A practical method for evaluating worker allocations in large-scale dual resource constrained job shops

Benjamin J. Lobo\textsuperscript{a}, James R. Wilson\textsuperscript{b}, Kristin A. Thoney\textsuperscript{c}, Thom J. Hodgson\textsuperscript{b} & Russell E. King\textsuperscript{b}

\textsuperscript{a} Department of Systems and Information Engineering, University of Virginia, P.O. Box 400747 Olsson Hall, Room 102A Charlottesville, VA 22904-4747, USA E-mail:
\textsuperscript{b} Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University, Campus Box 7906, Raleigh, NC 27695-7906, USA
\textsuperscript{c} Department of Textile and Apparel Technology and Management, North Carolina State University, Campus Box 8301, Raleigh, NC 27695-8301, USA

Accepted author version posted online: 12 Feb 2014. Published online: 28 Jul 2014.

To cite this article: Benjamin J. Lobo, James R. Wilson, Kristin A. Thoney, Thom J. Hodgson & Russell E. King (2014) A practical method for evaluating worker allocations in large-scale dual resource constrained job shops, IIE Transactions, 46:11, 1209-1226, DOI: 10.1080/0740817X.2014.892231

To link to this article: http://dx.doi.org/10.1080/0740817X.2014.892231

PLEASE SCROLL DOWN FOR ARTICLE
In two recent articles, Lobo et al. present algorithms for allocating workers to machine groups in a Dual Resource Constrained (DRC) job shop so as to minimize $L_{\text{max}}$, the maximum job lateness. Procedure LBSA delivers an effective lower bound on $L_{\text{max}}$, while the heuristic HSP delivers an allocation whose associated schedule has a (usually) near-optimal $L_{\text{max}}$ value. To evaluate an HSP-based allocation’s quality in a given DRC job shop, the authors first compute the gap between HSP’s associated $L_{\text{max}}$ value and LBSA’s lower bound. Next they refer this gap to the distribution of a “quasi-optimality” gap that is generated as follows: (i) independent simulation replications of the given job shop are obtained by randomly sampling each job’s characteristics; and (ii) for each replication, the associated quasi-optimality gap is computed by enumerating all feasible allocations. Because step (ii) is computationally intractable in large-scale problems, this follow-up article formulates a revised step (ii) wherein each simulation invokes HSP2, an improved version of HSP, to yield an approximation to the quasi-optimality gap. Based on comprehensive experimentation, it is concluded that the HSP2-based distribution did not differ significantly from its enumeration-based counterpart; and the revised evaluation method was computationally tractable in practice. Two examples illustrate the use of the revised method.

Keywords: Job shop scheduling, dual resource constrained system, worker allocation, maximum job lateness

1. Introduction

In conventional job shop scheduling, by assumption the system’s ability to complete jobs on time is constrained only by the number of machines that are available to process jobs. This perspective on the job shop scheduling problem ignores the additional constraint imposed by the number of workers that are available to operate the machines. Dual Resource Constrained (DRC) systems are subject to limits on the availability of both machines and manpower (Treleven and Elvers, 1985). Lobo et al. (2013a, 2013b) address the problem of minimizing $L_{\text{max}}$, the maximum job lateness, in a DRC job shop. Because the job shop scheduling problem with even a single constrained resource is NP-hard (Lenstra and Rinnooy Kan, 1979), Lobo et al. (2013a, 2013b) develop search procedures (mainly heuristics) to find the following: (i) the most promising (hoped to be optimal) allocation of available workers to the machine groups (departments) in the job shop; and (ii) the best achievable schedule for the job shop based on allocation (i). The objective of this article is to develop a practical method for applying the approach of Lobo et al. (2013b) to large-scale DRC job shop problems.

Given an allocation $\vartheta$ of workers to machine groups, Lobo et al. (2013a) derive a corresponding lower bound $\text{LB}^\vartheta$ on the value of $L_{\text{max}}$ for all schedules based on that allocation. They also develop LBSA, a search algorithm that is guaranteed to deliver an allocation $\vartheta^*$ whose corresponding lower bound $\text{LB}^\vartheta^*$ exactly equals the smallest value of $\text{LB}^\vartheta$ taken over all feasible allocations. The authors use $\text{LB}^\vartheta^*$ as a benchmark for evaluating heuristic solutions to the DRC job shop scheduling problem. Lobo et al. (2013b) establish criteria for verifying that an allocation is optimal—i.e., the allocation corresponds
to a schedule that minimizes $L_{\text{max}}$. For situations in which the LBSA-delivered allocation $\vartheta^*$ does not satisfy the optimality criteria, Lobo et al. (2013b) develop HSP, a heuristic search procedure that delivers an allocation $\vartheta^{\text{HSP}}$, which generally results in an improved schedule. Specifically, $\vartheta^{\text{HSP}}$ usually enables the Virtual Factory (an iterative heuristic scheduler developed by Hodgson et al. (1998)) to generate a schedule whose $L_{\text{max}}$ value, denoted by $\text{VF}_{\vartheta^{\text{HSP}}}$, is less than $\text{VF}_{\vartheta^*}$, the corresponding $L_{\text{max}}$ value for the Virtual Factory–generated schedule based on allocation $\vartheta^*$. Ideally, $\vartheta^{\text{HSP}}$ enables the Virtual Factory to generate a schedule for which $\text{VF}_{\vartheta^{\text{HSP}}}$ is close not only to the minimum feasible value of $L_{\text{max}}$ but also to the lower bound $\text{LB}_{\vartheta^*}$. The Virtual Factory is used as the heuristic scheduler because of its proven track record in successfully generating optimal or near-optimal schedules in job shop scheduling problems for which the primary objective is to minimize $L_{\text{max}}$ (Weintraub et al., 1999; Hodgson et al., 2000; Thoney et al., 2002; Zozom et al., 2003; Hodgson et al., 2004; Schultz et al., 2004). Nevertheless, Lobo et al. (2013a, 2013b) note that their approach can work with other heuristic schedulers.

It should be emphasized that throughout this article, we restrict attention to situations in which the LBSA-delivered allocation $\vartheta^*$ does not satisfy the optimality criteria of Lobo et al. (2013b). In an application for which $\vartheta^*$ satisfies either of the optimality criteria of Lobo et al. (2013b), we know with certainty that $\vartheta^*$ is an optimal worker allocation so that the probabilistic method developed in this article is not needed; but, as explained below, this method will confirm the optimality of $\vartheta^*$ if it is used in such an application.

A “VF-best” allocation $\vartheta^{\text{VF}}$ enables the Virtual Factory to generate a schedule whose corresponding “quasi-optimal” $L_{\text{max}}$ value, denoted by $\text{VF}_{\vartheta^{\text{VF}}}$, is the smallest value of $L_{\text{max}}$ that can be achieved by the Virtual Factory for a given DRC job shop scheduling problem, taken over all feasible worker allocations for that problem. The value $\text{VF}_{\vartheta^{\text{VF}}}$ is termed a “quasi-optimal” $L_{\text{max}}$ value because although all feasible allocations are explored, the Virtual Factory is not guaranteed to generate a schedule yielding the exact minimum value of $L_{\text{max}}$ for each of those worker allocations. By exploiting the generally superior quality of schedules generated by the Virtual Factory, Lobo et al. (2013b) seek to answer the following question:

For a given worker allocation $\vartheta$ in a real instance of the DRC job shop scheduling problem, what is our degree of confidence that $\vartheta$ is a VF-best allocation? In other words, for a simulation replication of the given job shop in which the key characteristics of each job are randomly sampled, what is the probability that the resulting (random) quasi-optimality gap $\text{VF}_{\vartheta^{\text{VF}}} - \text{LB}_{\vartheta^*}$ for the simulation replication is at least as large as the (fixed) value $\text{VF}_{\vartheta^*} - \text{LB}_{\vartheta^*}$ for the real DRC job shop scheduling problem that must actually be solved?

$$\{(1)\}$$

Given a simulation-based estimator $\tilde{F}^{\text{VF}}(x)$ of the cumulative distribution function (c.d.f.) $F^{\text{VF}}(x) = \Pr[\text{VF}_{\vartheta^{\text{VF}}} - \text{LB}_{\vartheta^*} \leq x]$ for all $x$, question (1) can be answered by evaluating $1 - F^{\text{VF}}(\text{VF}_{\vartheta^*} - \text{LB}_{\vartheta^*})$ as an easily interpreted measure of our degree of confidence that the allocation $\vartheta$ is a VF-best allocation for the specific problem at hand.

