. Tel.: +852-2766-7364; fax: +852-2774-3679.

E-mail address: lgtctng@polyu.edu.hk (C.T. Ng).

0377-2217/$ - see front matter © 2004 Published by Elsevier B.V.
doi:10.1016/j.ejor.2004.03.017

difficult. The best ratio guarantee for any of the proposed heuristics is $n/2$. On the other hand, Szwarc and some other researchers [9,10] focused on the development of exact algorithms, which are mainly based on either dynamic programming or the branch-and-bound method, or their combination. Koulamas [5] also gave a brief survey of enumerative algorithms for SMTTP. A more detailed review on this aspect was presented by Chen et al. [1]. The most recent result can be found in Szwarc et al. [10], which contains a branch-and-bound algorithm that handles instance with up to 500 jobs.

Since SMTTP is NP-hard in the ordinary sense, it is worth investigating its polynomially solvable cases. However, there exist only a few such results. For example, the common due date problem $1|d_i = d|\sum T_i$ can be solved by the shortest processing time (SPT) rule according to Lawler and Moore [8], while the equal processing time problem $1|p_i = p|\sum T_i$ is obviously solvable by the earliest due date (EDD) rule, since it is identical to the unit processing time problem $1|p_i = 1|\sum T_i$ in the case of identical release dates. In the case that the due dates and processing times are agreeable, the problem $1|(d_i, p_i)|\sum T_i$ is solvable in $O(n \log n)$ time by the SPT or EDD rule according to Theorem 3 of Lawler [7], where (d_i, p_i) denotes the agreeability of due dates and processing times. Emmons [3] identified three special cases where SMTTP can be polynomially solved. Lawler [7] extended two of Emmons' results. Koulamas [6] reviewed the polynomially solvable cases of SMTTP and developed the other results of Emmons.

In this paper, we first recall a famous theorem in Lawler [7] and identify some optimality properties. Then, a special case of SMTTP with a given number of distinct due dates is proved polynomially solvable. The problem addressed in this paper can be formally stated as follows: A set of n jobs $N = \{1, \ldots, n\}$ has to be processed on a single machine that can perform only one job at a time. Each job i has a processing time p_i and a due date d_i. All p_i and d_i are integers. The objective is to schedule the jobs so as to minimize total tardiness $\sum T_i$, where $T_i = \max\{0, C_i - d_i\}$ and C_i is the completion time of job i. Using the three-field notation of Graham et al. [4] and Chen et al. [1], the problem can be denoted as $1\|\sum T_i$.

The remainder of this paper is organized as follows: In Section 2, some optimality properties and complexity analysis for a special case are presented. The results are extended to the problem with release dates in Section 3.

2. Optimality properties and complexity analysis

Lawler's [7] pseudo-polynomial algorithm is based on Theorem 3 of [7], which is in turn based on two preliminary results (Theorems 1 and 2 of [7]). The sensitivity of an optimal schedule to the due dates is considered in Theorem 1 of [7], while a dominance rule is presented in Theorem 2 of [7], which is also a result from Theorem 1 of Emmons [3].

Theorem 2.1 (Theorem 3 of Lawler [7]). *Suppose the jobs are agreeably weighted and numbered in nondecreasing due date order, i.e., $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_n$. Let job k be such that $p_k = \max_j \{p_j\}$. There is some integer δ, $0 \leqslant \delta \leqslant n - k$, such that there exists an optimal schedule π in which k is preceded by all jobs j such that $j \leqslant k + \delta$, and followed by all jobs j such that $j > k + \delta$.*

We note that the above theorem holds for a more general agreeably weighted case of $1\|\sum w_i T_i$, i.e. $p_i < p_j$ implies $w_i \geqslant w_j$. Similarly, all the following discussion holds for the agreeably weighted case, too.

