Available online at www.sciencedirect.com

SCIENCE DIRECT®

European Journal of Operational Research 159 (2004) 558–573

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/dsw

Discrete Optimization

Mean flow time minimization with given bounds of processing times

Tsung-Chyan Lai [a], Yuri N. Sotskov [b], Nadezhda Sotskova [c], Frank Werner [c,*]

[a] Department of Industrial and Business Management, National Taiwan University, 50, Lane 144, Sec. 4, Keelung Rd., Taipei 106, Taiwan

[b] United Institute of Informatics Problems, Surganov St. 6, 220012 Minsk, Belarus

[c] Fakultät für Mathematik, Institut für Mathematische Optimierung, Otto-von-Guericke-Universität, PSF 4120, 39016 Magdeburg, Germany

Received 22 November 1999; accepted 24 June 2003
Available online 23 September 2003

Abstract

We consider a job shop scheduling problem under uncertain processing times and fixed precedence and capacity constraints. Each of the random processing times can take any real value between given lower and upper bounds. The goal is to find a set of schedules which contains at least one optimal schedule (with mean flow time criterion) for any admissible realization of the random processing times. In order to compute such a set of schedules efficiently and keep it as small as possible, we develop several exact and heuristic algorithms and report computational experience based on randomly generated instances.

Keywords: Scheduling; Mean flow time; Uncertainty

1. Introduction

The job shop problem under consideration can be described as follows. Each job $J_i \in J = \{J_1, J_2, \dots, J_n\}$ has to be processed on each machine $M_j \in M = \{M_1, M_2, \dots, M_m\}$ exactly once. Technological routes (machine orders) are fixed for all jobs, but may vary from job to job. A job $J_i \in J$ consists of a chain of m uninterrupted operations $(O_{ii_1}, O_{ii_2}, \dots, O_{ii_m})$, where O_{ii_k} denotes the operation of job J_i on machine $M_{i_k} \in M$ and $(i_1, i_2, \dots, i_m)$ is a permutation of the indices $\{1, 2, \dots, m\}$. At the *scheduling stage*, the processing time p_{ij} of operation O_{ij} ($J_i \in J$, $M_j \in M$) is not fixed: p_{ij} may take any real value between a given lower bound $a_{ij} \geqslant 0$ and a given upper bound $b_{ij} \geqslant a_{ij}$. The probability distribution of the random processing times is unknown. The criterion is mean flow time minimization. By adopting the three-field notation [10], we denote this problem by $J|a_{ij} \leqslant p_{ij} \leqslant b_{ij}|\Phi$, where Φ is the objective function depending on the job completion times

* Corresponding author. Tel.: +49-391-6712025; fax: +49-391-6711171.
E-mail address: frank.werner@mathematik.uni-magdeburg.de (F. Werner).

0377-2217/$ - see front matter
doi:10.1016/S0377-2217(03)00424-7

C_i, $J_i \in J$, with $\Phi(C_1, C_2, \ldots, C_n) = \sum_{i=1}^{n} C_i = \sum C_i$.

Problem $J|a_{ij} \leqslant p_{ij} \leqslant b_{ij}|\Phi$ aims to complement but not to replace other models of uncertain scheduling environments, e.g. a stochastic model [13] and a fuzzy model [14]. A former model is useful when we have enough prior information to characterize the probability distributions of random processing times and there is a large number of realizations of similar processes, but it may have a limited significance for a small number of realizations. In [6], uncertainty of the processing times was modeled by means of fuzzy numbers for mean flow time minimization and for other criteria. The known results and current trends in the field of *scheduling under fuzziness* have been presented in [14]. To model scheduling in an uncertain environment, a two-person non-zero sum game was introduced in [5], where the decision maker was considered as player 1 and 'nature' as player 2.

We have to emphasize that the random processing times in the problem $J|a_{ij} \leqslant p_{ij} \leqslant b_{ij}|\Phi$ are due to external forces in contrast to scheduling problems with *controllable* processing times [7,17], when the objective is to determine optimally the processing times (which are under control of a decision maker) and the schedule at the same time. Another related yet different problem is the *hoist scheduling problem*, see [4,12] for the definition along with results and Section 3 in [11] for a condensed survey. Hoist scheduling problems arise in chemical, electroplating and medical industries, where the objective is to minimize the cycle time of a repetitive process. The assigned hoist has to travel to the tank, wait if necessary, lift the job up at the suitable time, and travel to the next tank on the technological route. Due to the nature of the chemical process, the processing time at each tank has to be strictly controlled (by a decision maker) within given lower and upper bounds.

For problem $J|a_{ij} \leqslant p_{ij} \leqslant b_{ij}|\sum C_i$, a set of m sequences $(O_{i_1k}, O_{i_2k}, \ldots, O_{i_nk})$ of the operations processed on machines M_k, $k = 1, 2, \ldots, m$, defines a *schedule*, if none of these sequences contradicts to others and to the technological routes. For fixed processing times, these m sequences define the *earliest starting* and *earliest completion times* of all operations, i.e. these sequences define a *semiactive* schedule [10]. In what follows, we use a circuit-free digraph G_s to represent a semiactive schedule for problem $J|a_{ij} \leqslant p_{ij} \leqslant b_{ij}|\sum C_i$.

The approach under consideration was originated in [8,9] for the makespan objective function: $\Phi(C_1, C_2, \ldots, C_n) = \max\{C_i \,|\, J_i \in J\} = C_{\max}$. As far as scheduling is concerned, the mean flow time seems to be more important than makespan. While $C_{\max}$ aims for minimizing the schedule duration, in industry this duration is often defined by the periodicity of the process, say, a working day or a working week. So, the criterion $\sum C_i$ is more suitable for a periodical scheduling environment. Theoretically speaking, the problem with $\sum C_i$ is more difficult to solve than that with $C_{\max}$ as evidenced by the facts that the two-machine open shop problem $O2\|C_{\max}$ and the flow shop problem $F2\|C_{\max}$ are polynomially solvable while problems $O2\|\sum C_i$ and $F2\|\sum C_i$ are NP-hard (see e.g. [1,10]). (Obviously, the NP-hardness of the latter problem implies the NP-hardness of problem $J|a_{ij} \leqslant p_{ij} \leqslant b_{ij}|\sum C_i$ under consideration). In academic research, criterion $C_{\max}$ has mainly been considered for multi-stage systems (open, flow and job shops) while criterion $\sum C_i$ for single-stage systems, i.e. systems with one machine or with a set of *parallel* (usually identical) machines. One explanation could be that makespan is easier to attack than mean flow time.

The main technical issues of the research presented in [8,9] were to simplify the circuit-free digraph G_s due to the dominance relation between its paths. In this paper, we perform a further step in this direction by focusing on a dominance relation between circuit-free digraphs. This step is useful for scheduling problems for both criteria $C_{\max}$ and $\sum C_i$ since it allows to reduce the number of candidate schedules for a solution. Moreover, for the criterion $\sum C_i$ this step may be more important since the comparison of digraphs G_s with respect to $\sum C_i$ is essentially more complicated than that for $C_{\max}$ (see [3,15]). Note also that while the simplification of a digraph G_s may be done in polynomial time [8], to find a solution to problem $J|a_{ij} \leqslant p_{ij} \leqslant b_{ij}|\sum C_i$ or $J|a_{ij} \leqslant p_{ij} \leqslant b_{ij}|C_{\max}$ is NP-hard.

As follows from [2], a reduction of semiactive schedules may be significant for the case of all non-negative perturbations of the processing times: