Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Operational Research 157 (2004) 307–321
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/dsw

Discrete Optimization

Development and analysis of cost-based dispatching rules for job shop scheduling

M.S. Jayamohan, Chandrasekharan Rajendran *

Industrial Engineering and Management Division, Department of Humanities and Social Sciences, Indian Institute of Technology Madras, Chennai 600 036, India

Received 7 September 2000; accepted 21 January 2003

Abstract

Most dispatching rules for job shop scheduling assume that the cost of holding per unit time is the same for all jobs. Likewise, it is assumed that the cost of tardiness per unit time is the same for all jobs. In other words, it is implied that the holding cost of a job is directly proportional to its flowtime, and the tardiness cost of a job is directly proportional to its positive lateness. These assumptions may not hold good in all situations. Some attempts were made to overcome this deficiency, and a couple of dispatching rules were proposed by considering different weights or penalties for different jobs. However, these dispatching rules assume that the holding and tardiness costs per unit time of a given job are the same, even though these costs may differ from job to job in practice. In this study, we propose dispatching rules by explicitly considering different weights or penalties for flowtime and tardiness of a job. Many measures of performance related to weighted flowtime and weighted tardiness of jobs are considered, and the results of simulation are presented.

Keywords: Job shop; Scheduling; Weighted flowtime; Weighted tardiness; Dispatching rules

1. Introduction

The problem of scheduling in dynamic job shops has been extensively investigated over many years. In spite of the research work that has been carried out in the area of job shop scheduling, the problem remains attractive to academicians and practitioners. The primary reasons are that the performance of a dispatching rule will be influenced by various parameters such as the utilization level of shop floor and allowance factor, and that no single rule has been found to be the best under all conditions. Various state-of-the-art survey articles (see Day and Hottenstein, 1970; Panwalker and Iskander, 1977; Blackstone et al., 1982; Haupt, 1989; Ramasesh, 1990) discuss some critical aspects of simulation experiments, impact of due-date setting methodologies, allowance factors and shop floor utilization on the performance of dispatching rules by

* Corresponding author. Tel.: +91-44-22578431; fax: +91-44-22578430.
E-mail address: craj@iitm.ac.in (C. Rajendran).

0377-2217/$ - see front matter
doi:10.1016/S0377-2217(03)00204-2

considering several objectives of scheduling. The choice of a dispatching rule also depends upon the objective of scheduling. There are objectives of scheduling that are time based. Time-based objectives are commonly considered in the literature because these measures of performance are surrogate measures for cost-based measures of performance. Some of the common time-based measures of performance are the minimization of mean, maximum and variance of flowtime of jobs, percentage of tardy jobs, and mean, maximum and variance of tardiness of jobs. Some of the popular rules are the SPT, EDD and FIFO rules. The research on dispatching rules has seen the development of many dispatching rules that are more efficient than these popular dispatching rules. Some of the noteworthy rules are the MOD rule (see Baker and Kanet, 1983), ATC rule (see Vepsalainen and Morton, 1987), COVERT rule (see Russell et al., 1987), CR + SPT and S/RPT + SPT rules (Anderson and Nyirenda, 1990), and RR rule (Raghu and Rajendran, 1993). Recently, Holthaus and Rajendran (1997), Rajendran and Holthaus (1999), and Holthaus and Rajendran (2000) have come up with new rules considering the objectives of minimizing the mean, maximum and variance of flowtime, and maximum and variance of tardiness of jobs. Their studies have included a number of existing rules in the performance evaluation and the results have shown the superiority of the new rules in terms of many measures of performance.

In the era of global competition, the focus of any firm's operation is on satisfying (and more on delighting) its customers. It is quite natural that any firm will have many customers for whom the firm attaches importance to varying degrees. Accordingly, different tardiness penalties or weights are assigned to jobs of different customers so as to measure the true performance of the firm with respect to overall customer service. Some times, a multiple class or priority system is developed by firms to categorize different jobs. However, many times, it is practically easy to assign a customer importance index in the form of job tardiness penalties so that the resulting scalar tardiness penalty values can be incorporated in scheduling rules. Specifically, we consider the case where tardiness penalties can be measured on a numerical and continuous scale. Likewise, the costs of holding different jobs per unit time will also be different. It is therefore important that any scheduling decision should reflect these 'penalties or weights' of jobs when dispatching rules are made use of in shop floor operations. In such cases, it is a customary practice in the literature on job shop scheduling to associate weights for flowtime and tardiness of jobs. The measures of performance of dispatching rules would then be the minimization of weighted mean flowtime, maximum weighted flowtime, weighted variance of flowtime, weighted mean tardiness, maximum weighted tardiness and weighted variance of tardiness of jobs.

Some attempts have been made to develop dispatching rules for minimizing the weighted flowtime and weighted tardiness of jobs in job shop scheduling problems. The weighted version of the SPT rule (i.e. dividing the process time of a job for a given operation by the job weight, and choosing the job with the minimum weighted process time) is a popular rule for minimizing the weighted mean flowtime. Vepsalainen and Morton (1987) reported a notable work on the development of dispatching rules minimizing the weighted mean tardiness of jobs. They presented the results of performance of the weighted version of apparent tardiness cost rule, called the WATC rule, and the weighted COVERT rule. They observed that the use of a look-ahead parameter that could scale the job slack (in place of the linear slack in COVERT rule) and the exponential function of the job slack resulted in enhancing the performance of the WATC rule. Jensen et al. (1995) conducted simulation experiments to investigate the effect of the shape and dispersion of tardiness penalty distribution on the performance of a dispatching rule. It is concluded that the shape of tardiness penalty distribution does not affect choice of a dispatching rule, and that the performance of a dispatching rule is only mildly affected by the dispersion of tardiness penalty distribution. Jensen et al. also observed that the WATC and weighted COVERT rules emerged to the best with respect to the minimization of weighted mean tardiness of jobs. Further investigation on the performance of WATC and weighted COVERT rules has been carried out by Kutanoglu and Sabuncuoglu (1999).