To estimate the c.d.f. $F^{\text{VF}}(\cdot)$ for a given DRC job shop scheduling problem by means of a simulation experiment, Lobo et al. (2013b) perform the following steps:

(i) a total of $Q = 500$ independent simulation replications of the given DRC job shop scheduling problem are performed; and

(ii) for each simulation replication, the VF-best allocation $\vartheta^{\text{VF}}$ and its associated quasi-optimality gap $\text{VF}_{\vartheta^{\text{VF}}} - \text{LB}_{\vartheta^*}$ are obtained by enumeration of the allocation search space—that is, by invoking the Virtual Factory to generate a schedule for every feasible allocation.

However, because the size of the allocation search space grows exponentially with an increase in either the number of machine groups or the number of machines, the computational complexity of enumeration is prohibitive in large-scale DRC job shop scheduling problems. Experimental results supporting this conclusion can be found in Lobo et al. (2013b).

In this article we formulate and evaluate a revised method for rapidly assessing the quality of a given worker allocation in a specific instance of the DRC job shop scheduling problem. The revised method is based on HSP2, a version of HSP that is designed for large-scale applications. Specifically, in step (ii) of the revised method, HSP2 is used instead of enumeration to yield the allocation $\vartheta^{\text{HSP2}}$ and its associated gap $\text{VF}_{\vartheta^{\text{HSP2}}} - \text{LB}_{\vartheta^*}$, where the latter quantity is taken as an approximation to the quasi-optimality gap $\text{VF}_{\vartheta^{\text{VF}}} - \text{LB}_{\vartheta^*}$. The revised method has proved to be computationally tractable in problems of realistic size and complexity. Through comprehensive experimentation using progressively larger DRC job shop scheduling problems, we compare the empirical reference distributions delivered by enumeration with the corresponding distributions delivered by HSP2, where the comparisons are based on visual inspection as well as the Mann–Whitney test. Except for situations in which the staffing level is so high as to render question (1) moot, we concluded that the difference between the HSP2-based distribution and its enumeration-based counterpart was not significant from either a practical or a statistical perspective. Two numerical examples illustrate the use of the new method for probabilistically evaluating the quality of a given worker allocation in a given instance of the DRC job shop scheduling problem.

The rest of this article is organized as follows. Section 2 provides background on the solution approach of Lobo et al. (2013a, 2013b) for a given instance of the DRC job shop scheduling problem. In Section 3 we present our revised method for evaluating the quality of a given worker allocation. In Section 4 we describe the techniques used to validate the adequacy and reliability of the revised method by visual and statistical comparisons of the distribution of $\text{VF}_{\vartheta^{\text{HSP2}}} - \text{LB}_{\vartheta^*}$ with that of $\text{VF}_{\vartheta^{\text{VF}}} - \text{LB}_{\vartheta^*}$; and in Section 5 we
summarize the experimental results obtained by applying these validation techniques. Section 6 contains two numerical examples illustrating the intended use of the revised method in practical applications. The main findings of this work are summarized in Section 7. The Online Supplement to this article contains the full set of experimental results.

2. Solving the DRC job shop scheduling problem

2.1. Background and motivation

Beginning with Nelson (1967), there has been extensive research activity expended on DRC systems, especially DRC job shops. Treliven (1989), Gargeya and Deane (1996), and Hottenstein and Bowman (1998) survey research on DRC job shop scheduling through the 1990s. Hottenstein and Bowman (1998) find that there are two main questions regarding the allocation of workers:

- When should workers transfer from one machine group to another?
- To which machine group (where) should those workers move?

The authors find that most of the cited research addresses these questions through investigation of worker flexibility, centralization of control, worker allocation rules, queue discipline, and the cost of transferring workers.

Felan et al. (1993) examine the effect that labor flexibility and staffing levels have on key measures of job shop performance. The authors consider a homogeneous group of workers whose flexibility comes from each worker’s ability to work in a varying number of departments within the job shop; moreover, the authors assume that for a given staffing level, each department has the same number of workers assigned to it. Although Felan et al. (1993) allow workers to transfer between departments, the authors are concerned with finding the optimal combination of staffing level and workforce flexibility level, rather than optimizing the allocation of workers to departments given a staffing level. Current work on DRC job shop scheduling has focused on incorporating more realistic assumptions about worker behavior—e.g., learning, fatigue, and forgetfulness (Malhotra et al., 1993; Kher et al., 1999; Kher, 2000; Jaber and Neumann, 2010).

To allocate fully cross-trained workers in a DRC job shop, Lobo et al. (2013a, 2013b) formulate methods that are motivated by the authors’ examination of, and involvement with, the operations in the U.S. Navy’s Aviation Depots, a large-scale apparel producer, and a major high-end furniture manufacturer. In each case, cross-trained workers are allocated on an ad hoc basis; and in each of these disparate organizations, there is a clearly recognized need for a more systematic and effective approach to the allocation of workers. The approach taken in Lobo et al. (2013a, 2013b) and in this article makes extensive use of the Virtual Factory, which was developed by Hodgson et al. (1998) as part of their solution to the $\text{JM}\mid L_{\text{max}}$ single resource constrained job shop scheduling problem. Incorporating a transient, deterministic simulation of a job shop, the Virtual Factory is based on an approach proposed by Vepsäläinen and Morton (1988). The Virtual Factory is an iterative procedure that sequences jobs using a revised slack calculation on the first iteration. On each subsequent iteration, the Virtual Factory sequences the jobs using an iteratively revised slack calculation that includes revised queuing time estimates from the previous iteration. Iterations continue until a previously computed lower bound on $L_{\text{max}}$ is achieved or until the number of iterations reaches a prespecified limit; then the Virtual Factory delivers the best schedule encountered over all iterations.

2.2. Lower bound on $L_{\text{max}}$ and allocation optimality criteria of Lobo et al. (2013a, 2013b)

The DRC job shop scheduling problem considered by Lobo et al. (2013a, 2013b) encompasses $N$ jobs, each of which must be processed on some of the $M$ machines in the job shop, where there are $W$ workers available to operate the machines; and $W < M$, so that the job shop is not fully staffed. Each job has its own route specifying the order in which the job visits its required machines, as well as the processing times required by the job on those machines. Each job also has its own due-date, the time by which all processing must be finished for the job to be considered on time; and thus the job’s lateness $L$ is the job’s completion time minus its due-date. The job shop is divided into $S$ machine groups, and the main problem is to make an allocation $\vartheta$ of the $W$ workers to the $S$ machine groups so as to minimize $L_{\text{max}}$, the maximum lateness taken over all $N$ jobs. Following Pinedo (2012, Section 2.1), we use the notation $\text{JM}\mid W\mid L_{\text{max}}$ to refer to this DRC job scheduling problem.

As mentioned in Section 1, procedure LBSA of Lobo et al. (2013a) delivers a worker allocation $\vartheta^\circ$ and the associated value $L_{\text{BSA}}$ that is a lower bound on the value of $L_{\text{max}}$ over all feasible allocations and over all schedules based on those allocations. To verify that a given allocation is optimal (i.e., it corresponds to a schedule that minimizes $L_{\text{max}}$), Lobo et al. (2013b) establish two theorems providing optimality criteria that are readily checked in practice. Theorem 1 is a general criterion for verifying the optimality of any feasible allocation (including allocation $\vartheta^\circ$), while Theorem 2 is a criterion for verifying the optimality of allocation $\vartheta^\circ$ under certain termination conditions for LBSA.

To provide a meaningful indication of the likelihood that the LBSA-delivered allocation $\vartheta^\circ$ is optimal in practical applications, Lobo et al. (2013a) perform an extensive simulation experiment. The experimental job shop consists of $M = 80$ machines organized into $S = 10$ machine groups of equal size, and there are $N = 1200$ jobs to be completed. Each job has an average of eight operations, where each operation processing time is randomly sampled from the discrete uniform distribution on the integers $\{1, \ldots, 40\}$. The due-date range of the jobs varies from 200 to 3000 in
increments of 400. The staffing level is either 60, 70, 80, or 90%, where a 60% staffing level means that 48 workers are available for allocation. This experimental design yields 64 designated DRC job shop scheduling problems—i.e., 64 different combinations of job shop type, staffing level, and due-date range. For each designated DRC job shop scheduling problem, the authors generate $Q = 200$ simulation replications of the problem using the method given in Appendix A of Lobo et al. (2013a). Each simulation replications of a designated DRC job shop scheduling problem share the following “structural” characteristics:

- the number of jobs;
- the number of machines and the number of machine groups;
- the pattern of symmetric or asymmetric loading of the machine groups;
- the due-date range; and
- the level of staffing—i.e., the ratio of the number of workers to the number of machines expressed as a percentage between 0% and 100%.

However, for each simulation replication, the following characteristics of each job are randomly sampled from the appropriate probability distributions:

- the job’s due-date;
- the number of operations required to complete the job;
- the route of the job through the job shop (i.e., the ordered sequence of machines that the job must visit); and
- the processing time of each operation.

In addition, for each combination of job shop type and due-date range, the $Q = 200$ simulation replications are generated so that for $i = 1, \ldots, Q$, the $i$th randomly generated simulation replication is the same one used with each staffing level. This use of the Monte Carlo method of common random numbers enables sharper comparisons of system performance for different levels of staffing (Chang et al., 1985).

Lobo et al. (2013b) find that depending on the structural characteristics of a given DRC job shop scheduling problem, on a simulation replication of that problem there can be a substantial probability that $\theta^*$ does not satisfy their optimality criteria. This finding led to the development of heuristics to search for an allocation that enables the Virtual Factory to generate a schedule whose $L_{\text{max}}$ value is less than $\text{VF}_{\theta^*}$.

### 2.3. HSP: The heuristic of Lobo et al. (2013b) for solving the DRC job shop scheduling problem

The heuristic HSP of Lobo et al. (2013b) is designed to find a VF-best allocation for the DRC job shop scheduling problem when the LBSA-delivered allocation $\theta^*$ does not satisfy the relevant optimality criteria (i.e., Theorem 1 and Theorem 2 of Lobo et al. (2013b)). Three distinct search heuristics are incorporated into HSP—namely, the Local Neighborhood Search Strategy (LNSS), Queuing Time Search Strategy 1 (QSS1), and Queuing Time Search Strategy 2 (QSS2). LNSS makes use of the structure of the allocation search space to select promising allocations for evaluation. In selecting the next allocation to be evaluated, both QSS1 and QSS2 exploit the Virtual Factory’s estimates of job queuing time at each machine group. LNSS is initialized using allocation $\theta^*$; then QSS1 is initialized using the best allocation encountered by LNSS, where the best allocation is the one that enables the Virtual Factory to generate a feasible schedule with the smallest $L_{\text{max}}$ value encountered thus far. Next QSS2 is initialized using the best allocation encountered by either LNSS or QSS1. The allocation $\theta^{HSP}$ finally returned by HSP is the best allocation encountered by LNSS, QSS1, or QSS2.

The experimental results of Lobo et al. (2013b) indicate that across the 64 designated DRC job shop scheduling problems, the performance of HSP is typically very good. For each designated problem, in a substantial percentage of the simulation replications in which allocation $\theta^*$ does not satisfy the relevant optimality criteria, the HSP-delivered allocation $\theta^{HSP}$ is a VF-best allocation (in this case allocation $\theta^{VFB}$ is obtained via enumeration). However, the authors note that the difference $\text{VF}_{\theta^{HSP}} - \text{VF}_{\theta^*}$ is usually non-zero; therefore, if allocation $\theta^{VFB}$ and consequently the value $\text{VF}_{\theta^{VFB}}$ are both unknown, then we cannot be certain that $\theta^{HSP}$ is a VF-best allocation. To gauge the likelihood that, for a specific instance of the DRC job shop scheduling problem, allocation $\theta^{HSP}$ is in fact a VF-best allocation, Lobo et al. (2013b) develop a probabilistic method for evaluating allocation quality.

### 2.4. The probabilistic method of Lobo et al. (2013b) for evaluating allocation quality

For each DRC job shop scheduling problem used in their experimentation, Lobo et al. (2013b) generate $Q = 500$ simulation replications of that problem as detailed in Section 2.2. Attention is restricted to only the $Q_1$ simulation replications in which allocation $\theta^*$ does not satisfy the relevant optimality criteria (Theorem 1 and Theorem 2 of Lobo et al. (2013b)), where $0 \leq Q_1 \leq Q$. (In situations where $Q_1 = 0$ with high probability, there is of course no need to search for a VF-best allocation; therefore, in the following discussion, we assume that $Q_1 \geq 0$.) For the $i$th simulation replication $(i = 1, \ldots, Q_1)$, the partial enumeration strategy of Lobo et al. (2013b) is used to obtain $\theta^{VFB}$ and $\text{VF}_{\theta^{VFB}}$. Given an arbitrary allocation $\theta$, the quantity $\text{PLB}(\theta) = \text{VF}_{\theta} - \text{LB}_{\theta^*}$ measures the performance of $\theta$ relative to the lower bound on $L_{\text{max}}$. Computing $\text{PLB}(\theta^{VFB})$ for the $Q_1$ replications of interest yields the random sample

$$\Delta_{\text{VFB}} = \{\text{PLB}(\theta^{VFB}) : i = 1, \ldots, Q_1\},$$

which is used to estimate the distribution of the random variable $\text{PLB}(\theta^{VFB})$ taken over all possible realizations (i.e., simulation replications) of the given DRC job shop scheduling problem.

Given a real instance of the DRC job shop scheduling problem that must actually be solved and a specific worker allocation $\theta$ that yields the fixed quantity $\text{PLB}(\theta)$ for the real
problem, Lobo et al. (2013b) seek to test the null hypothesis that $\hat{\theta}$ is a VF-best allocation for the real problem. They test this hypothesis by estimating the probability that for a simulation replication of the given DRC job shop scheduling problem, the random variable $\text{PLB}(\hat{\theta}^{\text{VF}}) = V F_{\hat{\theta}^{\text{VF}}}-L B_{P}$ will be at least as large as the fixed value $\text{PLB}(\hat{\theta})$ for the real problem. This upper-tail probability can be viewed as the $p$-value or level of significance for a test of the null hypothesis (Bickel and Doksum, 2007, pp. 221–223). The upper-tail probability can also be interpreted as the user’s level of confidence that $\hat{\theta}$ is a VF-best allocation for the real problem. For example, if approximately 90% of the observations in the random sample $\Delta_{\text{VF}}^{+}$ are at least as large as $\text{PLB}(\hat{\theta})$ and $Q_{i}^{+}$ is reasonably large, then with roughly 90% confidence we may conclude that $\hat{\theta}$ is a VF-best allocation for the real DRC job shop scheduling problem to be solved.

Lobo et al. (2013b) and Lobo et al. (2014) fit a continuous distribution to the random sample $\Delta_{\text{VF}}^{+}$ in situations for which $\Delta_{\text{VF}}^{+}$ contains a sufficiently large subset of non-zero observations,

$$\Delta_{\text{VF}}^{+} = \{ \text{PLB}(\theta^{\text{VF}}) : \text{PLB}(\theta^{\text{VF}}) > 0 \text{ and } 1 \leq i \leq Q_{i}^{+} \} \text{ with } Q_{i}^{+} = |\Delta_{\text{VF}}^{+}|,$$

so as to enable a meaningful assessment of the adequacy of the fitted distribution. In particular, if $Q_{i}^{+} > \max \{0.9 Q_{i}, 50\}$, then the entire random sample $\Delta_{\text{VF}}^{+}$ is fitted using a continuous distribution. (The best-fitting continuous distributions identified by Lobo et al. (2013b) and Lobo et al. (2014) in various cases include the generalized beta distribution, the shifted gamma distribution, the bounded Johnson distribution, the unbounded Johnson distribution, the shifted Lognormal distribution and the shifted Weibull distribution.) On the other hand if $50 < Q_{i}^{+} \leq 0.9 Q_{i}$, then the authors use $\Delta_{\text{VF}}^{+}$ to fit a mixed distribution of the form

$$F^{\text{VF}}(x) = p_{0} F_{0}(x) + (1-p_{0}) F_{c}(x) \text{ for } -\infty < x < \infty,$$

where: (i) $p_{0} = \text{Pr}\{ \text{PLB}(\theta^{\text{VF}}) = 0 \}$; (ii) $F_{0}(x)$ is the c.d.f. of the degenerate distribution with unit probability mass at the origin so that

$$F_{0}(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases},$$

and (iii) $F_{c}(x)$ is the conditional c.d.f. of $\text{PLB}(\theta^{\text{VF}})$ given that $\text{PLB}(\theta^{\text{VF}}) > 0$, where $F_{c}(x)$ is estimated using the distribution from the above list that provides the best fit to the subsample $\Delta_{\text{VF}}^{+}$. Finally if $Q_{i}^{+} \leq 50$, then $F^{\text{VF}}(x)$ is simply estimated by the empirical c.d.f. of $\Delta_{\text{VF}}^{+}$.

Lobo et al. (2013b) and Lobo et al. (2014) use the simulation-generated random sample $\Delta_{\text{VF}}^{+}$ of size $Q_{i}^{+}$ to compute an estimate $\tilde{F}^{\text{VF}}(x)$ of the underlying c.d.f. $F^{\text{VF}}(x)$. Then for a given real instance of the associated DRC job shop scheduling problem and a specific worker allocation $\hat{\theta}$ that yields $\text{PLB}(\hat{\theta}) > 0$ in the real problem, the authors take $100\{1-\tilde{F}^{\text{VF}}(\text{PLB}(\hat{\theta}))\}$% as their approximate level of confidence that $\hat{\theta}$ is a VF-best allocation for the real problem; and if $\text{PLB}(\hat{\theta}) = 0$, then their confidence level is 100% (i.e., it is known with certainty) that $\hat{\theta}$ is a VF-best allocation for the real problem. Similarly, if $\text{PLB}(\hat{\theta}) > 0$, then $1-\tilde{F}^{\text{VF}}(\text{PLB}(\hat{\theta}))$ is the approximate $p$-value associated with the authors’ simulation-based test of the null hypothesis that $\hat{\theta}$ is a VF-best allocation for the real problem; and if $\text{PLB}(\hat{\theta}) = 0$, then the associated $p$-value is exactly 1.0.

3. A practical method for evaluating worker allocations in large-scale DRC job shops

To evaluate the quality of a given allocation, the method of Lobo et al. (2013b) requires some form of enumeration of the allocation search space on each simulation replication so as to compute the associated “quasi-optimality” gap. Because of its computational complexity, such enumeration is impractical for even a single simulation replication of a large-scale DRC job shop scheduling problem. In this article we develop a revised method for evaluating allocation quality based on the following idea: on the $i$th simulation replication of a given problem, the revised method delivers a readily computed, heuristic-based approximation to the enumeration-based observation $\text{PLB}(\hat{\theta}^{\text{VF}})$ that is too expensive to compute. In other words, we use “simulated” simulation replications of the random variable $\text{PLB}(\hat{\theta}^{\text{VF}})$ to evaluate the quality of a given worker allocation.

The usefulness of the revised method depends critically on a heuristic approximation to $\text{PLB}(\hat{\theta}^{\text{VF}})$ that is sufficiently fast and accurate on each simulation replication $i$ of a large-scale DRC job shop scheduling problem. For this purpose we formulated HSP2, an improved version of HSP whose performance is more stable across a broad range of DRC job shop scheduling problems with varying degrees of size and complexity. In the original HSP algorithm, elapsed execution time is one of the main stopping criteria. Among the search substrategies comprising HSP, the heuristic LNSS has up to 45 seconds for searching different allocations, whereas the heuristics QSS1 and QSS2 together have a total of 15 seconds for this purpose, ensuring that HSP’s total execution time is at most 1 minute per allocation. However, each of these substrategies invokes the Virtual Factory for each allocation searched; and the execution time required by the Virtual Factory to generate a schedule based on a specific allocation increases with the size of the associated job shop. Therefore, the time needed by HSP to search a fixed number of allocations increases with the size of the job shop, and the performance of HSP may vary substantially across DRC job shop scheduling problems of different sizes if the execution time is kept constant.

HSP2 is an improved version of HSP in which the main stopping criterion for each of its constituent search
substrategies is based on the number of allocations searched, not the execution time. In HSP2, the substrategy LNSS is allowed to search up to 10 allocations, while QSS1 and QSS2 are each allowed to search up to four allocations. These limits are based on the average number of allocations that we have found are necessary to ensure adequate performance of HSP2 in DRC job shop scheduling problems of the largest size considered in all our current and previous work.

In the revised method for evaluating allocation quality, attention is focused on the \( Q_i \leq Q \) simulation replications of a given DRC job shop scheduling problem in which allocation \( \hat{\vartheta}^* \) does not satisfy the relevant optimality criteria. Applying HSP2 to the \( i \)th such simulation replication (\( i = 1, \ldots, Q' \)), we obtain the random sample

\[
\Delta_{\text{HSP2}} = \{ \text{PLB}(\hat{\vartheta}_{i,\text{HSP2}}) : i = 1, \ldots, Q' \}. \tag{6}
\]

The c.d.f. \( F_{\text{HSP2}}(x) \) underlying \( \Delta_{\text{HSP2}} \) should be reasonably close to that of \( \Delta_{\text{VFB}} \) because HSP2 is designed to ensure that for \( i = 1, \ldots, Q' \), the HSP2-delivered observation \( \text{PLB}(\hat{\vartheta}_{i,\text{HSP2}}) \) is a reasonably close approximation to the observation \( \text{PLB}(\hat{\vartheta}_{i,\text{VFB}}) \) that would have been generated by jointly using the Virtual Factory and enumeration of the allocation search space on the \( i \)th simulation replication of the given problem.

Next we fit a c.d.f. \( \tilde{F}_{\text{HSP2}}(x) \) to the simulation-generated random sample \( \Delta_{\text{HSP2}} \) along the same lines as in the discussion of Equations (3) to (5) in the previous section. In particular, we fit a continuous distribution to \( \Delta_{\text{HSP2}} \) only if \( \Delta_{\text{HSP2}} \) contains a sufficiently large subset

\[
\Delta_{\text{HSP2}}^+ = \{ \text{PLB}(\hat{\vartheta}_{i,\text{HSP2}}) : \text{PLB}(\hat{\vartheta}_{i,\text{HSP2}}) > 0 \text{ and } 1 \leq i \leq Q' \}
\]

where \( Q'_+ = |\Delta_{\text{HSP2}}| \),

such that

\[
Q' > \max\{0.9Q'_+, 50\}; \tag{7}
\]

and the rest of the discussion of Equations (4) and (5) applies similarly.

Given a real instance of the associated DRC job shop scheduling problem and a specific worker allocation \( \vartheta \) that yields \( \text{PLB}(\vartheta) > 0 \) in the real problem, we take 100 \( \{1 - \tilde{F}_{\text{HSP2}}[\text{PLB}(\vartheta)]\} \% \) as our approximate level of confidence that \( \vartheta \) is a VF-best allocation for the real problem; and if \( \text{PLB}(\vartheta) = 0 \), then our confidence level is 100\% (i.e., we know with certainty) that \( \vartheta \) is a VF-best allocation for the real problem. The main advantage of the revised method for evaluating allocation quality is that it requires orders of magnitude less execution time to yield the simulation-generated observations \( \{ \text{PLB}(\hat{\vartheta}_{i,\text{HSP2}}) : i = 1, \ldots, Q' \} \) on which the estimated c.d.f. \( \tilde{F}_{\text{HSP2}}(x) \) is based. In Section 6 we present two numerical examples that illustrate the application of the revised method to two moderately large DRC job shop scheduling problems.

4. Techniques for validating the revised method

As a first step in gauging the differences and similarities between the two distributions respectively underlying \( \Delta_{\text{VFB}} \) and \( \Delta_{\text{HSP2}} \) for a given DRC job shop scheduling problem, we compute the usual sample statistics for the following paired-sample differences computed from \( \Delta_{\text{VFB}} \) and \( \Delta_{\text{HSP2}} \):

\[
D_i = \text{PLB}(\hat{\vartheta}_{i,\text{HSP2}}) - \text{PLB}(\hat{\vartheta}_{i,\text{VFB}}) = \text{VF}_{\text{HSP2}} - \text{VF}_{\text{VFB}}
\]

for \( i = 1, \ldots, Q' \); \tag{9}

and the paired observations \( \text{PLB}(\hat{\vartheta}_{i,\text{VFB}}) \) and \( \text{PLB}(\hat{\vartheta}_{i,\text{HSP2}}) \) are computed from the same simulation replication \( i \) of the given DRC job shop scheduling problem for \( i = 1, \ldots, Q' \), where \( Q' = Q'_1 = Q'_2 \). In particular, we compute the sample mean and variance of the paired-sample differences,

\[
\bar{D} = \frac{1}{Q'} \sum_{i=1}^{Q'} D_i \quad \text{and} \quad S_D^2 = \frac{1}{Q' - 1} \sum_{i=1}^{Q'} (D_i - \bar{D})^2, \tag{10}
\]

together with paired-sample Student’s \( t \)-statistic

\[
t = \frac{\bar{D}}{S_D / \sqrt{Q'}} \tag{11}
\]

for each designated DRC job shop scheduling problem in the overall simulation experiment. Thus, we can use Equation (11) to test the hypothesis that for a given DRC job shop scheduling problem, \( \text{E}[\bar{D}] = 0 \) so that the underlying distributions from which \( \Delta_{\text{VFB}} \) and \( \Delta_{\text{HSP2}} \) are sampled must have the same mean. Since we know that \( \text{PLB}(\hat{\vartheta}_{i,\text{VFB}}) \leq \text{PLB}(\hat{\vartheta}_{i,\text{HSP2}}) \) when both random variables are computed from the same simulation replication of a given DRC job shop scheduling problem, the appropriate alternative hypothesis based on the paired-sample Student’s \( t \)-statistic (11) is that \( \text{E}[\bar{D}] > 0 \) so that the distribution corresponding to \( \Delta_{\text{HSP2}} \) has a larger mean than that for the distribution corresponding to \( \Delta_{\text{VFB}} \).

The following two techniques are used to test the hypothesis that the random samples \( \Delta_{\text{VFB}} \) and \( \Delta_{\text{HSP2}} \) are both drawn from the same underlying distribution. The first technique is a plot of the empirical c.d.f. of each random sample on the same graph. The degree of overlap between the two empirical c.d.f.s as judged by the reader provides an indication of how likely it is that the two data sets are drawn from the same underlying distribution. For the random sample \( \{ \text{PLB}(\hat{\vartheta}_{i,\text{VFB}}) : i = 1, \ldots, Q'_1 \} \) based on \( Q'_1 \) independent simulation replications of a designated DRC job shop scheduling problem, the corresponding empirical c.d.f. is defined as

\[
\hat{F}_{\text{VFB}}(x) = \frac{1}{Q'_1} \sum_{i=1}^{Q'_1} I\{ \text{PLB}(\hat{\vartheta}_{i,\text{VFB}}) \leq x \} \quad \text{for all real } x, \tag{12}
\]

where for the condition \( C \), the indicator function \( I(C) = 1 \) if condition \( C \) is true, and \( I(C) = 0 \) otherwise. For a random sample \( \{ \text{PLB}(\hat{\vartheta}_{i,\text{HSP2}}) : j = 1, \ldots, Q'_2 \} \), the corresponding empirical c.d.f. \( \hat{F}_{\text{HSP2}}(x) \) is defined similarly. In this graphical validation technique, the samples \( \{ \text{PLB}(\hat{\vartheta}_{i,\text{VFB}}) : i = 1, \ldots, Q'_1 \} \) and \( \{ \text{PLB}(\hat{\vartheta}_{j,\text{HSP2}}) : j = 1, \ldots, Q'_2 \} \) should be overlaid to the extent possible.
... $Q_j$ and $\{PLB(\theta^{HSP2}) : j = 1, \ldots, Q_5\}$ are computed from the same set of $Q$ simulation replications of a given DRC job shop scheduling problem, and we have $Q'_0 = \hat{Q}' = \hat{Q}'_0$.

We also use the Mann–Whitney (or Wilcoxon) nonparametric statistical test (Conover, 1999, pp. 272–281) to test the null hypothesis that $F_VFB(x) = F^{HSP2}(x)$ for all real $x$. As explained in the preceding paragraph, the Mann–Whitney test should be a one-sided test, where the alternative hypothesis is that $PLB(\theta^{HSP2})$ is stochastically larger than $PLB(\theta^{VFB})$, meaning that $F^{HSP2}(x) \leq F^{VFB}(x)$ for all real $x$ (Lehmann, 1975, p. 66). Because the Mann–Whitney test requires that the random samples $\Delta_{VFB}$ and $\Delta_{HSP2}$ are mutually independent, for each designated DRC job shop scheduling problem we used two independent sets of simulation replications to compute the random samples $\Delta_{VFB}$ and $\Delta_{HSP2}$. The Mann–Whitney test was performed using the function scipy.stats.mannwhitneyu in SciPy (Oliphant, 2007).

5. Experimental validation results

An extensive simulation experiment was performed to provide credible evidence that, regardless of the size of DRC job shop considered, the random sample obtained using the new method ($\Delta_{HSP2}$) and the random sample obtained using the original method ($\Delta_{VFB}$) are both drawn from the same underlying distribution. Section 5.1 defines the experimental design and the DRC job shop scheduling problems used, while the results of the extensive simulation experiment are presented in the rest of the section.

5.1. Test problems used in the experimental validation

The DRC job shop used in Lobo et al. (2013a, 2013b) has $N = 1200$ jobs and $M = 80$ machines spread uniformly across $S = 10$ departments; and in our experience, this combination of $M$, $S$, and $N$ can arguably be considered as the specification for a “large” DRC job shop to be used in the experimental validation. Similarly, we selected DRC job shops of “medium” and “small” sizes for the experimental validation based on our experience with such systems; and the defining features of the three job shop sizes are summarized in Table 1. In each selected DRC job shop, there are $M/S$ machines in every department (machine group). In Table 1, the column headings Min., Max., and Same under the column heading No. job operations respectively denote the minimum and maximum number of operations for each job and the maximum number of the job’s operations that can occur in the same machine group. In the DRC job shop of medium size with $M = 64$ machines, we set the number of workers $W$ so as to approximate the desired staffing level percentages as closely as possible.

The definition of a designated DRC job shop scheduling problem must be expanded to include the size of the job shop being considered, so that a designated problem now refers to the combination of the size of the job shop, the type of job shop, the staffing level, and the due-date range.

In the symmetric (or balanced) version of the selected DRC job shops, on average each machine group has $100/5\%$ of the job shop’s workload pass through it. Asymmetry in a selected job shop is created by altering the probability that a machine group is on each job’s route. Thus, if machine group $i$ has probability $0.075$ of being on each job’s route, then on average machine group $i$ will have $7.5\%$ of the job shop’s workload pass through it. Table 2 contains the probabilities of each job being routed through each machine group in the asymmetric (or unbalanced) version of each selected DRC job shop.

In all DRC job shops considered, the due-date of each job is a function of the due-date range, which varies from 200 to 3000 in increments of 400; and the processing time of each operation is randomly sampled from the discrete uniform distribution on the integers $\{1, \ldots, 40\}$.

5.2. Paired-sample Student’s $t$-tests

For the small, medium, and large asymmetric DRC job shop scheduling problems described in Tables 1 and 2, Table 3 summarizes the sample statistics $\{D_i\}$, $\bar{D}$, $S_D$, and $t$ as defined by Equations (9), (10), and (11). The sample statistics for the small, medium, and large symmetric DRC job shop scheduling problems can be found in Table 4. All results are expressed in the units of an average operation processing time, which is equal to 20.5 time units because the operation processing time is uniformly distributed on the integers $\{1, \ldots, 40\}$. Also note that because the paired-sample Student’s $t$-statistic is being computed, the random samples $\Delta_{VFB}$ and $\Delta_{HSP2}$ are not mutually independent: for each designated DRC job shop scheduling problem both $\Delta_{VFB}$ and $\Delta_{HSP2}$ are constructed from the same set of $Q = 500$ simulation replications.

At the $1\%$ level of significance, the results indicated that at the $80\%$ and $90\%$ staffing levels, there were statistically significant differences in the means of the underlying populations from which $\Delta_{VFB}$ and $\Delta_{HSP2}$ were respectively sampled for both symmetric and asymmetric job shops and across all due-date ranges. A similar conclusion applied to the large asymmetric job shop with $70\%$ staffing. On the other

| Job shop size | $M$ | $S$ | $N$ | Min. | Max. | Same | $W$ for each staffing level |
|---------------|-----|-----|-----|------|------|------|---------------------------|
| Small         | 40  | 5   | 1200| 3    | 5    | 2    | 24  28  32  36 |
| Medium        | 64  | 8   | 1200| 5    | 8    | 3    | 38  45  51  58 |
| Large         | 80  | 10  | 1200| 6    | 10   | 3    | 48  56  64  72 |

Table 1. Defining features of the three different size DRC job shops
hand, for the rest of the DRC job shop scheduling problems (asymmetric and symmetric DRC job shop scheduling problems with 60% and 70% staffing, and across all due-date ranges), the results indicate that there were no practically significant differences between the expected values of the two underlying populations from which \( \Delta_{VFB} \) and \( \Delta_{HSP2} \) were respectively sampled.

