A hospital staff scheduling problem under stochastic operation times

Subhash C. Sarin*, Hanif D. Sherali and Seon Ki Kim

Grado Department of Industrial and Systems Engineering,
Virginia Tech,
Blacksburg, VA 24061, USA
Email: sarins@vt.edu
Email: hanifs@vt.edu
Email: kim.seon@gmail.com
*Corresponding author

Abstract: This paper is concerned with a hospital staff scheduling problem wherein surgical operation times are stochastic. We formulate a complete recourse, scenario-based model for this problem, and develop a branch-and-cut (B&C) algorithm that is further hybridised with the Monte Carlo method (MCM). The proposed hybrid MCM-based approach is demonstrated to substantially reduce the required computational effort over a purely branch-and-cut methodology while producing near-optimal solutions having relatively small optimality gaps.

Keywords: hospital staff scheduling problem; Monte Carlo method; MCM; stochastic programming.

Reference to this paper should be made as follows: Sarin, S.C., Sherali, H.D. and Kim, S.K. (2016) 'A hospital staff scheduling problem under stochastic operation times', *Int. J. Planning and Scheduling*, Vol. 2, No. 3, pp.210–232.

Biographical notes: Subhash C. Sarin is the Paul T. Norton Endowed Professor in the Grado Department of Industrial and Systems Engineering at Virginia Polytechnic Institute and State University, Blacksburg, Virginia. His research interests are in the areas of production scheduling, applied mathematical programming, and design and analysis of manufacturing systems. He has published several papers in the industrial engineering and operations research journals and has co-authored three books in the production scheduling area. He is a recipient of several prestigious awards at the university, state, and national levels. He has served on the editorial boards of many journals. He is a Fellow of the Institute of Industrial Engineering and a Full Member of the Institute of Operations Research and the Management Sciences.

Hanif D. Sherali is a University Distinguished Professor Emeritus in the Industrial and Systems Engineering Department at Virginia Polytechnic Institute and State University. His areas of research interest are in mathematical optimisation modelling, analysis, and design of algorithms for specially structured linear, nonlinear, and continuous and discrete non-convex programs, with applications to transportation, location, engineering and network design, production, economics, and energy systems. He has published over 343 refereed articles in various operations research journals and has (co-)

authored nine books, with a total Google Scholar citation count of over 27,025 and an H-index of 62. He is an elected member of the National Academy of Engineering, a fellow of both INFORMS and IIE, and a member of the Virginia Academy of Science Engineering and Medicine.

Seon Ki Kim is a graduate of Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg. He completed his Doctoral degree in 2009. Currently, he is working for United Airlines as a Systems Analyst.

1 Introduction

This paper addresses a hospital staff scheduling problem that is concerned with the scheduling of surgeons to operating rooms so as to maximise revenues. Operating rooms are among the most revenue-generating resources of a hospital, and their proper utilisation plays a critical role in the fiscal viability of the hospital. We assume that the time required by a surgeon to perform an operation is stochastic, and we accordingly designate the underlying problem as a stochastic hospital staff scheduling problem (SHSSP). We formulate a two-stage stochastic mixed-integer programming (MIP) model for this problem and design a branch-and-cut (B&C) method to solve it. Furthermore, we integrate the developed B&C algorithm with a Monte Carlo method (MCM), where the latter affords the use of a fewer number of scenarios while producing near-optimal solutions, and we demonstrate the effectiveness of the proposed hybridised B&C-MCM approach.

Existing studies on hospital scheduling problems have focused on two areas: staff scheduling (Beliën and Demeulemeester, 2006; Beliën, 2007; Berrada et al., 1996; Bosi and Milano, 2001; Burke et al., 2002; Chen and Yeung, 1993; Jaszkiewicz, 1997; Jaumard et al., 1998; Sherali et al., 2002; Sherali and Smith, 2002; Trivedi and Warner, 2001) and operating room scheduling (Batun et al., 2011; Blake and Donald, 2002; Denton et al., 2010; Dexter et al., 1999, 2001; Guinet and Chaabane, 2003; Jebali et al., 2006; Lamiri et al., 2008; Lapierre et al., 1999; Litvak and Long, 2000; Marcon et al., 2003; Ozakarahan, 1989, 1995, 2000; Verweij et al., 2003). In this context, the term, *staff* usually refers to nurses, trainees, residents, and surgeons. These studies have been motivated by the desire to improve the efficiency of a hospital system and reduce costs, which promotes affordable, yet competitive, health care service.

The nurse scheduling problem (Berrada et al., 1996; Burke et al., 2002; Chen and Yeung, 1993; Jaszkiewicz, 1997; Jaumard et al., 1998; Sherali and Smith, 2002) pertains to determining work periods (or shifts) for nurses on a daily basis from among three shifts: day-shift, night-shift, and late-night shift. Alternative objective functions include the minimisation of operating costs, the maximisation of a preference function of nurses, or the minimisation of conflicts between the schedule and nurse preferences. Constraint logic programming methods (Bosi and Milano, 2001), metaheuristics (Burke et al., 2002; Jaszkiewicz, 1997), and expert system approaches (Chen and Yeung, 1993) have been applied to optimise such problems. Venkataraman and Brusco (1996) presented an integrated approach to nurse staffing and scheduling. Beliën (2007) examined the trainee scheduling problem and Beliën and Demeulemeester (2006) developed a branch-and-price solution procedure for this problem. Sherali et al. (2002) studied the resident scheduling problem that is concerned with assigning shifts to residents based on departmental staffing and skill requirements as well as residents' preferences.

The operating room scheduling problem addresses the assignment of a surgical group to operating rooms. The pertinent entities include surgical groups, operating rooms, and the duration for which each surgical group needs an operating room. Blake and Donald (2002) developed an integer programming model to assign each surgical group to an operating room, while penalising the gap between the preferable operating time and the actual assigned time. Santibanez et al. (2005) formulated a model that jointly considers the availabilities of operating rooms and patient beds, as well as the order of patients on the waiting list. Cardeon et al. (2009) employed a branch-and-price-based methodology to generate a comprehensive schedule of surgeons, operation rooms, instruments, and beds, while incorporating medical requirements, number of beds in recovery rooms, and different patterns of surgeries. A number of case studies have also been reported on the operating room scheduling problem (Dexter et al., 1999, 2001; Guinet and Chaabane, 2003; Jebali et al., 2006; Lapierre et al., 1999; Marcon et al., 2003; Ozakarahan, 1989, 1995; Verweij et al., 2003). The goal of operating room scheduling is to guarantee the availability of operating rooms over sufficient durations for each surgical group, and to schedule the operating rooms as efficiently as possible. There are a number of elements of the hospital staff scheduling problem that are stochastic in nature. This includes the time that a surgeon requires to perform an operation, and also, the arrival time of the patients. To offer a high quality of service, it is essential to allow sufficient time, albeit uncertain, for a procedure to be completed satisfactorily. Typically, hospitals have maintained records of time required for each surgical procedure to minimise the gap between the predetermined allotted time and the actual time required to perform the procedure. However, a better approach is to consider stochastic operation times *a priori* while generating a schedule. Belien and Demeulemeester (2007) have further considered two other sources of stochasticity: the number of patients for each operating room and the length of stay for each patient. Their work focused on the cyclic master surgery schedule based on two case studies pertaining to Belgian hospitals. A MIP model was formulated for this problem, which was solved using a simulated annealing approach. Lamiri et al. (2008) considered two types of demand for surgeries: elective surgeries and emergency surgeries. Emergency surgeries were assumed to be stochastic in nature and were required to be performed on the same day as identified. On the other hand, the time required to perform an elective surgery was assumed to be deterministic. The available aggregate operating room capacity (out of an aggregate amount of total available capacity required for emergency surgeries) was assumed to follow a known probability density function. A stochastic mathematical programming model was formulated for scheduling operating rooms, and a combination of the MCM and a MIP approach was adopted for the solution of the problem. Our work differs from that of Lamiri et al. (2008), and also from that presented in the other papers discussed above, in that:

1. we consider the available resources as a set of multiple operating rooms (not as a single aggregate facility)
2. we assume that the operation times are stochastic
3. we use a mathematical programming-based approach that beneficially integrates a B&C algorithm with the MCM.

In a recent article, Denton et al. (2010) have addressed a problem of allocating surgeries to operating rooms in which the time to perform a surgery is assumed to be stochastic. They formulated the problem as a two-stage stochastic program to minimise the cost of utilising the operating rooms plus the expected overtime cost. They also considered the case of minimising the worst possible cost over all realisations of surgery durations within an uncertainty set. In another recent article, Batun et al. (2011) have considered a multi-operating room scheduling problem where the surgery durations are uncertain, but some portions of the surgeries can be performed in parallel (due to the availability of assistants), and the main decisions to be made concern the number of operating rooms to open, the assignment of surgeries to operating rooms, the sequence of surgeries within each operating room, and the times at which surgeons start their first surgery of the day, with the objective of minimising the total expected operating cost. A two-stage stochastic programming approach with recourse was used for solving the problem, where the authors employed valid inequalities to ensure second-stage feasibility of the first-stage solution. The objective function that we use is different from that considered by Denton et al. (2010) and Batun et al. (2011) in that we include revenues, costs, as well as penalties for overlapping operating room and surgeon assignments and for surgeries exceeding expected duration times, and we also demonstrate the effectiveness of a proposed methodology based on combining a B&C procedure with the MCM. Our motivation for including penalty costs is to minimise the overlapping of operations performed by a surgeon and also that of operations assigned to an operating room. Our aim is also to minimise under-assigned operation times for surgeons to perform operations. All of these conflicts are unavoidable in view of stochastic operation times, and so the inclusion of penalty costs helps find solutions that have a lower chance of becoming infeasible for scenarios realised in practice. Shylo et al. (2012) addressed batch operating room scheduling within a block booking system that maximises the expected utilisation of operating room resources subject to a set of probabilistic capacity constraints. As such, their problem is different from that considered in this paper. The reader is referred to the book edited by Denton (2013) for work in health care operations management.

The remainder of this paper is organised as follows. In Section 2, we formulate a two-stage stochastic programming model for the SHSSP. In Section 3, we propose a hybrid B&C and MCM-based algorithm for the solution of this problem, which is theoretically assured to converge to an optimal solution. Related computational results are presented in Section 4. Finally, Section 5 provides a summary with some concluding remarks.

2 Stochastic hospital staff scheduling problem

The SHSSP problem that we address in this paper can be concisely defined as follows: *Given a set of operating rooms and a set of surgeons, the probability distributions for the times taken by surgeons to perform operations, and the operations to be performed by each surgeon, determine the start time of each operation and the operating room in which to perform this operation by the designated surgeon so as to maximise the net profits earned by the hospital*. We use the terms *staff* (or *surgery groups*) and *surgeons* interchangeably, and the same is true for *resources* and *operating rooms*. The hospital earns a profit from each operation that is performed. We assume that the time required to perform an operation by a surgeon is stochastic, and that we have a discretised approximation of the joint probability density function as delineated by a set of scenarios S, where each scenario $s \in S$ occurs with a probability p_s with $\sum_{s \in S} p_s = 1$. Furthermore, penalty costs are introduced to reduce overlapping of operating room and surgeon assignments, overtime services of surgeons, and violations of pre-specified operation times.

2.1 *Model formulation for the SHSSP*

Consider the following notations:

Indices and parameters

$i \in I$	set of surgeons
$j \in J$	set of operations to be performed
$m \in M$	set of operating rooms
$s \in S$	set of scenarios governing operating times
TP	overall planning period
$t = 1,\ldots,TP$	set of time periods
T	threshold that defines *regular time* if $t = 1,..,T$ and *overtime if* $t = T + 1,\ldots,TP$
r_j	profit accruing from operation j, $\forall j \in J$
c_i	extra cost incurred per unit time if surgeon i works overtime, $\forall i \in I$
$J(i)$	set of operations that surgeon i can perform, $\forall i \in I$
$I(j)$	set of surgeons who can perform operation j, $\forall j \in J$
$d_{ijs} \geq 1$	length of operation time required by surgeon i to perform operation j under scenario s, $\forall i \in I, j \in J(i), s \in S$
p_s	probability of occurrence of scenario s, $s \in S$
e^M	penalty per unit for assigning more than one operation at a time in an operating room
e^I	penalty per unit for assigning more than one operation at a time to a surgeon
e^D	penalty per unit for violating pre-specified operation times.

Decision variables

x_{ijmt}	binary variable, which equals 1 if surgeon i is scheduled to start operation j in operating room m during period t, and 0, otherwise, $\forall i \in I, j \in J(i)$, $m \in M$, $t = 1,\ldots, TP$
y_{ijmts}	binary variable, which equals 1 if surgeon i performs operation j in operating room m during period t, under scenario s, and equals 0 otherwise, $\forall i \in I, j \in J(i), m \in M, t = 1,\ldots,TP, s \in S$
w^{M}_{mtsS}	continuous variable representing the excess number of operations (exceeding one) that are assigned to operating room m during period t under scenario s, $\forall m \in M, t = 1,\ldots,TP, s \in S$
w^{I}_{its}	continuous variable representing the excess number of operations (exceeding one) that are assigned to surgeon i during period t under scenario s, $\forall i \in I, t = 1,\ldots,TP, s \in S$
w_{ijs}	continuous slack variable representing the duration of operation time for surgeon i to perform operation j under scenario s, $\forall i \in I, j \in J(i), s \in S$
x	vector of first stage variables, x_{ijmt}, $\forall i \in I, j \in J(i), m \in M, t = 1,\ldots,TP$
$f_s(x)$	recourse function (as defined below) for scenario s, $\forall s \in S$
$\text{Exp}[\tilde{f}(x)]$	expected cost for operations performed during overtime.

We use a two-stage approach for the SHSSP, where the variables are split into first and second stage variables. The principal x_{ijmt}-variables determine the starting time and the operating room for an operation, and are hence designated as the first stage variables. The y_{ijmts}-variables determine the subsequent time intervals over which surgeons perform operations, and together with the other continuous overlap and slack/surplus w-variables, jointly constitute the second stage variables.

2.1.1 SHSSP1

$$\text{Maximise} \quad \sum_{i\in I}\sum_{j\in J(i)}\sum_{m\in M}\sum_{t=1}^{TP} r_j x_{ijmt} - \text{Exp}\left[\tilde{f}(x)\right] \tag{2.1a}$$

$$\text{subject to} \quad \sum_{i\in I(j)}\sum_{m\in M}\sum_{t=1}^{TP} x_{ijmt} \le 1, \qquad \forall j \in J \tag{2.1b}$$

$$x_{ijmt} \in \{0,1\}, \qquad \forall i \in I,\ j \in J(i), m \in M, t = 1,\ldots,TP. \tag{2.1c}$$

The objective function in (2.1a) maximises the net profit minus the expected penalty costs as detailed below. Constraint (2.1b) requires each potential operation to be performed at most once, and constraint (2.1c) enforces binary restrictions on the first stage x_{ijmt} variables. The second stage recourse model for each scenario s, which evaluates the recourse function $f_s(x)$, $\forall s \in S$, is given as follows:

Recourse model for SHSSP1

$$f_s(x) \equiv \text{Minimise} \sum_{i \in I} \sum_{j \in J(i)} \left[\sum_{m \in M} \sum_{t=T+1}^{TP} \left(c_i y_{ijmts} \right) \right] + \sum_{t=1}^{TP} \left(e^M \sum_{m \in M} w_{mts}^M + e^I \sum_{i \in I} w_{its}^I \right) + e^D \sum_{i \in I} \sum_{j \in J(i)} w_{ijs} \quad (2.2a)$$

subject to $$\sum_{i \in I} \sum_{j \in J(i)} y_{ijmts} - w_{mts}^M \leq 1, \quad \forall m \in M, t = 1, \ldots, TP \quad (2.2b)$$

$$\sum_{j \in J(i)} \sum_{m \in M} y_{ijmts} - w_{its}^I \leq 1, \quad \forall i \in I, t = 1, \ldots, TP \quad (2.2c)$$

$$y_{ijm1s} = x_{ijm1}, \quad \forall i \in I, j \in J(i), m \in M \quad (2.2d)$$

$$y_{ijmts} - y_{ijm(t-1)s} \leq x_{ijmt}, \quad \forall i \in I, j \in J(i), m \in M, t = 2, \ldots, TP \quad (2.2e)$$

$$\sum_{m \in M} \sum_{t=1}^{TP} y_{ijmts} + w_{ijs} \geq d_{ijs} \sum_{m \in M} \sum_{t=1}^{TP} x_{ijmt}, \quad \forall i \in I, j \in J(i) \quad (2.2f)$$

$$y_{ijmts} \in \{0, 1\}, \quad \forall i \in I, j \in J(i), m \in M, t = 1, \ldots, TP \quad (2.2g)$$

$$w_{mts}^M \geq 0, \quad \forall m \in M, t = 1, \ldots, TP \quad (2.2h)$$

$$w_{its}^I \geq 0, \quad \forall m \in M, t = 1, \ldots, TP \quad (2.2i)$$

$$w_{ijs} \geq 0, \quad \forall i \in I, j \in J(i). \quad (2.2j)$$

The objective function (2.2a) records the total overtime surgery cost plus the respective penalty costs for overlapping operating room usage, overlapping surgeon usage, and under-assigned surgery times. Constraints (2.2b), (2.2c), (2.2h), (2.2i), along with the objective function, assert that

$$w_{mts}^M = \max\left\{0, \sum_{i \in I} \sum_{j \in J(i)} y_{ijmts} - 1\right\}, \forall m \in M, t = 1, \ldots, TP, s \in S \text{ and}$$
$$w_{its}^I = \max\left\{0, \sum_{j \in J(i)} \sum_{m \in M} y_{ijmts} - 1\right\}, \quad (2.3a)$$

respectively. In other words, w_{mts}^M and w_{its}^I become positive once multiple operations are assigned at a time to an operating room and a surgeon, respectively, for a given scenario. Also, constraints (2.2f) and (2.2j) along with the objective function yield

$$w_{ijs} = \max\left\{0, d_{ijs} \sum_{m \in M} \sum_{t=1}^{TP} x_{ijmt} - \sum_{m \in M} \sum_{t=1}^{TP} y_{ijmts}\right\}, \quad \forall i \in I, j \in J(i), s \in S. \quad (2.3b)$$

When $\sum_{m\in M}\sum_{t=1}^{TP} y_{ijmts} < d_{ijs}\sum_{m\in M}\sum_{t=1}^{TP} x_{ijmt}$, i.e., the operation time assigned to a surgeon i to perform operation j under scenario s is less than d_{ijs}, given that this surgeon-operation assignment is made, then w_{ijs} becomes positive. Note that for relatively large values of e^M and e^I, the w_{mts}^M – and w_{its}^I – variables will take a value of zero, which would preclude any overlaps among assignments of surgeons and operating rooms. However, for more representative penalty cost parameters, even though some w_{mts}^M – and w_{its}^I – variables might be positive, thereby implying overlaps in the assignments of surgeons and operating rooms, this might increase the net profit for the hospital. In the same vein, profits can potentially increase by permitting the scheduling of additional surgeries when $w_{ijs} > 0$, albeit at a penalty cost. Constraint (2.2d) requires surgeon i to be engaged in performing operation j in room m during the first period if and only if $x_{ijm1} = 1$. For subsequent periods $t = 2,\ldots,TP$, and for each $i \in I, j \in J(i)$, and $m \in M$, Constraint (2.2e) assures that if $y_{ijm(t-1)s} = 0$ and $y_{ijmts} = 1$, then it must be that $x_{ijmt} = 1$. Hence, in light of (2.1b), the stream of y_{ijmts}-variables over $t = 1,\ldots,TP$ can only follow the pattern of having at most a single string of consecutive ones, with the remaining variables being zeros.

Accordingly, $\text{Exp}[\tilde{f}(x)]$ can be expressed as follows:

$$Exp\left[\tilde{f}(x)\right] = \sum_{s\in S} p_s f_s(x) = \left\{\begin{array}{l} \sum_{i\in I}\sum_{j\in J(i)}\left[\sum_{m\in M}\sum_{t=T+1}^{TP}\left(c_i\sum_{s\in S} p_s y_{ijmts}\right)\right] \\ +\sum_{s\in S} p_s\left[\begin{array}{l}\sum_{t=1}^{TP}\left(e^M\sum_{m\in M} w_{mts}^M + e^I\sum_{i\in I} w_{its}^I\right) \\ +e^D\sum_{i\in I}\sum_{j\in J(i)} w_{ijs}\end{array}\right]\end{array}\right., \tag{2.4}$$

where the y- and w-variables represent optimal values to the recourse model given by (2.2a)–(2.2j), for each $s \in S$. Correspondingly, the overall deterministic equivalent of the two-stage stochastic program is given as follows:

2.1.2 SHSSP2

$$\begin{aligned}\text{Maximise}\quad & \sum_{i\in I}\sum_{j\in J(i)}\left(\sum_{m\in M}\sum_{t=1}^{TP} r_j x_{ijmt}\right) - \sum_{i\in I}\sum_{j\in J(i)}\left[\sum_{m\in M}\sum_{t=T+1}^{TP}\left(c_i\sum_{s\in S} p_s y_{ijmts}\right)\right] \\ & -\sum_{s\in S} p_s\left[\sum_{t=1}^{TP}\left(e^M\sum_{m\in M} w_{mts}^M + e^I\sum_{i\in I} w_{its}^I\right) + e^D\sum_{i\in I}\sum_{j\in J(i)} w_{ijs}\right]\end{aligned} \tag{2.5a}$$

$$\text{subject to}\quad \sum_{i\in I(j)}\sum_{m\in M}\sum_{t=1}^{TP} x_{ijmt} \le 1, \qquad \forall j\in J \tag{2.5b}$$

$$\sum_{i\in I}\sum_{j\in J(i)} y_{ijmts} - w_{mts}^M \le 1, \qquad \forall m\in M, t = 1,\ldots,TP, s\in S \tag{2.5c}$$

$$\sum_{j\in J(i)}\sum_{m\in M} y_{ijmts} - w^{I}_{its} \le 1, \qquad \forall i\in I, t=1,\ldots,TP, s\in S \tag{2.5d}$$

$$y_{ijm1s} = x_{ijm1}, \qquad \forall i\in I, j\in J(i), m\in M, s\in S \tag{2.5e}$$

$$y_{ijmts} - y_{ijm(t-1)s} \le x_{ijmt}, \qquad \forall i\in I, j\in J(i), m\in M, t=2,\ldots,TP, s\in S \tag{2.5f}$$

$$\sum_{m\in M}\sum_{t=1}^{TP} y_{ijmts} + w_{ijs} \ge d_{ijs}\sum_{m\in M}\sum_{t=1}^{TP} x_{ijmt}, \qquad \forall i\in I, j\in J(i), s\in S \tag{2.5g}$$

$$x_{ijmt}\in\{0,1\}, \qquad \forall i\in I, j\in J(i), m\in M, t=1,\ldots,TP \tag{2.5h}$$

$$y_{ijmts}\in\{0,1\}, \qquad \forall i\in I, j\in J(i), m\in M, t=1,\ldots,TP, s\in S \tag{2.5i}$$

$$w^{M}_{mts} \ge 0, \qquad \forall m\in M, t=1,\ldots,TP, s\in S \tag{2.5j}$$

$$w^{I}_{its} \ge 0, \qquad \forall i\in I, t=1,\ldots,TP, s\in S \tag{2.5k}$$

$$w_{ijs} \ge 0, \qquad \forall i\in I, j\in J(i), s\in S. \tag{2.5l}$$

The formulated two-stage stochastic program is a *complete recourse model* in the sense that for any binary x-vector, there exists a feasible solution to (2.2b)–(2.2j). This is established below.

Proposition 1: The formulated two-stage stochastic program is a complete recourse model.

Proof: Let x be any binary vector. For each $s \in S$, by setting $y_{ijmts} = x_{ijmt}$, $\forall i \in I, j \in J(i)$, $m \in M$, and $t = 1,\ldots,TP$, and then computing the w-variables accordingly via equations (2.3a) and (2.3b), we obtain a feasible solution to (2.2b)–(2.2j). Hence, the two-stage stochastic program SHSSP1 (or SHSSP2) is a complete recourse model. □

2.2 Performance enhancing constraints

2.2.1 Symmetry breaking constraints

Note that the formulation SHSSP2 can be improved by inhibiting the occurrence of symmetric solutions [which can otherwise severely impair the performance of branch-and-bound/cut approaches due to the redundant enumeration of symmetric reflections of essentially identical solutions – see Sherali and Smith (2002), for example]. Problem SHSSP2 has an inherent symmetry with respect to rooms $m \in M$, which can be alleviated by imposing the following set of hierarchical constraints:

$$\sum_{i\in I}\sum_{j\in J(i)}\sum_{t=1}^{TP} r_j x_{ijmt} \ge \sum_{i\in I}\sum_{j\in J(i)}\sum_{t=1}^{TP} r_j x_{ij(m+1)t}, \quad \forall m\in 1,\ldots,|M|-1 \tag{2.5m}$$

Observe that constraint (2.5m) requires a non-increasing hierarchy of profits generated via the rooms labelled as $m = 1,\ldots,|M|$. We assume henceforth that SHSSP1 and SHSSP2 additionally accommodate constraint (2.5m).

Similarly, symmetry can be broken with respect to operating hours. For each $(i, j) \in I \times J$, let $\delta \equiv (i - 1)|J| + j$. Hence, $\delta = 1, 2,\ldots,|I||J|$ continuously indexes the combinations (i, j) for $i = 1, 2,\ldots,|I|$, $j = 1, 2,\ldots,|J|$. Let Δ denote the set of (δ_1, δ_2) such that $\delta_1 < \delta_2$ and both δ_1 and δ_2 represent viable assignments of surgeons to operations and let d_{ij}^{min} denote the minimum dedicated time (without overlaps) to be allotted for surgeon i to perform operation j. Then, we can assert that for each room $m \in M$, if both δ_1 and δ_2 are performed in room m for $(\delta_1, \delta_2) \in \Delta$, then δ_1 must precede δ_2. This can be stated as follows:

$$\sum_{t=1}^{TP} t x_{\delta_1 mt} + d_{\delta_1}^{min} - \sum_{t=1}^{TP} t x_{\delta_2 mt} \le TP\left[2 - \sum_{t=1}^{TP} x_{\delta_1 mt} - \sum_{t=1}^{TP} x_{\delta_2 mt}\right]$$

$$\text{i.e., } \sum_{t=1}^{TP} (t + TP) x_{\delta_1 mt} - \sum_{t=1}^{TP} (t - TP) x_{\delta_2 mt} \le 2TP - d_{\delta_1}^{min} \tag{2.5n}$$

for all $m \in M, (\delta_1, \delta_2) \in \Delta$.

2.2.2 Zeroing-out redundant variables

Given the restriction on d_{ij}^{min} and noting constraint (2.5n), we can additionally assert the following:

$$x_{ijmt} \equiv 0, \ \forall i \in I, \ \forall j \in J(i), \ m \in M, \ t = TP - d_{ij}^{min} + 2,\ldots,TP \tag{2.5o}$$

2.2.3 Tightening constraints

Likewise, the (recourse) model can be tightened by incorporating the following constraint:

$$w_{ijs} \le \left(d_{ijs} - d_{ij}^{\min}\right) \sum_{m \in M} \sum_{t=1}^{TP} x_{ijmt} \qquad \forall i \in I, \forall j \in J(i), \forall s \in S \tag{2.5p}$$

3 An integrated B&C and MCM-based approach

In this section, we propose a MCM-based approach for solving SHSSP2, where the sub-problems are solved using a standard B&C methodology, as for example, that implemented in the commercial software CPLEX. The MCM is known for its effectiveness in solving stochastic programming problems because it employs only a subset of scenarios. To design this approach, let K and N denote the number of replications and the number of independent random scenarios in each replication, respectively. We can express the k^{th} replication of SHSSP2 as follows:

3.1 k^{th} replication of SHSSP2

$$\text{Maximise} \sum_{i\in I}\sum_{j\in J(i)}\left(\sum_{m\in M}\sum_{t=1}^{TP} r_j x_{ijmt}\right) - \frac{1}{N}\sum_{i\in I}\sum_{j\in J(i)}\left[\sum_{m\in M}\sum_{t=T+1}^{TP}\left(c_i \sum_{s\in S^k} y_{ijmts}\right)\right]$$
$$-\frac{1}{N}\sum_{s\in S^k}\left[\sum_{t=1}^{TP}\left(e^M \sum_{m\in M} w_{mts}^M + e^I \sum_{i\in I} w_{its}^I\right) + e^D \sum_{i\in I}\sum_{j\in J(i)} w_{ijs}\right] \quad (3.1a)$$

$$\text{subject to} \quad \sum_{i\in I(j)}\sum_{m\in M}\sum_{t=1}^{TP} x_{ijmt} \le 1, \qquad \forall j\in J \quad (3.1b)$$

$$\sum_{i\in I}\sum_{j\in J(i)} y_{ijmts} - w_{mts}^M \le 1, \qquad \forall m\in M, t=1,\ldots,TP, s\in S^k \quad (3.1c)$$

$$\sum_{j\in J(i)}\sum_{m\in M} y_{ijmts} - w_{its}^I \le 1, \qquad \forall i\in I, t=1,\ldots,TP, s\in S^k \quad (3.1d)$$

$$y_{ijm1s} = x_{ijm1}, \qquad \forall i\in I, j\in J(i), m\in M, s\in S^k \quad (3.1e)$$

$$y_{ijmts} - y_{ijm(t-1)s} \le x_{ijmt}, \qquad \forall i\in I, j\in J(i), m\in M, t=2,\ldots,TP, s\in S^k \quad (3.1f)$$

$$\sum_{m\in M}\sum_{t=1}^{TP} y_{ijmts} + w_{ijs} \ge d_{ijs}\sum_{m\in M}\sum_{t=1}^{TP} x_{ijmt}, \qquad \forall i\in I, j\in J(i), s\in S^k \quad (3.1g)$$

$$\sum_{i\in I}\sum_{j\in J(i)}\sum_{t=1}^{TP} r_j x_{ijmt} \ge \sum_{i\in I}\sum_{j\in J(i)}\sum_{t=1}^{TP} r_j x_{ij(m+1)t}, \quad \forall m\in 1,\ldots,|M|-1 \quad (3.1h)$$

$$\sum_{t=1}^{TP}(t+TP)x_{\delta_1 mt} - \sum_{t=1}^{TP}(t-TP)x_{\delta_2 mt} \le 2TP - d_{\delta_1}^{min}$$
$$\forall(\delta_1,\delta_2)\in\Delta, m\in M \quad (3.1i)$$

$$x_{ijmt} \equiv 0, \qquad \forall i\in I, \forall j\in J(i), m\in M, t = TP - d_{ij}^{min}+2,\ldots,TP \quad (3.1j)$$

$$w_{ijs} \le \left(d_{ijs} - d_{ij}^{min}\right)\sum_{m\in M}\sum_{t=1}^{TP} x_{ijmt}, \qquad \forall i\in I, \forall j\in J(i), \forall s\in S^k \quad (3.1k)$$