The present work focuses on the development of new rules that incorporate weights for flowtime and tardiness of jobs. This work is perhaps the first of its kind in developing dispatching rules to address the problem of scheduling with different weights for flowtime and tardiness of jobs. The motivation for addressing this type of problem is that the costs of holding and tardiness of a given job per unit time need not be the same in real-life situations. The primary reason for this difference is that the holding cost (or carrying cost) and the penalty cost for belated delivery of a job need not necessarily be related. As it is evident from the literature survey, it is a customary practice for researchers to assume that the holding and tardiness costs per unit time of a given job to be the same. The main reason for such an assumption is to simplify the development of dispatching rules. However, in reality, these cost components of any job can be different, and hence, this study attempts to develop efficient dispatching rules to minimize the measures related to weighted flowtime and weighted tardiness of jobs.

2. Terminology and measures of performance

To understand the various rules in this study, we present the terminology used in this paper as follows:

τ	time instant at which dispatching decision is made
T_{ij}	process time for operation j of job i
m_i	total number of operations on job i
D_i	due-date for job i
ODD_{ij}	operation due-date for operation j of job i
A_i	time of arrival of job i in the shop floor
Z_i	priority index of the job i at instant τ (the job with the least Z_i is chosen for loading)
Z_i'	priority index of the job i at instant τ (the job with the largest Z_i' is chosen for loading)
$C_{i,j-1}$	completion time of the previous operation (i.e., operation $j-1$) of job i
c	due-date allowance factor
h_i	penalty or weight relating to flowtime of job i (or equivalently, the per-unit-time cost of holding job i)
r_i	customer priority or weight relating to tardiness of job i (or equivalently, the per-unit-time cost of tardiness of job i)

When n jobs are to be scheduled, the aggregate weighted performance measures relating to flowtime are the weighted mean flowtime, maximum weighted flowtime, and weighted variance of flowtime of jobs. The weighted performance measures relating to tardiness are the weighted mean tardiness, maximum weighted tardiness and weighted variance of tardiness of jobs. In addition, the objective of minimizing the percentage of tardy jobs is considered in the current study.

3. Identification of rules for evaluation

We first identify the best known rules for minimizing various weighted measures of performance, followed by the development of the proposed rules.

3.1. Best existing rules

The best existing rules for minimizing the weighted mean flowtime and weighted mean tardiness of jobs are given below.

3.1.1. WSPT (weighted shortest process time) rule

The SPT rule is probably one of the simplest and oldest in design, which is commonly used to minimize mean flow time and percentage of tardy jobs (see Conway, 1965; Baker, 1974; Rochette and Sadowski, 1976; Blackstone et al., 1982; French, 1982; Haupt, 1989). The WSPT (weighted shortest process time) rule has been identified in the literature to perform very well in terms of minimizing the weighted mean flow time, as well as reducing the percentage of tardy jobs, especially under highly loaded conditions (see Vepsalainen and Morton, 1987; Jensen et al., 1995).

In this study, the WSPT rule is denoted as the WSPTF rule, when used as a benchmark rule for minimizing the weighted mean flow time, and the priority index of job i is given as follows:

$$Z_i = T_{ij}/h_i. \tag{1}$$

The same rule, when used as benchmark rule for minimizing the weighted mean tardiness, is denoted as the WSPTT rule in this study, with the priority index for jobs calculated as follows:

$$Z_i = T_{ij}/r_i. \tag{2}$$

The job with the minimum value of Z_i is chosen for loading in both cases. The WSPTF rule seeks to minimize the weighted mean flow time, and the WSPTT rule aims at minimizing the weighted mean tardiness of jobs.

3.1.2. WCOVERT (weighted cost over time) rule

This is one of the most popular rules in the current literature, mostly used in the unweighted version (see Russell et al., 1987; Holthaus and Rajendran, 1997). The weighted version of the rule, incorporating weights for customer priorities relating to tardiness, is also quite popular (see Vepsalainen and Morton, 1987; Jensen et al., 1995).The priority index of job i, according to the COVERT rule (see Russell et al., 1987 for details), is given as follows:

$$Z'_i = \begin{cases} (r_i/T_{ij})\left(\sum_{q=j+1}^{m_i} W'_{iq} - s_i\right) \Big/ \sum_{q=j+1}^{m_i} W'_{iq} & \text{if } 0 \leqslant s_i < \sum_{q=j+1}^{m_i} W'_{iq}, \\ 0 & \text{if } s_i \geqslant \sum_{q=j+1}^{m_i} W'_{iq}, \\ (r_i/T_{ij}) & \text{if } s_i < 0. \end{cases} \tag{3}$$

Note that W'_{iq} denotes the expected waiting time of job i for its operation q, and s_i denotes the slack of job i. The job with the largest value of Z'_i is chosen for loading. In the study by Vepsalainen and Morton (1987), this rule is stated to perform well in dynamic job shops, in terms of tightening the job tardiness and reducing the percentage of tardy jobs.

3.1.3. WATC (weighted apparent tardiness cost) rule

This rule, presented by Vepsalainen and Morton (1987), is a very popular benchmark in literature in view of reducing the percentage of tardy jobs and mean tardiness (Pinedo, 1995). This rule is an extension to the popular rule COVERT (Carroll, 1965) by the introduction of an exponential look-ahead (instead of the linear look-ahead in COVERT) that scales the slack according to the expected number of competing jobs. A value of $k = 2$ for static shops and a value of $k = 3$ for dynamic shops have been recommended as the parameter values to be used in the rule. This exponential look-ahead works by ensuring the timely completion of short duration jobs (steep increase of priority close to the due-date) and by extending the look-ahead far enough to prevent long tardy jobs from overshadowing long clusters of shorter jobs (see Vepsalainen and Morton, 1987 for details). The priority index of job i is computed as follows:

$$Z'_i = (r_i/T_{ij}) \times \left\{ \exp\left\{ -\max\left\{ \left[D_i - \sum_{q=j+1}^{m_i} (W'_{iq} + T_{iq}) - \tau - T_{ij} \right] \Big/ (k \times T_{\text{bar}}) \right\}; 0 \right\} \right\}, \tag{4}$$

where k is the exponential look-ahead parameter to scale the slack according to the expected number of competing jobs and T_{bar} is the average processing time of the imminent operations of the competing jobs at the service facility. In this study, parameter values are set in accordance with the recommendations and observations of Vepsalainen and Morton (1987), and Anderson and Nyirenda (1990). The job with the maximum value of Z_i' is taken up for loading.

3.2. Proposed rules

3.2.1. Proposed rule 1: WSLACK rule

We modify the conventional slack rule (used in the unweighted form for minimizing the maximum tardiness and variance of tardiness of jobs) by taking into account the weights for tardiness of jobs. The proposed weighted version of the slack rule is presented below.

$$Z_i = \begin{cases} s_i \times r_i & \text{if } s_i \leqslant 0, \\ s_i / r_i & \text{otherwise.} \end{cases} \tag{5}$$

The slack of job i, s_i, is computed as follows:

$$s_i = D_i - \tau - \text{sum of process times of remaining operations of job } i. \tag{6}$$

The job with the minimum value of Z_i is taken up for loading. This rule attempts to minimize the maximum weighted tardiness and weighted variance of tardiness of jobs.

3.2.2. Proposed rule 2: WFDD rule

A concept, called flow due-date, is used in this paper. We define a due date of completion for every operation of a job by considering the arrival time of the job and its process times. The flow due-date of job i for operation j, denoted by FDD_{ij}, is defined as follows:

$$\mathrm{FDD}_{ij} = A_i + \sum_{k=1}^{j} T_{ik}. \tag{7}$$

By considering the weight for flow time, the priority index of job i for operation j is computed as follows:

$$Z_i = \begin{cases} (\mathrm{FDD}_{ij} - \tau) \times h_i & \text{if } \mathrm{FDD}_{ij} \leqslant \tau, \\ (\mathrm{FDD}_{ij} - \tau)/h_i & \text{otherwise.} \end{cases} \tag{8}$$

The job with the minimum value of Z_i is taken for loading. In this rule, we regulate the job completion according to the job's flow due date, apart from reckoning with the weight for flow time. Hence, this rule attempts to yield minimum values for the maximum weighted flow time and weighted variance of flow time of jobs. The rule is called the WFDD rule.

3.2.3. Proposed rule 3: WODD rule

This proposed rule is the weighted version of the ODD rule developed by Kanet and Hayya (1982). We call this weighted version as the *WODD* rule. According to this rule, the priority index of job i is obtained as follows:

$$Z_i = \begin{cases} (\mathrm{ODD}_{ij} - \tau) \times r_i & \text{if } \mathrm{ODD}_{ij} \leqslant \tau, \\ (\mathrm{ODD}_{ij} - \tau)/r_i & \text{otherwise,} \end{cases} \tag{9}$$

where

$$\mathrm{ODD}_{ij} = A_i + c \times \sum_{k=1}^{j} T_{ik}. \tag{10}$$

The job with the minimum value of Z_i is taken for loading. This rule is aimed at minimizing the maximum weighted tardiness and weighted variance of tardiness of jobs. Basically, this rule seeks to regulate the job completion at every operation according to the operation due date of the job, apart from reckoning with the weight for job tardiness.

3.2.4. Proposed Rule 4: PT + PW(WF) rule

The flow time of a job is determined by process times and waiting times of the job for different operations. In order to minimize the flow time, we consider the process time of the job for the imminent operation and the waiting time of the job at the current operation. In addition, we consider the weight for job flow time. In the process, we attempt the minimization of the weighted mean flowtime of jobs. The rule, thus developed, is called the PT + PW(WF) rule. The priority index is computed as follows (with the job having the minimum Z_i being chosen):

$$Z_i = [T_{ij} + (\tau - C_{i,j-1})]/h_i. \tag{11}$$

3.2.5. Proposed rule 5: PT + PW(WF + WT) rule

The PT + PW(WF) rule is developed to minimize the weighted mean flow time of jobs. In order to improve the performance of the rule with respect to the minimization of the weighted mean tardiness of jobs as well, we modify the rule in the following way:

$$Z_i = \begin{cases} [T_{ij} + (\tau - C_{i,j-1})]/(h_i) & \text{if } \tau \leqslant \mathrm{ODD}_{ij}, \\ [T_{ij} + (\tau - C_{i,j-1})]/(h_i + r_i) & \text{otherwise.} \end{cases} \tag{12}$$

This rule, called PT + PW(WF + WT) rule, reckons with the weight for tardiness only when the present time τ is more than the operation due-date of the job. Thus, the rule attempts to perform well with respect to the minimization of the weighted mean flow time and weighted mean tardiness of jobs. The job with minimum value of Z_i will be chosen as the next job for loading.

