Computers & Industrial Engineering 90 (2015) 186–196

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

A hybrid estimation of distribution algorithm for simulation-based scheduling in a stochastic permutation flowshop

K. Wang [a,*], S.H. Choi [b], H. Lu [c]

[a] Department of Management Science and Engineering, Economics and Management School, Wuhan University, Wuhan, China
[b] Department of Industrial and Manufacturing Systems Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong
[c] Department of Logistics Management, School of Logistics Engineering, Wuhan University of Technology, Wuhan, China

ARTICLE INFO

Article history:
Received 13 January 2015
Received in revised form 8 September 2015
Accepted 9 September 2015
Available online 16 September 2015

Keywords:
Permutation flowshop scheduling
Stochastic processing times
Estimation of distribution algorithm
Genetic algorithm
Meta-model

ABSTRACT

The permutation flowshop scheduling problem (PFSP) is NP-complete and tends to be more complicated when considering stochastic uncertainties in the real-world manufacturing environments. In this paper, a two-stage simulation-based hybrid estimation of distribution algorithm (TSSB-HEDA) is presented to schedule the permutation flowshop under stochastic processing times. To deal with processing time uncertainty, TSSB-HEDA evaluates candidate solutions using a novel two-stage simulation model (TSSM). This model first adopts the regression-based meta-modelling technique to determine a number of promising candidate solutions with less computation cost, and then uses a more accurate but time-consuming simulator to evaluate the performance of these selected ones. In addition, to avoid getting trapped into premature convergence, TSSB-HEDA employs both the probabilistic model of EDA and genetic operators of genetic algorithm (GA) to generate the offspring individuals. Enlightened by the weight training process of neural networks, a self-adaptive learning mechanism (SALM) is employed to dynamically adjust the ratio of offspring individuals generated by the probabilistic model. Computational experiments on Taillard's benchmarks show that TSSB-HEDA is competitive in terms of both solution quality and computational performance.

1. Introduction

The permutation flowshop scheduling problem (PFSP) is a well-known and well-studied combinatorial optimisation problem (Gupta & Stafford, 2006; Vallada & Ruiz, 2009). In the classical PFSP, jobs arrive at the shop floor simultaneously and then follow the same processing order on each of the machines. The PFSP has been proven strongly NP-complete for more than two machines (Garey, Johnson, & Sethi, 1976). Due to its great significance in both academic and real-world applications, the PFSP has attracted considerable attention after the pioneering work of Johnson (1954).

Although a tremendous amount of effort has been devoted to addressing the PFSP, most of the research works consider a static environment, in which no unexpected events would occur to disturb job processing. Real-world manufacturing environments, however, tend to suffer a variety of uncertainties, including change of processing time, machine breakdown, rush orders, and job cancellations, etc. (Gholami, Zandieh, & Alem-Tabriz, 2009; Ouelhadj & Petrovic, 2009). Therefore, permutation flowshop scheduling under uncertainties has recently received an increasing attention.

Three types of approaches, namely exact algorithms, heuristics, and meta-heuristics, are commonly adopted to solve the PFSPs in the literature (Ruiz & Maroto, 2005; Xu, Yin, Cheng, Wu, & Gu, 2014). Exact algorithms aim to achieve the optimal solution, and hence are computationally expensive for large-sized PFSPs. Examples of such methods are branch and bound approaches (Chung, Flynn, & Kirca, 2002). In addition to exact algorithms, heuristics and meta-heuristics have also been introduced to find approximate solutions within reasonable computational cost. Since most existing heuristic methods, such as constructive heuristics and improvement heuristics, tend to perform poorly on large-sized PFSPs (Ceberio, Irurozki, Mendiburu, & Lozano, 2014), a wide range of meta-heuristics have been applied to address the PFSPs (Zobolas, Tarantilis, & Ioannou, 2009).

To deal with uncertainties in a flowshop, the simulation-based meta-heuristics (SBM) have been successfully developed to construct and evaluate candidate solutions. In these approaches, a discrete-event simulator is usually incorporated into a meta-heuristic (Wang, Choi, Qin, & Huang, 2013), such as genetic algorithm (GA) (Dugardin, Yalaoui, & Amodeo, 2010), immune

* Corresponding author. Tel.: +86 27 68753063.
E-mail addresses: kai.wang@whu.edu.cn (K. Wang), shchoi@hku.hk (S.H. Choi), lvhaili@whut.edu.cn (H. Lu).

http://dx.doi.org/10.1016/j.cie.2015.09.007
0360-8352/

algorithm (Zandieh & Gholami, 2009), ant colony optimisation (ACO) algorithm (Ahmadizar, Ghazanfari, & Ghomi, 2010), and hybrid meta-heuristics (Safari & Sadjadi, 2011). As an iterative procedure, the meta-heuristic guides its subordinate heuristics to iteratively produce high-quality candidate solutions until a termination criterion is met. In the SBM, the performance of candidate solutions is estimated over iterations using the discrete-event simulator. Accordingly, the computation time of such an evaluation process inevitably greatly increases with the growth of the number of candidate solutions or simulation replications. The main disadvantage of SBM technique therefore lies in the large computation time required for performance evaluation under uncertainties (Dugardin et al., 2010).