Consider a special case of $1\|\sum T_i$ in which there are m distinct due dates $d'_1 < \cdots < d'_m$, where m is a fixed positive integer. To reflect the restriction on the due dates, we add $d_i \in \{d'_1, \ldots, d'_m\}$ to the second field of the three-field notation $\alpha|\beta|\gamma$ of Chen et al. [1] and Graham et al. [4]. Define N_i to be the subset of the $n_i = |N_i|$ jobs with the common due date $d'_i, i = 1, \ldots, m$. Also, suppose the n_i jobs in N_i are indexed by the SPT rule as $J_{i1}, \ldots, J_{in_i}$. It is easily proven, by a simple interchange technology, that there exists an optimal schedule in which $J_{i1} \rightarrow \ldots \rightarrow J_{in_i}$ holds for all $i = 1, \ldots, m$. Then, following Theorem 2.1, we can state Theorem 2.2 as follows:

Theorem 2.2. *Suppose there are m distinct due dates $d'_1 < \cdots < d'_m$, where $m < n$ is a given positive integer, and the jobs in the set N_i of the jobs with due date d'_i are indexed by the SPT rule as $J_{i1}, \ldots, J_{in_i}$. Let J_{kn_k} be such a job in N_k that $p_{kn_k} = \max_i \{p_{in_i}\}$. There is some integer δ, $0 \leqslant \delta \leqslant m - k$, such that there exists an optimal schedule π in which J_{kn_k} is preceded by $N_k \setminus \{J_{kn_k}\}$ and all N_i such that $i = 1, \ldots, k-1, k+1, \ldots, k+\delta$, and followed by all N_i such that $i > k + \delta$.*

Now, we consider the time complexity of Lawler's dynamic programming algorithm for this special case. Equation (3.1) in Lawler [7] is recalled below. We note that the notation in this equation is independent of our notation in Theorem 2.2:

$$T(S(i,j,k),t) = \min_{\delta}\{T(S(i,k+\delta,k'),t) + w_{k'}\max(0, C_{k'}(\delta) - d_{k'}) + T(S(k'+\delta+1,j,k'), C_{k'}(\delta))\}, \quad (1)$$

where k' is such that

$$p_{k'} = \max\{p_{j'} | j' \in S(i,j,k)\}.$$

First of all, according to Theorem 2.2, each Eq. (1) requires minimization over at most m (not n) alternatives and an $O(m)$ running time.

Next, consider any Eq. (1). For the sake of convenience, we say the set $S(i,j,k)$ is decomposed by job k' and call job k' the decomposition job. Suppose a set $N'_l \subseteq N_l$, is included in $S(i,j,k)$. By Theorem 2.2, all jobs in N'_l or a reduced set $N'_l \setminus \{k'\}$, where k' is the job with the largest index in N'_l and is selected as the decomposition job, should be wholly included in either $S(i,k+\delta,k')$ or $S(k'+\delta+1,j,k')$. As the decomposition job in each recursion is the one with the largest index among the jobs in the selected subset, the set N'_l is always in the form of $\{J_{l1}, J_{l2}, \ldots\}$. In other words, job i can only be one of the m jobs $J_{i1}, i = 1, \ldots, m$. Moreover, it is easily seen that, in the case of $i = m$, all jobs in $S(i,j,k)$ are from N_m, and decomposition of $S(i,j,k)$ is unnecessary. Hence, there are no more than $m - 1$ values for the index i in the set $S(i,j,k)$.

Then, any $J_{i1}, i \leqslant m - 1$, is preceded only by the jobs in N_l, $l < i$. So, the possible values of t are not more than $(n_1 + 1) \times \cdots \times (n_{m-2} + 1) \leqslant [(n + m - 2)/(m-2)]^{m-2}$ in the case of $m > 2$. In the case of $m = 2$, there is only one possible value of t, i.e., $t = 0$.

In conclusion, the problem $1|d_i \in \{d'_1, \ldots, d'_m\}|\sum T_i$ can be solved by Lawler's algorithm in $O(m^2n^2[(n+m-2)/(m-2)]^{m-2}) = O(m^{4-m}n^m)$ time in the case of $2 < m < n$.

Finally, we consider the case of $m = 2$. According to Theorem 2.2, in each recursion, if $k' \in N_1$ job k' is either immediately followed by J_{21} or immediately preceded by the job with the largest index among the remaining jobs in N_2, which should also be followed by job k' in the case of $k' \in N_2$. We note that in the case of $k' \in N_2$ and the latter case of $k' \in N_1$ the jobs following job k' are already sequenced in the SPT order. So, there are only $n_1 + 1$ schedules that need to be considered, i.e., for $j = 1, \ldots, n_1$, the first j jobs in N_1 (in SPT order), followed by the rest of the jobs in N_1 and all jobs in N_2 in SPT order. As the time complexity of the SPT ordering is $O(n \log n)$, the problem $1|d_i \in \{d'_1, d'_2\}|\sum T_i$ can be solved in $O(n \log n)$ time.