3.2.6. Proposed rule 6: W(PT + PW + ODD) rule

The priority index for this proposed rule is as follows:

$$Z_i = [T_{ij} + (\tau - C_{i,j-1}) - (\tau - \mathrm{ODD}_{ij})]/r_i,$$

or in simpler terms, (13)

$$Z_i = [T_{ij} - C_{i,j-1} + \mathrm{ODD}_{ij}]/r_i.$$

This rule is a modification of the previous rule with a focus on tardiness related measures. The purpose of addition of ODD_{ij} is to maintain the tardiness of different jobs at low levels, thereby minimizing the weighted mean tardiness. The job with the minimum value of Z_i is selected for loading.

In all, we have identified three existing rules, and proposed six new rules.

In all the cases, if the priority index of job i is indicated as Z_i, then the job with the minimum value of Z_i is chosen for loading. If the index of job i is symbolized by Z_i', then the job with the maximum value Z_i' will be chosen for loading.

4. Experimental design for the simulation study

Job shops can be classified into two: open and closed shops, based on routeing patterns existing in the shop. If jobs on arrival follow one of the fixed routeings, it is a closed job shop, and if the jobs have no restriction on routings, it is an open job shop. Our study focuses on open job shop configurations with standard assumptions (Baker, 1974; French, 1982; Haupt, 1989; Ramasesh, 1990). We assume the presence of ten machines in the shop. In both the phases of the study, the number of operations for an entering job is randomly sampled in the range 5–9, and the corresponding machine visitations are randomly generated among ten machines. It is to be noted that no two consecutive operations can be performed on the same machine and that a machine can be revisited by a job for a later operation.

In all our experimental studies, process times are sampled from a discrete uniform distribution in the range 1–50. The total work-content method of due-date setting is utilized (see Blackstone et al., 1982; Kanet and Hayya, 1982; Baker, 1984; Ragatz and Mabert, 1984; Haupt, 1989) with allowance factors of 3, 5 and 7 to represent different levels of due-date settings. The job arrivals are generated using an exponential distribution for interarrival times. Two levels of machine utilization are tested in the experiments, viz., 85% and 95%. Weights or penalties per unit of flowtime (representative of holding costs) are randomly sampled in the range 1–9. Likewise, tardiness weights (representative of penalties of job completion time beyond the due-date) are randomly sampled in the range 1–9. It has been found that dispatching rules are insensitive to both the dispersion and the shape of the tardiness penalty distribution (Jensen et al., 1995). Hence we have chosen to sample the weights for flowtime and tardiness in this manner. Thus, we have three due-date settings and two different utilization levels, thereby making six different simulation experiments for every dispatching rule.

In our study, each simulation experiment consists of twenty replications (or runs). In each replication, shop is continually loaded with job-orders numbered on their arrival. In order to ascertain when the system attains steady state, we have observed shop parameters such as utilization level values, weighted mean flowtime of jobs, etc. It has been observed that steady state has been reached in the shop after the arrival of about 500 job orders. Typically, the total sample size of a job shop simulation study is of the order of tens of thousands of job completions (Conway et al., 1960; Blackstone et al., 1982). For a given sample size, it is preferable to have a smaller number of replications and a larger run length (Law and Kelton, 1984). Following these guidelines, the number of replications and the run-length are taken as 20 and 1500 completed job-orders per run respectively. Statistical analysis of the experimental data using the one-way ANOVA and the Duncan's multiple range tests (see Montgomery, 1991; Lorenzen and Anderson, 1993) has shown the chosen sample size yield a variance which results in a Type I error of at most 1%. As for the computation of statistics from a given replication, we have collected data from job-orders numbered 501–2000, and the shop is continued to be loaded with jobs until the completion of these 1500 numbered job orders. This procedure helps us to overcome the problem of censored data (see Conway, 1965). Duncan's multiple range test has been used to identify the best set (marked with * in the tables) and second best set of mutually exclusive rules (marked with # in the tables), *after conducting the test on the four best performing rules in each category of performance measures*. The simulation program has been written in C++ and implemented on an IBM RS-6000 system on a UNIX environment.

5. Results and discussion

The performance of all the rules under study are evaluated with respect to weighted mean, maximum weighted, and weighted variance of flowtime; weighted mean, maximum weighted and weighted variance of tardiness; and percentage of tardy jobs in the shop. The absolute values of performance measures have been obtained by taking the mean of the values obtained for the twenty replications carried out for each of the settings. *It is to be noted that even though every proposed rule has been developed for a specific objective (or for a specific set of objectives), we have chosen to observe how various dispatching rules perform with respect to all objectives under consideration. In other words, every dispatching rule is evaluated with respect to all measures of performance in order to obtain a complete picture of a rule's overall performance.*

After a statistical analysis of the absolute values of the best four performing rules with respect to a given performance measure using the one-way ANOVA and Duncan's multiple range tests, we have tabulated some typical results in Tables 1–4 for different utilization levels and due-date settings. Our studies reveal

Table 1
Performance of rules with respect to weighted flow time and weighted tardiness (utilization level: 95%; c : 3)