### 5.3. Graphical and statistical comparisons of distributions underlying \( \Delta_{VFB} \) and \( \Delta_{HSP2} \)

For the empirical c.d.f. plots that follow, for each designated DRC job shop scheduling problem the random samples \( \Delta_{VFB} \) and \( \Delta_{HSP2} \) used are the same as those used for the results in Tables 3 and 4, which is to say that both \( \Delta_{VFB} \) and \( \Delta_{HSP2} \) are constructed from the same set of simulation replications. Thus, \( Q'_1 = Q'_2 = Q' \). The empirical c.d.f. plots corresponding to the 179 designated DRC job shop scheduling problems for which \( Q' > 25 \) can be found in the Online Supplement, while selected representative plots are presented in the following sections. The software used to create the empirical c.d.f. plots draws the empirical c.d.f. of the random sample \( \Delta_{VFB} \) first and the empirical c.d.f. of the random sample \( \Delta_{HSP2} \) second. As a result, the empirical c.d.f. corresponding to \( \Delta_{HSP2} \) is drawn over the empirical c.d.f. corresponding to \( \Delta_{VFB} \) when the two empirical c.d.f.s overlap.

| Due-date range | 60% Staffing | 70% Staffing | 80% Staffing | 90% Staffing |
|----------------|--------------|--------------|--------------|--------------|
| Q' | \( \overline{D} \) | \( S_D \) | t | Q' | \( \overline{D} \) | \( S_D \) | t | Q' | \( \overline{D} \) | \( S_D \) | t | Q' | \( \overline{D} \) | \( S_D \) | t |
| Small 200 | 43 0.000 0.000 — | 142 0.001 0.008 1.004 | 6 0.000 0.000 — | 0 0.000 0.000 — | 200 | 43 0.000 0.000 — | 142 0.001 0.008 1.004 | 6 0.000 0.000 — | 0 0.000 0.000 — |
| Medium 200 | 43 0.000 0.000 — | 123 0.035 0.239 1.620 | 36 0.198 0.644 1.843 | 0 0.000 0.000 — | 360 42 0.000 0.000 — | 140 0.048 0.300 1.882 | 47 0.146 0.582 1.723 | 12 0.123 0.300 1.723 | 479 48 0.000 0.000 — | 181 0.034 0.265 1.712 | 137 0.095 0.264 4.202 | 100 0.284 0.324 8.753 | 871 73 0.000 0.000 — | 236 0.032 0.300 1.617 | 231 0.115 0.219 7.974 | 195 0.179 0.244 10.272 |
| Large 200 | 159 0.000 0.000 — | 219 0.193 0.600 4.770 | 2 0.000 0.000 — | 2 0.000 0.000 — | 600 151 0.000 0.000 — | 258 0.129 0.467 4.418 | 14 0.070 0.157 1.658 | 14 0.038 0.079 1.824 |
| 1000 | 159 0.001 0.008 1.003 | 291 0.230 0.665 5.909 | 35 0.102 0.170 3.550 | 37 0.171 0.233 4.822 | 1800 | 191 0.006 0.063 1.348 | 318 0.314 0.936 5.976 | 74 0.153 0.270 4.871 | 73 0.271 0.345 6.694 |
| 1400 | 176 0.000 0.000 — | 311 0.283 0.744 6.717 | 64 0.133 0.232 4.598 | 66 0.162 0.257 5.111 | 2200 | 207 0.002 0.024 1.135 | 345 0.239 0.661 6.701 | 101 0.123 0.259 4.781 | 97 0.281 0.396 6.975 |
| 2600 | 219 0.000 0.000 — | 361 0.312 0.975 6.081 | 143 0.157 0.240 7.813 | 141 0.280 0.354 9.408 | 3000 | 242 0.000 0.000 — | 377 0.305 0.764 7.752 | 188 0.166 0.211 10.802 | 191 0.215 0.251 11.853 |
The Mann–Whitney test requires that the random samples \( \Delta_{VFB} \) and \( \Delta_{HSP2} \) are mutually independent. Thus, for the results presented in Tables 5 and 6, two independent sets of \( Q = 500 \) simulation replications were used to generate the random samples \( \Delta_{VFB} \) and \( \Delta_{HSP2} \). For each designated DRC job shop scheduling problem, Table 5 and Table 6 contain the following information:

(i) \( Q_1 = |\Delta_{VFB}| \), the number of replications in the first set of \( Q = 500 \) simulation replications in which allocation \( \theta^* \) did not satisfy the optimality criteria (Theorem 1 and Theorem 2 of Lobo et al. (2013b));

(ii) \( Q_2 = |\Delta_{HSP2}| \), the number of replications in the second set of \( Q = 500 \) simulation replications in which allocation \( \theta^* \) did not satisfy the optimality criteria; and

(iii) the \( p \)-value corresponding to the one-sided Mann–Whitney test as performed on the mutually independent random samples \( \Delta_{VFB} \) and \( \Delta_{HSP2} \).

Table 5 presents the results for the asymmetric DRC job shop scheduling problems, and Table 6 presents the results for the symmetric DRC job shop scheduling problems.

5.3.1. Asymmetric job shop

Figures 1 and 2 contain empirical c.d.f.s that are representative of the two different types of plots encountered with asymmetric DRC job shop scheduling problems. Figure 1 displays representative asymmetric DRC job shop scheduling problems in which both the empirical c.d.f.s and the Mann–Whitney \( p \)-values indicate a high level of confidence that both random samples are drawn from the same underlying distribution. Figure 2 displays representative asymmetric DRC job shop scheduling problems in which the empirical c.d.f.s and the Mann–Whitney \( p \)-values do not indicate a high level of confidence that both random samples are drawn from the same underlying distribution.

In asymmetric DRC job shop scheduling problems with lower staffing levels (60% and 70%), there is strong evidence that the new method results in an excellent approximation to the distribution of the quasi-optimality gap. However, in asymmetric DRC job shop scheduling problems with higher staffing levels (80% and 90%), the evidence indicates that the approximation to the distribution of the quasi-optimality gap achieved by the new method is not especially...
convincing. In particular, the approximation to the distribution of the quasi-optimality gap achieved by the new method is less than satisfactory for the following asymmetric DRC job shop scheduling problems:

- small job shop with 90% staffing and due-date range 1800–3000;
- medium job shop with 90% staffing and due-date range 1400–3000;
- large job shop with 80% staffing and due-date range 1400–3000; and
- large job shop with 90% staffing and due-date range 1400–3000.

However, in all of the aforementioned asymmetric DRC job shop scheduling problems, the vast majority of PLB(s, HSP2) values is less than a single average operation processing time (20.5 time units). Thus, the asymmetric DRC job shop scheduling problems in which the new method does not result in an excellent approximation to the distribution of the quasi-optimality gap are also problems that do not necessarily require the approximation to begin with.

5.3.2. Symmetric job shop

Figures 3 and 4 contain empirical c.d.f.s that are representative of the two different types of plots encountered with symmetric DRC job shop scheduling problems of all sizes. Figure 3 displays representative symmetric DRC job shop scheduling problems in which both the empirical c.d.f.s and the Mann–Whitney p-values indicate that there is a high level of confidence that both random samples are drawn from the same underlying distribution. Figure 4 displays representative symmetric DRC job shop scheduling problems in which the empirical c.d.f.s and the Mann–Whitney p-values do not indicate a high level of confidence that both random samples are drawn from the same underlying distribution.

In general, in symmetric DRC job shop scheduling problems there is strong evidence that the new method results in an excellent approximation to the distribution of the quasi-optimality gap. In the symmetric DRC job shop scheduling problems with 60, 70, and 80% staffing, the degree of overlap between empirical c.d.f.s on these plots gives the appearance that there is a single empirical c.d.f. plotted; and this congruence is reflected in the associated Mann–Whitney p-values. In symmetric DRC job shop scheduling problems with a staffing level of 90%, although there is strong agreement in the overall shape of the two empirical c.d.f.s, the discrepancy between the two empirical c.d.f.s is manifested as a downward shift of the empirical c.d.f. corresponding to the random sample \( \Delta_{HSP2} \) relative to the empirical c.d.f.