$$x_{ijmt}\in\{0,1\}, \qquad \forall i\in I, j\in J(i), m\in M, t=1,\ldots,TP \quad (3.1l)$$

$$y_{ijmts}\in\{0,1\}, \qquad \forall i\in I, j\in J(i), m\in M, t=1,\ldots,TP, s\in S^k \quad (3.1m)$$

$$w_{mts}^M \ge 0, \qquad \forall m\in M, t=1,\ldots,TP, s\in S^k \quad (3.1n)$$

$$w_{its}^I \ge 0, \qquad \forall i\in I, t=1,\ldots,TP, s\in S^k \quad (3.1o)$$

$$w_{ijs} \geq 0, \qquad \forall i \in I, j \in J(i), s \in S^k \qquad (3.1p)$$

where $S^k \equiv \{s_1^k, s_2^k, \ldots, s_N^k\}$ is the set of N scenarios that are randomly generated for the k^{th} replication according to the specified discrete probability distribution. Note that constraints (3.1g)–(3.1k) correspond to the performance-enhancing constraints (2.5m)–(2.5p) in the previous section.

Proposition 2: The optimal solution obtained for the formulation (3.1) designated as the k^{th} replication of SHSSP2 can be augmented to derive a complete feasible solution to the original formulation SHSSP2.

Proof: By fixing the x-variables according to the solution obtained by solving the k^{th} replication of SHSSP2, the resultant problem SHSSP2 decomposes by scenarios, each of which is feasible by Proposition 1, and can be optimised over the remaining scenarios $s \in S \setminus S^k$, noting that the model directly determines the corresponding optimal solutions for $s \in S^k$. □

Now, consider the following notation for implementing the proposed MCM-based solution approach for the SHSSP':

$\hat{X}_N^k = \{(\hat{x}_{ijmt})_N^k\}$	vector of x_{ijmt}-variables that represents an optimal solution to the k^{th} replication of SHSSP2
$\hat{Y}_N^k = \{(\hat{y}_{ijmts})_N^k\}$	vector of y_{ijmts}-variables that represents an optimal solution to the k^{th} replication of SHSSP2
$(\hat{W}^I)_N^k = \{(\hat{w}_{its}^I)_N^k\}$	vector of w_{its}^I-variables that represents an optimal solution to the k^{th} replication of SHSSP2
$(\hat{W}^M)_N^k = \{(\hat{w}_{mts}^M)_N^k\}$	vector of w_{mts}^M-variables that represents an optimal solution to the k^{th} replication of SHSSP2
$(\hat{W})_N^k = \{(\hat{w}_{ijs})_N^k\}$	vector of w_{ijs}-variables that represents an optimal solution to the k^{th} replication of SHSSP2
X^*	vector of x_{ijmt}-variables that represents an optimal solution to SHSSP2
Y^*	vector of y_{ijmts}-variables that represents an optimal solution to SHSSP2
W^{I*}	vector of w_{its}^I-variables that represents an optimal solution to SHSSP2
W^{M*}	vector of w_{mts}^M-variables that represents an optimal solution to SHSSP2
W^*	vector of w_{ijs}-variables that represents an optimal solution to SHSSP2

$h_N^k(\cdot)$	objective function for the k^{th} replication of SHSSP2
$h(\cdot)$	objective function for SHSSP2
$h_N^k(\hat{X}_N^k)$	optimal objective function value for the k^{th} replication of SHSSP2
$h(\hat{X}_N^k)$	'true' objective function value corresponding to $\hat{X}_N^k$ for SHSSP2 (which could, in theory, be derived by determining an optimal completion to $\hat{X}_N^k$ as alluded to in the proof of Proposition 2)
$h(X^*)$	'true' (unknown) optimal objective function value for SHSSP2
$\overline{h}_N$	sample average (mean) of $h_N^1(\hat{X}_N^1), h_N^2(\hat{X}_N^2), \cdots, h_N^K(\hat{X}_N^K)$
$V(\overline{h}_N)$	sample variance of $h_N^1(\hat{X}_N^1), h_N^2(\hat{X}_N^2), \cdots, h_N^K(\hat{X}_N^K)$
$\tilde{N}$	number of scenarios used to determine an unbiased estimate of the true objective function value once a solution is found using N scenarios, where $\tilde{N} >> N$
$h_{\tilde{N}}^k(\hat{X}_N^k)$	unbiased estimate of $h_{\tilde{N}}^k(\hat{X}_N^k)$, given by the sample average of $h_{\tilde{n}}^k(\hat{X}_N^k)$, $\tilde{n} = 1, \ldots, \tilde{N}$, where $h_{\tilde{n}}^k(\hat{X}_N^k)$ is the value of $h(\hat{X}_N^k)$ obtained via SHSSP2 by fixing $x \equiv \hat{X}_N^k$ and using scenario $\tilde{n}$ by itself
$V(h_{\tilde{N}}^k(\hat{X}_N^k))$	sample variance of $h_{\tilde{n}}^k(\hat{X}_N^k)$, $\tilde{n} = 1, \ldots, \tilde{N}$, as given by $\frac{1}{\tilde{N}(\tilde{N}-1)}[\sum_{\tilde{n}=1}^{\tilde{N}}(h_{\tilde{n}}^k(\hat{X}_N^k) - h_{\tilde{N}}^k(\hat{X}_N^k))^2]$
$(\hat{X}_N)^*$	estimate for a 'true' optimal solution that is obtained from the set of optimal solutions to each replication as determined by: $(\hat{X}_N)^* \in \underset{\hat{X}_N^k,\, k=1,\ldots,K}{\arg\max} \{h_{\tilde{N}}^k(\hat{X}_N^k)\}$
$h_{\tilde{N}}((\hat{X}_N)^*)$	estimate for the 'true' optimal solution value as determined by $(\hat{X}_N)^*$, where $h_{\tilde{N}}((\hat{X}_N)^*) \equiv \max\{h_{\tilde{N}}^k(\hat{X}_N^k) : k = 1, \ldots, K\}$
$V[h_{\tilde{N}}((\hat{X}_N)^*)]$	$V(h_{\tilde{N}}^k(\hat{X}_N^k))$ corresponding to $k \in \{1, \ldots, K\}$ for which $(\hat{X}_N)^* = (\hat{X}_N^k)$