Rule	wmean_F	wmax_F	wvar_F	wmean_T	wmax_T	wvar_T	% Tardy
WSPTF	#672.27	#26 957	#16 064 500	522.88	134 783	127 345 681	#45.60
WSPTT	1284.53	135 954	41 003 408	#269.93	#17 478	#7 023 457	#45.67
WATC	1009.64	#15 300	#4 962 606	485.30	#10 287	#2 123 268	94.87
WCOVERT	1055.70	41 895	39 034 074	459.71	22 560	10 525 436	94.00
WSLACK	976.67	35 793	28 103 034	335.19	*4768	*480 047	96.07
WFDD	834.03	*8872	*1 626 130	583.56	28 672	23 453 418	90.47
WODD	1153.36	36 540	29 030 405	416.57	*4419	*403 325	98.87
W(PT + PW + ODD)	#696.62	51 516	59 104 050	*155.57	20 960	12 343 218	46.80
PT + PW(WF)	*531.66	#24 813	#14 104 030	#295.65	96 345	212 334 352	*37.40
PT + PW(WF + WT)	#*567.52	38 052	32 356 732	#*192.35	33 040	20 342 345	#*41.53

Table 2
Performance of rules with respect to weighted flow time and weighted tardiness (utilization level: 95%; c : 5)

Rule	wmean_F	wmax_F	wvar_F	wmean_T	wmax_T	wvar_T	% Tardy
WSPTF	#672.27	26 957	16 064 500	291.20	129 563	121 334 451	#27.40
WSPTT	1284.53	135 954	41 003408	212.29	#16 914	#6 324 567	28.20
WATC	996.05	51 786	59 075 643	237.67	22 987	234 334 55	51.53
WCOVERT	1032.50	#20 781	#9 292 235	199.69	#13 086	#4 654 557	63.47
WSLACK	982.46	#20 736	#9 271 432	#126.86	*3703	*299 201	70.13
WFDD	834.03	*8872	*1 626 130	347.43	26 032	24 334 223	50.80
WODD	1200.00	35 424	27 107 344	218.93	*3612	#*277 945	86.40
W(PT + PW + ODD)	#715.15	51 237	58 273 436	*92.35	21 462	13 343 456	*21.27
PT + PW(WF)	*531.66	#24 813	#14 104 030	#195.61	93 753	211 443 528	*20.87
PT + PW(WF + WT)	#*566.66	38 169	32 136 732	#*110.07	24 773	11 034 537	#*22.07

Table 3
Performance of rules with respect to weighted flow time and weighted tardiness (utilization level: 85%; c : 3)

Rule	wmean_F	wmax_F	wvar_F	wmean_T	wmax_T	wvar_T	% Tardy
WSPTF	#433.72	#*9429	#*1 919 431	210.75	54 712	72 334 435	30.60
WSPTT	616.00	61 101	83 045 007	80.74	#7356	#1 213 452	#30.27
WATC	566.47	22 168	11 112 387	101.32	10 571	3 423 342	46.73
WCOVERT	561.65	#16 002	#5 583 092	93.32	#7424	#1 435 645	47.73
WSLACK	550.60	#12 015	#3 117 143	#64.18	*2488	*136 103	53.53
WFDD	490.80	*6867	*994 667	129.51	13 401	4 123 434	44.20
WODD	596.58	17 248	6 489 389	#77.45	*2395	*125 377	62.00
W(PT + PW + ODD)	#454.72	25 146	14 345 678	*37.03	9375	2 233 348	*24.53
PT + PW(WF)	*381.29	17 127	6 470 057	87.48	17 307	7 234 356	*23.13
PT + PW(WF + WT)	*386.05	17 640	6 865 780	#52.44	16 200	6 234 534	*23.93

Table 4
Performance of rules with respect to weighted flow time and weighted tardiness (utilization level: 85%; c : 5)

Rule	wmean_F	wmax_F	wvar_F	wmean_T	wmax_T	wvar_T	% Tardy
WSPTF	#433.72	#9429	#*1 919 431	128.21	50 834	63 577 665	16.13
WSPTT	616.00	61 101	83 045 007	42.91	5984	800 558	15.53
WATC	544.64	#9801	#2 055 899	16.38	#4352	#424 636	11.07
WCOVERT	543.15	14 904	4 837 316	#4.17	#3584	#288 715	*4.40
WSLACK	561.09	13 959	4 228 962	*1.29	*707	*11 228	#*6.27
WFDD	490.80	*6867	*994 667	46.85	11 475	3 343 443	13.80
WODD	630.34	#13 167	#3 736 595	#9.22	#*1350	#*40 758	14.13
W(PT + PW + ODD)	#461.57	26 334	15 344 634	#14.72	9612	2 341 223	#*7.33
PT + PW(WF)	*381.29	17 127	6 470 057	45.86	18 808	8 213 456	9.27
PT + PW(WF + WT)	*385.49	22 671	11 084 654	22.50	13 986	4 223 454	#9.00

that there is no marked difference between the relative performances of the various rules with the due-date allowance factors of 5 and 7. Hence, we do not report the results corresponding to $c = 7$. We now address the performance among the various rules when jobs have independent weights for flowtime and tardiness.

5.1. Performance evaluation of the rules

5.1.1. Weighted mean flowtime (see under the column wmean_F in tables)

The proposed rule PT + PW(WF) closely followed by the PT + PW(WF + WT), appear to be the best (or not significantly worse than the best performing rule) in almost all cases. The reason for the good performance of the PT + PW(WF) and PT + PW(WF + WT) rule is that these seek to hasten the job with the minimum process time plus the waiting time and large weight for flowtime, thereby leading to the minimum sum of weighted flowtimes of jobs.

5.1.2. Maximum weighted flowtime (see under the column wmax_F in tables)

Among the various rules tested, the new dispatching rule, WFDD emerges to be the best for minimizing the maximum weighted flowtime. This rule performs very well because the rule seeks to minimize the deviation of the job completion time from the job's flow due-date, apart from considering the weight for flowtime of the job, thereby minimizing the maximum weighted flowtime of jobs.