Recall that in the case of $m = 1$, the problem addressed is known as $1|d_i = d|\sum T_i$ and can be solved in $O(n \log n)$ by the SPT rule [8]. We have the following theorem:

Theorem 2.3. $1|d_i \in \{d'_1, \ldots, d'_m\}|\sum T_i$ *is polynomially solvable for a fixed value of m. Especially, the time complexity is* $O(n \log n)$ *for the case of* $m = 1, 2$ *and* $O(m^{4-m}n^m)$ *for the case of* $2 < m < n$.

3. An extension of the results to a problem with release dates

Consider a special case of the preemptive single machine total tardiness problem where the release dates and due dates are strictly agreeable, in the sense that $r_i < r_j$ implies $d_i \leqslant d_j$, and $r_i = r_j$ implies $d_i = d_j$. In such a case, we denote the problem as $1|(r_i, d_i)^=, pmtn|\sum T_i$ (P1). By assumption, all jobs simultaneously released are given equal slacks, but the slacks for two jobs with different release dates are not necessarily equal. It is easy to see that the problem $1|r_i, d_i = r_i + d, pmtn|\sum T_i$ is a special case of P1. The ordinary *NP*-hardness of this CON due date problem [11] implies that P1 cannot be polynomially solved. On the other hand, as a special case of $1|(r_i, d_i), pmtn|\sum T_i$, where (r_i, d_i) denotes the agreeability of the release dates and due dates, P1 cannot be strongly *NP*-hard [11]. Hence, we have the following result.

Theorem 3.1. $1|(r_i, d_i)^=, pmtn|\sum T_i$ *is* NP-*hard in the ordinary sense.*

Moreover, it is easy to see that the properties identified in the last section hold even for the case that all jobs in a subset $N_i, i = 1, \ldots, m$, are released simultaneously at r_i, but $r_i < r_{r+1}$ holds for any subset pair N_i and $N_{i+1}, 1 \leqslant i < m$. So, Theorem 2.3 also holds for $1|(r_i, d_i)^=, pmtn|\sum T_i$.

Acknowledgements

This research was supported by The Hong Kong Polytechnic University for Area of Strategic Development under grant number A628. We thank three anonymous referees for their constructive comments on an earlier version of this paper.

References

[1] B. Chen, C. Potts, G. Woeginger, A review of machine scheduling: Complexity, algorithms and approximability, in: Handbook of Combinatorial Optimization, Kluwer Academic Publishers, Boston, 1998, pp. 21–169.

[2] J.Z. Du, J.Y.-T. Leung, Minimizing total tardiness on one machine is *NP*-hard, Mathematics of Operations Research 15 (1990) 483–495.

[3] H. Emmons, One-machine sequencing to minimize certain functions of job tardiness, Operations Research 17 (1969) 701–715.

[4] R.L. Graham, E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, Optimization and approximation in deterministic sequencing and scheduling: A survey, Annals of Operations Research 5 (1979) 287–326.

[5] C. Koulamas, The total tardiness problem: Review and extensions, Operations Research 42 (1994) 1025–1041.

[6] C. Koulamas, Polynomially solvable total tardiness problems: Review and extensions, Omega 25 (1997) 235–239.

[7] E.L. Lawler, A 'pseudopolynomial' algorithm for sequencing jobs to minimize total tardiness, Annals of Discrete Mathematics 1 (1977) 31–342.

[8] E.L. Lawler, J.M. Moore, A functional equation and its application to resource allocation and sequencing problems, Management Science 16 (1969) 77–84.

[9] W. Szwarc, F. Della Croce, A. Grosso, Solution of the single machine total tardiness problem, Journal of Scheduling 2 (1999) 55–71.

[10] W. Szwarc, A. Grosso, F. Della Croce, Algorithmic paradoxes of the single-machine total tardiness problem, Journal of Scheduling 4 (2001) 93–104.

[11] Z.J. Tian, C.T. Ng, T.C.E. Cheng, Preemptive scheduling with agreeable release dates and due dates to minimize total tardiness, Proceedings of the 5th International Conference on Optimization Techniques and Applications, Hong Kong, 2001, pp. 405–412.