To overcome such drawback, some effective approaches have recently been proposed for scheduling under uncertainties. Instead of estimating the performance of all candidate solutions, these approaches only evaluate a number of promising candidate solutions by a time-consuming simulator. Zhang, Song, and Wu (2012) developed a hybrid particle swarm optimisation (PSO) algorithm for stochastic job shop scheduling problems. They first adopted the lower bound of the objective value to give a quick performance evaluation on candidate solutions, and then only the ones in the satisfactory regions were further estimated using a discrete-event simulator. Moreover, to determine the promising candidate solutions for further optimisation under uncertainties, both Horng, Lin, and Yang (2012) and Juan, Barrios, Vallada, Riera, and Jorba (2014) applied the stochastic simulation model with less simulation replications for performance evaluation.

Despite an increasing amount of research interest in developing new SBMs, few research works have been reported on improving their computational performance for scheduling problems under uncertainties.

This paper therefore presents an effective SBM to address the PFSPs under stochastic processing times. Enlightened by the works of Zhang et al. (2012), Horng et al. (2012), and Juan et al. (2014), we incorporate an efficient two-stage simulation-based model into a hybrid estimation of distribution algorithm (EDA) to generate good-quality schedules with less computational effort.

The EDA was first introduced by Mühlenbein and Paass (1996) as an alternative to conventional evolutionary algorithms (EA). Different from conventional EAs, EDA adopts a probabilistic model to generate the offspring. This model is established by learning from an elite set of individuals in the previous population. As an effective method to inherit good genes over generations, EDA has recently been successfully used to a wide range of combinatorial optimisation problems (Hauschild & Pelikan, 2011), such as the PFSP and its variants. Jarboui, Eddaly, and Siarry (2009) presented an efficient EDA to solve the PFSP with total flow time minimisation. For the same scheduling problem, Zhang and Li (2011) improved the EDA efficiency by incorporating the longest common subsequence into the probabilistic model. Wang, Wang, Liu, and Xu (2013) developed an effective EDA to minimise the makespan of the distributed permutation flowshop. More recently, Ceberio et al. (2014) introduced a probabilistic distance-based ranking exponential model, named the Mallows model, to construct EDA solutions. To further investigate the performance of EDA for the PFSP, hybridisation of EDA with other meta-heuristics has also been studied. Liu, Gao, and Pan (2011) hybridised EDA with PSO to allow social information sharing among candidate solutions. Moreover, Tzeng, Chen, and Chen (2012) incorporated the idea of ant colony system (ACS) into EDA to schedule a permutation flowshop.

The proposed two-stage simulation-based hybrid EDA (TSSB-HEDA) differentiates itself from the conventional EDA by two mechanisms, namely a two-stage simulation model (TSSM) and a self-adaptive learning mechanism (SALM). To reduce the computation cost of TSSB-HEDA, TSSM first employs a regression-based meta-model to provide a rough estimation of candidate solutions, and only a number of promising ones are identified and further evaluated using a discrete-event simulator. Moreover, to prevent EDA from early search stagnation, TSSB-HEDA employs both the probabilistic model of EDA and genetic operators of GA to produce offspring individuals. Motivated by the idea of neural network training, SALM dynamically adjusts the ratio of offspring generated by the probabilistic model to avoid being trapped into premature convergence. An extensive search of literature on PFSP suggests that not much research effort has been devoted to applying EDA to schedule the permutation flowshop under uncertainties.

The rest of the paper is organised as follows. Section 2 presents the mathematical formulation of PFSP. Section 3 describes the proposed TSSB-HEDA in details. To validate the performance of TSSB-HEDA under stochastic processing times, simulations are conducted and the computation results are analysed in Section 4. Finally, in Section 5, we conclude the paper and discuss some topics for future research.

2. Problem description

The PFSP is a well-known combinatorial optimisation problem. In the classical PFSP, a finite set $J = \{1, 2, \ldots, n\}$ of n jobs are firstly released simultaneously to the shop floor, and then are processed on a finite set $M = \{m_1, m_2, \ldots, m_m\}$ of m machines with no preemption allowed. Each job j, $j \in J$, consists of m operations that have to be processed on the machines in the order of $m_1, m_2, \ldots, m_m$. All the jobs have deterministic processing times and follow the same processing order on each machine.

In the real-world manufacturing environments, however, a variety of unexpected events, such as tool wear, equipment failure, operator unavailability, and quality issues, may lead to uncertain processing times (Lawrence & Sewell, 1997). This paper describes the processing time uncertainty using the level of processing time variation (LPTV), which is described as follows:

$$LPTV = \sigma / E(P) \tag{1}$$

where $E(P)$ and σ indicate the expected value and the standard deviation of processing time, respectively. According to formula (1), a larger LPTV may result in a large deviation between the expected and the actual processing times. For example, suppose $E[P]$ of a job equals 15 time units, LPTV values of 0.2 and 0.4 lead the standard deviation of actual processing time from $E[P]$ to be 3 and 6 times units respectively.

The objective of PFSP in this study is to determine a feasible permutation π to minimise the makespan, i.e. the maximum completion time of all operations. With consideration of processing time uncertainty, we formulate the PFSP as follows:

$$min\{E[C(\pi_n, m)]\} \tag{2}$$

Subject to the following constraints:

$$C(\pi_1, 1) = SP(\pi_1, 1) \tag{3}$$

$$C(\pi_j, 1) = C(\pi_{j-1}, 1) + SP(\pi_j, 1), \quad j = 2, \ldots, n \tag{4}$$

$$C(\pi_1, i) = C(\pi_1, i-1) + SP(\pi_1, i), \quad i = 2, \ldots, m \tag{5}$$

$$C(\pi_j, i) = max\{C(\pi_{j-1}, i), C(\pi_j, i-1)\} + SP(\pi_j, i), \quad i = 2, \ldots, m; \ j = 2, \ldots, n \tag{6}$$