5.1.3. Weighted variance of flowtime (see under the column wvar_F in the tables)

The WFDD rule is the best for this measure in almost all cases. The reason for the good performances of the rules is the same as that ascribed to its performance with respect to maximum weighted flowtime of jobs.

5.1.4. Weighted mean tardiness (see under the column wmean_T in tables)

The best performing rule in this category is the W(PT + PW + ODD) rule, followed by the PT + PW(WF + WT) and WSLACK rules as the next best performing rules. The W(PT + PW + ODD) rule, being process-time as well as due-date based, performs as expected. Likewise, the PT + PW(WF + WT) rule, being primarily a process-time based rule, performs well, especially at high utilization levels and tight due-date settings. This performance is similar to the performance of the SPT rule with respect to the minimization of the unweighted mean tardiness of jobs. The WSLACK rule also performs quite well. The good performance of the WSLACK rule (which is primarily a due-date based rule) and its dominance over the process-time based rules are expected when the due-date settings are loose (see Blackstone et al. (1982) and Haupt (1989) for a similar observation on the performance of the due-date based rules without weights incorporated in them when due-date settings are loose).

5.1.5. Maximum weighted tardiness (see under the column wmax_T in the tables)

The WODD rule at higher levels and the WSLACK rule at lower levels of utilization are the best rules (or significantly not worse than the best performing rules) for this measure of performance. It is also observed that the WSLACK rule, using the slack information, performs very well when the due-date setting is loose. This is in conformance with the fact that slack-based rules perform extremely well at loose due-date settings in the unweighted case (see Blackstone et al., 1982; Haupt, 1989). The WODD rule seeks to minimize the deviation of the completion time of a job from its operation due-date, and hence good performance for this rule is observed for this measure of performance.

5.1.6. Weighted variance of tardiness (see under the column wvar_T in the tables)

The rules that can be rated as the best for this measure are the WSLACK and WODD rules in all cases. Generally, the performance pattern of the dispatching rules for this objective is similar to the rules' performance pattern with respect to minimization of the maximum tardiness of jobs.

5.1.7. Percentage of tardy jobs (see under the column % tardy in the tables)

The proposed rules PT + PW(WF) and PT + PW(WF + WT) appear to be the best for this objective under the conditions of high utilization level and relatively tight due-date setting (i.e., with the utilization level of 95% and $c = 3$ and 5). However, the WCOVERT and WSLACK rules perform very well under other conditions (especially when the utilization level is relatively low at 85% and with loose due-date setting). Our observations are in conformance with the literature findings (see Blackstone et al., 1982; Haupt, 1989) that the process-time based rules perform very well with respect to minimizing the proportion of tardy jobs under the conditions of high utilization level and tight due-date setting.

5.1.8. Normalized measures of weighted flowtime and weighted tardiness

In the current study, we have computed the following normalized measures of performance:

$$\text{normalized weighted flowtime (Norm\_WF)} = (\text{sum of weighted flowtime of 1500 jobs, averaged over 20 replications})/(n \times o \times t \times h), \tag{14}$$

$$\text{normalized weighted tardiness of 1500 jobs (Norm\_WT)} = (\text{sum of weighted tardiness of 1500 jobs, averaged over 20 replications})/(n \times o \times t \times r), \tag{15}$$

$$\begin{aligned}&\text{normalized sum of weighted flowtime and weighted tardiness}\\&\text{(or normalized total cost, denoted by Norm\_TC)}\\&\quad = (\text{sum of weighted flowtime and weighted tardiness of 1500 jobs,}\\&\qquad \text{averaged over 20 replications})/(n \times o \times t \times (h + r)),\end{aligned} \tag{16}$$

where n indicates 1500 jobs taken up for the computation of statistics during the steady state, o indicates the average number of operations (equal to 7 in the current study), t refers to the average process time of a job per operation (equal to 25 in the current study), and h and r refer to the mean weight for holding and mean weight for tardiness respectively (equal to 5 in this phase of the study).

The results are presented in Tables 5–8. It can be observed that the relative performances of different rules for normalized measures of weighted flowtime and weighted tardiness measures are similar to those of rules with respect to minimization of the weighted mean flowtime and weighted mean tardiness respectively. As for the normalized sum of weighted flowtime and weighted tardiness, we find that the

Table 5
Normalized performance of rules (utilization level: 95%; c : 3)

Rule	Norm_WF	Norm_WT	Norm_TC
WSPTF	3.85	2.96	3.45
WSPTT	7.34	1.55	4.45
WATC	5.74	2.75	4.25
WCOVERT	6.07	2.64	4.35
WSLACK	5.61	1.90	3.75
WFDD	4.75	3.37	4.06
WODD	6.62	2.40	4.51
W(PT + PW + ODD)	4.00	0.90	2.46
PT + PW(WF)	3.05	1.71	2.35
PT + PW(WF + WT)	3.26	1.10	2.14

Table 6
Normalized performance of rules (utilization level: 95%; c : 5)

Rule	Norm_WF	Norm_WT	Norm_TC
WSPTF	3.85	1.87	2.95
WSPTT	7.34	1.15	4.26
WATC	5.71	1.40	3.53
WCOVERT	5.89	1.15	3.50
WSLACK	5.61	0.70	3.16
WFDD	4.75	2.00	3.40
WODD	6.88	1.25	4.07
W(PT + PW + ODD)	4.09	0.55	2.30
PT + PW(WF)	3.04	1.13	2.10
PT + PW(WF + WT)	3.25	0.65	1.95

Table 7
Normalized performance of rules (utilization level: 85%; c : 3)