Table 5. Mann-Whitney p-values between mutually independent random samples \( \Delta_{VFR} \) and \( \Delta_{HSP2} \), asymmetric job shops

| Due-date range | 60% Staffing | 70% Staffing | 80% Staffing | 90% Staffing |
|---------------|--------------|--------------|--------------|--------------|
|               | \( Q_1 \) | \( Q_2 \) | p-Value | \( Q_1 \) | \( Q_2 \) | p-Value | \( Q_1 \) | \( Q_2 \) | p-Value | \( Q_1 \) | \( Q_2 \) | p-Value |
| Small         | 200         | 36           | 0.35       | 6           | 9           | —           | 0           | 3           | —           | 142         | 155         | 0.173       |
| Medium        | 600         | 35           | 0.378      | 15          | 19          | —           | 6           | 12          | —           | 143         | 145         | 0.349       |
| Large         | 1000        | 50           | 0.211      | 50          | 49          | 0.073       | 43          | 45          | 0.003       | 176         | 181         | 0.364       |
|               | 1400        | 45           | 0.476      | 182         | 196         | 0.426       | 90          | 98          | 0.154       | 300         | 300         | 0.091       |
|               | 3000        | 50           | 0.366      | 242         | 222         | 0.491       | 377         | 382         | 0.124       | 141         | 148         | 0.023       |

\( \Delta_{VFR} \) and \( \Delta_{HSP2} \) are also problems that do not necessarily result in an excellent approximation to the distribution of the quasi-optimality gap. In the symmetric DRC job shop scheduling problems with 60, 70, and 80% staffing, the degree of overlap between empirical c.d.f.s on these plots gives the appearance that there is a single empirical c.d.f. plotted; and this congruence is reflected in the associated Mann–Whitney p-values.
corresponding to the random sample $\Delta_{VFB}$. However, in general, this discrepancy is not reflected in the Mann–Whitney test $p$-values.

5.3.3. Overall results

DRC job shop scheduling problems with lower staffing levels and with large optimality gaps represent the problems for which a probabilistic characterization of the optimality gap is necessary and for which question (1) is important. A larger optimality gap translates into a higher level of uncertainty regarding the quality of the allocation, making question (1) more relevant. The results of the large-scale simulation experiment indicate that the revised method results in a distribution that is neither practically nor statistically different from its enumeration-based counterpart. The revised method works well across a wide variety of both asymmetric and symmetric DRC job shop scheduling problems, and works equally well on small problems as it does on medium and large problems. Thus, there is good reason to believe that the new method will work equally well on DRC job shop scheduling problems that are much larger than the largest size problems considered in this article.

6. Examples illustrating the revised method for evaluating allocation quality

The first numerical example involves a large asymmetric job shop with 70% staffing and a due-date range of 200. The revised method was used to generate the random sample $\Delta_{HSP2}$ and fit a distribution to that sample. The resulting fit is the mixed distribution $F_{HSP2}^{\tilde{Q}_i} = 0.301 F_0(x) + 0.699 F_1(x)$, where $F_i(x)$ is the Weibull distribution (1.47, 51.86, 1.00). Figure 5 shows two plots: (i) the fitted probability density function (p.d.f.) superimposed on the histogram formed from the subsample $\Delta_{HSP2}$; and (ii) the fitted c.d.f. $F_{HSP2}^{\tilde{Q}_i}$ superimposed on the empirical c.d.f. $F_{HSP2}^{\tilde{Q}_i}(x)$ formed from the subsample $\Delta_{HSP2}$. In the legend of Fig. 5, the label “Datapoints” refers to $\tilde{Q}_i$, and the $p$-value refers to the chi-squared goodness-of-fit test statistic for the fitted distribution.

For the given DRC job shop scheduling problem, allocation $\theta^*$ failed to satisfy either of the optimality criteria; and we obtained $\text{LB}_{\theta^*} = 3702$ and $\text{VF}_{\theta^*} = 3825$ so that $\text{PLB}(\theta^*) = 123$. From these results we know that

$$F^{VFB}[\text{PLB}(\theta^*)] = \Pr\{\text{PLB}(\theta^{VFB}) \leq 123\}$$
Fig. 1. Selected representative empirical c.d.f. plots, asymmetric job shop.
Fig. 2. Selected representative empirical c.d.f. plots, asymmetric job shop (continued).
Fig. 3. Selected representative empirical c.d.f. plots, symmetric job shop.
Practical method to evaluate worker allocations

Fig. 4. Selected representative empirical c.d.f. plots, symmetric job shop (continued).

\[ \approx \tilde{F}_{HSP2}^{\text{Q}_2}(123) \]
\[ = 0.301 + 0.699 \cdot \left[ 1 - \exp \left\{ - \left( \frac{123 - 1.00}{51.86} \right)^{1.47} \right\} \right] \]
\[ = 0.9792; \]

consequently, our level of confidence is only 2.08% that \( \theta^* \) is a VF-best allocation. On the other hand for the HSP2-generated allocation \( \theta^{HSP2} \), we obtained \( VF_{\theta^{HSP2}} = 3712 \) and \( PLB(\theta^{HSP2}) = 10 \) so that

\[ F_{\text{VF}}[PLB(\theta^{HSP2})] = \Pr\{PLB(\theta^{VF}) \leq 10\} \approx \tilde{F}_{Q_2}^{HSP2}(10) = 0.3523, \]
indicating we are 64.77% confident that allocation $\vartheta^{HSP2}$ is a VF-best allocation.

The second example involves a large symmetric job shop with 80% staffing and a due-date range of 2200. Again the revised method was used to generate the random sample $\Delta_{HSP2}$ and fit a distribution to that sample. The resulting fit is the continuous distribution $\tilde{F}_{HSP2} \sim \text{Johnson-SU} (-0.11, 1.72, 41.51, 86.11)$. Figure 6 shows two plots: (i) the fitted p.d.f. superimposed on the histogram formed from the complete random sample $\Delta_{HSP2}$; and (ii) the fitted c.d.f. $\tilde{F}_{HSP2}(x)$ superimposed on the empirical c.d.f. $\hat{F}_{HSP2}(x)$ formed from $\Delta_{HSP2}$. (The legend in Fig. 6 has the same meaning as for Fig. 5.) For this DRC job shop scheduling problem, allocation $\vartheta^*$ failed to satisfy either of the optimality criteria; and we obtained $\text{LB}_{\vartheta^*} = 1146$, $\text{VF}_{\vartheta^*} = 1337$, and $\text{PLB}(\vartheta^*) = 191$. In this situation we observed that

$$F_{\text{VF}}\{\text{PLB}(\vartheta^*)\} = \Pr\{\text{PLB}(\vartheta^{\text{VF}}) \leq 191\} \approx \tilde{F}_{Q_2}^{-1}(191)$$

$$= \Phi\{-0.11 + 1.72\ln\left(\frac{191 - 86.11}{41.51}\right) + \sqrt{(191 - 86.11)^2 + 1}\}$$

$$= 0.9969,$$

where $\Phi(\cdot)$ denotes the standard normal c.d.f. Therefore, our level of confidence is only 0.31% that $\vartheta^*$ is a
Practical method for evaluate worker allocations

VF-best allocation. The HSP2-generated allocation $\theta^{\text{HSP2}}$ yielded PLB($\theta^{\text{HSP2}}$) = 74 and
\[ F_{\text{VFB}}(\text{PLB}(\theta^{\text{HSP2}})) = \Pr\{\text{PLB}(\theta^{\text{VFB}}) \leq 74\} \approx F_{\text{HSP2}}^{\star}(74) = 0.2726, \]
indicating that we are 72.74% confident that $\theta^{\text{HSP2}}$ is a VF-best allocation.

In both of the foregoing examples, the revised method of Section 3 enabled rapid evaluation of the quality of a given worker allocation. In addition, the revised method allowed the user to quantify the increase in solution quality that is obtained by using one allocation instead of another: in both examples the HSP2-generated allocation $\theta^{\text{HSP2}}$ was substantially better than the baseline allocation $\theta^{\star}$.