Rule	Norm_WF	Norm_WT	Norm_TC
WSPTF	2.50	1.22	1.85
WSPTT	3.53	0.45	2.00
WATC	3.23	0.70	1.95
WCOVERT	3.21	0.55	1.85
WSLACK	3.13	0.35	1.74
WFDD	2.81	0.75	1.75
WODD	3.39	0.46	1.95
W(PT + PW + ODD)	2.61	0.20	1.40
PT + PW(WF)	2.20	0.52	1.35
PT + PW(WF + WT)	2.20	0.31	1.25

PT + PW(WF + WT) emerges to be the best. It means that the proposed rule PT + PW(WF + WT) is quite effective in minimizing the total cost of scheduling as well. It is to be noted that the proposed rules as well as the existing rules seek to minimize either the weighted flowtime or the weighted tardiness of jobs, and that they do not seek to minimize the total cost of scheduling. The problem of developing dispatching rules to minimize the total cost of scheduling is an interesting area of research, especially from a practical viewpoint, and this aspect can be viewed as a subject for further research. In such a case, the rules proposed in this study can be used as benchmark rules.

Table 8
Normalized performance of rules (utilization level: 85%; c : 5)

Rule	Norm_WF	Norm_WT	Norm_TC
WSPTF	2.51	1.03	1.65
WSPTT	3.52	0.35	1.90
WATC	3.10	0.20	1.60
WCOVERT	3.09	0.05	1.55
WSLACK	3.23	0.05	1.60
WFDD	2.80	0.32	1.55
WODD	3.64	0.15	1.85
W(PT + PW + ODD)	2.63	0.20	1.40
PT + PW(WF)	2.20	0.36	1.23
PT + PW(WF + WT)	2.20	0.24	1.18

5.2. Additional experimentation for performance evaluation of rules

We have conducted additional experiments in which we have sampled the weight for tardiness in the range [1, 99]. We have conducted the simulation experiments with the allowance factor for a job sampled randomly in the range [1, 5] (such that the average allowance factor is 3), and at 85% and 95% utilization levels, such that the settings reflect the some real-life situations in which tardiness penalties are large and due-date allowances for all jobs are not the same. The results of evaluation with respect to the normalized measures of weighted tardiness are presented in Tables 9 and 10. On the whole, the PT + PW(WF + WT), WSPTT and W(PT + PW + ODD) rules emerge to be the best with respect to the normalized measure of weighted tardiness.

5.3. Overall observations

Overall, we find that the PT + PW(WF + WT) rule performs very well with respect to the minimization of the weighted mean flowtime and weighted mean tardiness of jobs, followed by the W(PT + PW + ODD) rule. The same observation holds for the normalized measures of weighted flowtime and weighted tardiness of jobs as well. As for the maximum and standard deviation of weighted flowtime and weighted tardiness of jobs, it can be seen that, on the whole, the WSLACK, WODD, WCOVERT and WATC rules perform very well. However, the choice of a dispatching rule is influenced by shop parameters such as due-date setting and utilization levels, and hence, the shop floor manager can evaluate (with respect to the measure of

Table 9
Normalized performance of rules (utilization level: 95%; $c = [1, 5]$; weight for tardiness $= [1, 99]$)

Dispatching rule	Norm_WT
WSPTF	7.10
WSPTT	1.42
WATC	3.81
WCOVERT	3.96
WSLACK	3.09
WFDD	6.99
WODD	2.73
W(PT + PW + ODD)	2.58
PT + PW(WF)	4.96
PT + PW(WF + WT)	1.75

Table 10
Normalized performance of rules (utilization level: 85%; $c = [1,5]$; weight for tardiness = $[1,99]$)

Dispatching rule	Norm_WT
WSPTF	1.82
WSPTT	0.62
WATC	0.93
WCOVERT	1.06
WSLACK	0.80
WFDD	1.79
WODD	0.77
W(PT + PW + ODD)	0.81
PT + PW(WF)	1.44
PT + PW(WF + WT)	0.57

performance under consideration) one of the effective rules, such as the PT + PW(WF + WT), W(PT + PW + ODD), WSLACK and WCOVERT rules, in the actual shop under study.

5.4. Some specific observations

It has been found that the WSPT rule performs better than the rules based on process-time and due-date (such as the WCOVERT and WATC rules) in congested shops with tight due-date setting (e.g. see Kutanoglu and Sabuncuoglu, 1999). We too observe the same behavior in our study. For example, if we see Tables 1 and 5, the same pattern is exhibited by these rules (note that in the present study, the conventional WSPT rule is termed as WSPTF with respect to the objective of minimizing the weighted mean flowtime, and the WSPT rule is termed as the WSPTT rule with respect to the minimization of the weighted mean tardiness; see the performance evaluation of these rules in Tables 1–4). It is to be noted that at 95% utilization level and with $c = 5$, we find that the performances of WCOVERT, WATC and WSPTT are not significantly different from each other with respect to the weighted mean tardiness (see Table 2).

It is also known (see Kutanoglu and Sabuncuoglu, 1999) that the performance of the WSPTT rule (pure process-time based rule) is not better than the rules based on both process-time and due-date (such as WCOVERT and WATC rules) in less congested shops with loose due-dates with respect to the minimization of the weighted mean tardiness of jobs. For example, if we see Tables 3 and 7, the WSPTT and WCOVERT rules do not differ from each other significantly. However, WATC and WCOVERT perform better than the WSPTT rule when due-dates are loose (see Tables 4 and 8). It is also known (see Kutanoglu and Sabuncuoglu, 1999) that the WSPTT rule is more robust than the WCOVERT and WATC rules in terms of the performance across different experimental conditions (such as different utilization levels and due-date settings, in terms of tight and loose due-dates). We too observe the same pattern of performance of the WSPTT rule in our study.

6. Conclusion

Most of the research on development of dispatching rules for job shop scheduling has assumed that the holding cost of a job is directly proportional to its flowtime, and the tardiness cost of a job is directly proportional to its positive lateness. Only a few dispatching rules have been proposed by considering weights or penalties for holding and tardiness of jobs. However, these dispatching rules assume that the holding and tardiness weights of a given job are the same, even though these weights may differ from job to job in practice. In this study, we have proposed some dispatching rules by explicitly considering different weights for flowtime and tardiness of jobs. This work is perhaps the first of its kind in developing dispatching rules to address the problem of scheduling with different weights or penalties for flowtime and tardiness of jobs. In addition to this aspect, we have considered a number of measures of performance such as the minimization of weighted mean flowtime, maximum weighted flowtime, weighted variance of flowtime, weighted mean tardiness, maximum weighted tardiness, weighted variance of tardiness and the proportion of tardy jobs. This study is also unique in the sense that it makes use of a large number of measures of performance for the evaluation of every dispatching rule. The relative performance of the proposed rules and the existing rules is reported for all these measures. *It is to be noted again that even though every proposed rule has been developed for a specific objective (or for a specific set of objectives), we have chosen to observe how various dispatching rules perform with respect to all objectives under consideration. In other words, every dispatching rule is evaluated with respect to all measures of performance in order to obtain a complete picture of a rule's overall performance.*

Acknowledgements

The authors are thankful to the three reviewers whose comments and suggestions have been valuable to improve the earlier version of the paper.

References

Anderson, E.J., Nyirenda, J.C., 1990. Two new rules to minimize tardiness in a job shop. International Journal of Production Research 28, 2277–2292.

Baker, K.R., 1974. Introduction to Sequencing and Scheduling. Wiley, New York.

Baker, K.R., 1984. Sequencing rules and due-date assignments in a job shop. Management Science 30, 1093–1104.

Baker, K.R., Kanet, J.J., 1983. Job shop scheduling with modified due dates. Journal of Operations Management 4, 11–22.

Blackstone, J.H., Phillips, D.T., Hogg, G.L., 1982. A state of art survey of dispatching rules for manufacturing job shop operations. International Journal of Production Research 20, 27–45.

Carroll, D.C., 1965. Heuristic sequencing and multiple component jobs, Unpublished Ph.D. dissertation, Massachusetts Institute of Technology.

Conway, R.W., 1965. Priority dispatching and job lateness in a job shop. Journal of Industrial Engineering 16, 123–130.

Conway, R.W., Johnson, B.M., Maxwell, W.L., 1960. An experimental investigation of priority dispatching. Journal of Industrial Engineering 11, 221–230.

Day, J.E., Hottenstein, M.P., 1970. Review of sequencing research. Naval Research Logistics Quarterly 17, 11–39.

French, S., 1982. Sequencing and scheduling, an introduction to the mathematics of the JOB shop. Ellis Horwood, Chichester.

Haupt, R., 1989. A survey of priority rule based dispatching. O.R. Spektrum 11, 3–16.

Holthaus, O., Rajendran, C., 1997. Efficient dispatching rules for scheduling in job shop. International Journal of Production Economics 48, 87–105.

Holthaus, O., Rajendran, C., 2000. Efficient jobshop dispatching rules: Further developments. Production Planning and Control 11, 171–178.

Jensen, J.B., Philipoom, P.R., Malhotra, M.K., 1995. Technical note: Evaluation of scheduling rules with commensurate customer priorities in job shops. Journal of Operations Management 13, 213–228.

Kanet, J.J., Hayya, J.G., 1982. Priority dispatching with operation due-dates in a job shop. Journal of Operations Management 2, 167–176.

Kutanoglu, E., Sabuncuoglu, I., 1999. An analysis of heuristics in a dynamic job shop with weighted tardiness objectives. International Journal of Production Research 37, 165–188.

Law, A.M., Kelton, W.D., 1984. Confidence intervals for steady state Simulation: I. A survey of fixed sample survey procedures. Operations Research 32, 1221–1239.

Lorenzen, T.J., Anderson, V.L., 1993. Design of Experiments: A No-Name Approach. Marcel Dekker, New York.

Montgomery, D.C., 1991. Design and Analysis of Experiments, Third ed. John Wiley and Sons, New York.

Panwalker, S.S., Iskander, W., 1977. A survey of scheduling rules. Operations Research 25, 49–61.

Pinedo, M., 1995. Scheduling Theory, Algorithms and systems. Prentice Hall, Englewood Cliffs, NJ.

Ragatz, G.L., Mabert, V.A., 1984. A simulation analysis of due date assignment rules. Journal of Operations Management 5, 27–39.

Raghu, T.S., Rajendran, C., 1993. An efficient dispatching rule for scheduling in a job shop. International Journal of Production Economics 32, 301–313.

Rajendran, C., Holthaus, O., 1999. A comparative study of dispatching rules in dynamic flow shops and job shops. European Journal of Operational Research 116, 156–170.

Ramasesh, R., 1990. Dynamic job shop Scheduling: A survey of simulation research. OMEGA 16, 43–57.

Rochette, R., Sadowski, R.P., 1976. A statistical comparison of the performance of simple dispatching rules for a particular set of job shops. International Journal of Production Research 14, 63–75.

Russell, R.S., Dar-el, E.M., Taylor III, B.W., 1987. A comparative analysis of the COVERT job sequencing rule using various shop performance parameters. International Journal of Production Research 25, 1523–1540.

Vepsalainen, A.P.J., Morton, T.E., 1987. Priority rules for job shops with weighted tardiness costs. Management Science 33, 1035–1047.