7. Conclusions

In this article we formulated and validated a revised method for evaluating the quality of a given worker allocation in a given DRC job shop scheduling problem. Building on the original method of Lobo et al. (2013b), the revised method was designed to be applicable in large-scale problems. When used in conjunction with the Virtual Factory (Hodgson et al., 1998), the revised method has proven to be computationally tractable in problems of realistic size and complexity. Based on the results of our experimental validation of the revised method, we concluded that the method was reliable and effective except in situations for which the staffing level was so high as to make the search for near-optimal worker allocations practically unnecessary.

Supplemental material

Supplemental data for this article can be accessed on the publisher's website.

References

Bickel, P. and Doksum, K. (2007) *Mathematical Statistics: Basic Ideas and Selected Topics*, second edition, Pearson Prentice Hall, Upper Saddle River, NJ.

Chang, Y.-L., Sullivan, R., Bagchi, U. and Wilson, J. (1985) Experimental investigation of real-time scheduling in flexible manufacturing systems. *Annals of Operations Research*, 3(7), 355–377.

Conover, W. J. (1999) *Practical Nonparametric Statistics*, third edition, Wiley, New York, NY.

Felan, J., Fry, T. and Philipoom, P. (1993) Labor flexibility and staffing levels in a dual-resource constrained job shop. *International Journal of Production Research*, 31(10), 2487–2506.

Gargeya, V. and Deane, R. (1996) Scheduling research in multiple resource constrained job shops: a review and critique. *International Journal of Production Research*, 34(8), 2077–2097.

Hodgson, T., Cormier, D., Weintraub, A. and Zozom, A. (1998) Satisfying due dates in large job shops. *Management Science*, 44(10), 1442–1446.

Hodgson, T., King, R., Thoney, K., Stanislaw, N., Weintraub, A. and Zozom, A. (2000) On satisfying due-dates in large job shops: idle time insertion. *IIE Transactions*, 32, 177–180.

Hodgson, T., Melendez, B., Thoney, K. and Trainor, T. (2004) The deployment scheduling analysis tool (DSAT). *Mathematical and Computer Modelling*, 39(6-8), 905–924.

Hottentstein, M. and Bowman, S. (1998) Cross-training and worker flexibility: a review of DRC system research. *The Journal of High Technology Management Research*, 9(2), 157–174.

Jaber, M. and Neumann, W. (2010) Modelling worker fatigue and recovery in dual-resource constrained systems. *Computers & Industrial Engineering*, 59(1), 75–84.

Kher, H. (2000) Examination of flexibility acquisition policies in dual resource constrained job shops with simultaneous worker learning and forgetting effects. *The Journal of the Operational Research Society*, 51(5), 592–601.

Kher, H., Malhotra, M., Philipoom, P. and Fry, T. (1999) Modeling simultaneous worker learning and forgetting in dual resource constrained systems. *European Journal of Operational Research*, 115(1), 158–172.

Lehmann, E.L. (1975) *Nonparametrics: Statistical Methods Based on Ranks*, Holden-Day, Inc., San Francisco, CA.

Lenstra, J. and Rinnooy Kan, A. (1979) Computational complexity of discrete optimization problems, in *Discrete Optimization I: Proceedings of the Advanced Research Institute on Discrete Optimization and Systems Applications of the Systems Science Panel of NATO and of the Discrete Optimization Symposium*, Hammer, P., Johnson, E., and Korte, B. (eds), Elsevier, New York, NY, pp. 121–140.

Lobo, B., Hodgson, T., King, R., Thoney, K. and Wilson, J. (2013a) An effective lower bound on $L_{\text{max}}$ in a worker-constrained job shop, *Computers & Operations Research*, 40(1), 328–343.

Lobo, B., Hodgson, T., King, R., Thoney, K. and Wilson, J. (2013b) Allocating job-shop manpower to minimize $L_{\text{max}}$: optimality criteria, search heuristics, and probabilistic quality metrics. *Computers & Operations Research*, 40(10), 2569–2584.

Lobo, B., Thoney, K., Hodgson, T., King, R. and Wilson, J. (2014) A probabilistic characterization of allocation performance in a worker-constrained job-shop, in *Essays in Planning, Scheduling and Optimization: A Festschrift in Honor of Salah Elmaghraby*, Pulat, S., Sarin, S. and Uszoy, R. (eds), Springer, New York, NY, pp. 301–341.

Malhotra, M., Fry, T., Kher, H. and Donohue, J. (1993) The impact of learning and labor attrition on worker flexibility in dual resource constrained job shops. *Decision Sciences*, 24(3), 641–664.

Nelson, R. (1967) Labor and machine limited production systems. *Management Science*, 13(9), 648–671.

Oliphant, T. (2007) Python for scientific computing. *Computing in Science & Engineering*, 9(3), 10–20.

Pinedo, M. (2012) *Scheduling: Theory, Algorithms, and Systems*, fourth edition, Springer, New York, NY.

Schulz, S., Hodgson, T., King, R. and Thoney, K. (2004) Minimizing $L_{\text{max}}$ for large-scale, job-shop scheduling problems. *International Journal of Production Research*, 42(23), 4893–4907.

Thoney, K., Hodgson, T., King, R. and Wilson, A. (2002) Satisfying due-dates in large multi-factory supply chains. *IIE Transactions*, 34(9), 803–811.

Trelleven, M. (1989) A review of the dual resource constrained system research. *IIE Transactions*, 21(3), 279–287.

Trelleven, M. and Elvers, D. (1985) An investigation of labor assignment rules in a dual-constrained job shop. *Journal of Operations Management*, 6, 51–68.

Vepsäläinen, A. and Morton, T. (1988) Improving local priority rules with global lead-time estimates: a simulation study. *Journal of Manufacturing & Operations Research*, 1, 102–118.

Weintraub, A., Cormier, D., Hodgson, T., King, R., Wilson, J. and Zozom, A. (1999) Scheduling with alternatives: a link between process planning and scheduling. *IIE Transactions*, 31(11), 1093–1102.

Zozom, A., Hodgson, T., King, R., Weintraub, A. and Cormier, D. (2003) Integrated job release and shop-floor scheduling to minimize WIP and meet due-dates. *International Journal of Production Research*, 41(1), 31–45.
Biographies

Benjamin J. Lobo is a Research Associate in the Predictive Technology Laboratory (PTL) in the Systems and Information Engineering Department of the University of Virginia. Prior to joining the PTL he worked for 2 years as an Industrial Engineer for the U.S. Army Research Laboratory at Aberdeen Proving Grounds, MD. He received a Ph.D. in Industrial Engineering and a master’s in Operations Research from North Carolina State University. His research interests include scheduling, production planning, and logistics.

James R. Wilson is a Professor of the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University. His current research interests are focused on probabilistic and statistical issues in the design and analysis of simulation experiments, with special emphasis on applications in healthcare and production. He is a member of ACM, ASA, and SCS; and he is a Fellow of IIE and INFORMS. His Web page is http://www.ise.ncsu.edu/jwilson.

Kristin A. Thoney is an Associate Professor in the Textile and Apparel, Technology and Management Department at North Carolina State University (NCSU). She earned a Ph.D. in Industrial Engineering and Operations Research, as well as an M.S. in Operations Research from NCSU. She also has a B.S. in Mathematics from Valparaiso University. Her research interests include production scheduling, inventory management, logistics, and supply chain modeling.

Thom J. Hodgson is a Distinguished University Professor at North Carolina State University. He received a B.S.E., an M.B.A. and a Ph.D., all from the University of Michigan. His research interests include scheduling, logistics, production/inventory control, manufacturing systems, and applied military OR. He is a Fellow of IIE and INFORMS and a member of the National Academy of Engineering. He is a former Editor-in-Chief (1984–1988) of IIE Transactions. He is a handball player and a former skydiving instructor.

Russell E. King is the Edward P. Fitts Distinguished Professor of Industrial and Systems Engineering at North Carolina State University (NCSU). He is Co-Director of the Center for Additive Manufacturing and Logistics and serves as the Director of Graduate Programs for the Supply Chain Engineering and Management program. He is both Alumni Distinguished Graduate Professor and Alumni Distinguished Undergraduate Professor at NCSU and is a Fellow of the Institute of Industrial Engineers. He received a Ph.D. in Industrial Engineering from the University of Florida. His research interests include analysis, control, and scheduling of production and logistics systems.