The following results establish the relevant properties of the MCM approach, which we state without proofs since identical results are proven in Mak et al. (1999), Shapiro (2001, 2003), and Verweij et al. (2003).

1 Because SHSSP2 is a maximisation problem, the optimal objective function value of the k^{th} replication of SHSSP2, $h_N^k(\hat{X}_N^k)$, provides an unbiased estimate of an upper bound for the "true" optimal objective function value of SHSSP2, $h(X^*)$.

2 As $N \to \infty$, $E[h_N^k(\hat{X}_N^k)]$ approaches $h(X^*)$ from above. Moreover, $h_N^k(\hat{X}_N^k) \to h(X^*)$, $\bar{h}_N \to h(X^*)$, and $h((\hat{X}_N)^*) \to h(X^*)$. Furthermore, $(\hat{X}_N)^* \to X^*$ if X^* is unique.

3 As $K \to \infty$, $\bar{h}_N$ provides an unbiased estimate of an upper bound for the 'true' optimal objective function value of SHSSP2.

4 As $\tilde{N} \to \infty$, $h_{\tilde{N}}^k(\hat{X}_N^k)$ converges to $h(\hat{X}_N^k)$, the objective function value of $\hat{X}_N^k$ in SHSSP2, from above with probability one, and so, when $N \to \infty$ with $\tilde{N} > N$, we get $h_{\tilde{N}}^k(\hat{X}_N^k) \to h(X^*)$.

5 By (3) and (4), $\bar{h}_N - h_{\tilde{N}}((\hat{X}_N)^*) \to 0$ is an unbiased estimate of the gap between upper and lower bounds. Furthermore, by (2), $\bar{h}_N - h_{\tilde{N}}((\hat{X}_N)^*) \to 0$ as $N \to \infty$.

An unbiased estimate of the variance of the optimality gap is given by

$$V\left[\bar{h}_N - h_{\tilde{N}}\left((\hat{X}_N)^*\right)\right] = V[\bar{h}_N] + V\left[h_{\tilde{N}}\left((\hat{X}_N)^*\right)\right]. \tag{3.2}$$

The 100(1 – α) % confidence interval on the optimality gap is therefore given by the t-distribution with (K-1) degree of freedom as follows:

$$\left[0, \max\left\{0, \bar{h}_N - h_{\tilde{N}}\left((\hat{X}_N)^*\right)\right\} + t_{1-\alpha,K-1}\sqrt{V[\bar{h}_N] + V\left[h_{\tilde{N}}\left((\hat{X}_N)^*\right)\right]}\Big/\sqrt{K}\right] \tag{3.3}$$

From above, an unbiased estimator of the relative optimality gap is given by,

$$\frac{\bar{h}_N - h_{\tilde{N}}\left((\hat{X}_N)^*\right)}{\bar{h}_N} = 1 - \frac{h_{\tilde{N}}\left((\hat{X}_N)^*\right)}{\bar{h}_N}. \tag{3.4}$$

We refer to the expression in (3.4) as the *tolerance ratio*. The exterior sampling method suggested by Shapiro (2001, 2003) and Verweij et al. (2003) is adopted to build a subset of scenarios for each replication in the MCM approach as specified in the following stepwise procedure, where for each $k = 1,\ldots,K$, the k^{th} replication of SHSSP2 is solved via a B&C algorithm as provided in CPLEX using the default cut generation option.

Step 1 Specify values for N, $\tilde{N}$, and K. Set $k = 1$.

Step 2 Generate $S^k = \{s_1^k, s_2^k, \ldots, s_N^k\}$ using a random number generator.

Step 3 Formulate the k^{th} replication of SHSSP2 using S^k, and solve it. Store the resultant solution $\hat{X}_N^k$ and its objective value $h_N^k(\hat{X}_N^k)$.

Step 4 Generate $\tilde{S}^k = \{\tilde{s}_1^k, \tilde{s}_2^k, \ldots, \tilde{s}_{\tilde{N}}^k\}$, and compute $h_{\tilde{N}}^k(\hat{X}_N^k)$ and $V(h_{\tilde{N}}^k(\hat{X}_N^k))$.

Step 5 If $k = K$, go to step 6; otherwise, set $k \leftarrow k + 1$, and go to step 2.

Step 6 Compute $\bar{h}_N$ and $V(\bar{h}_N)$.

Step 7 Determine $(\hat{X}_N)^*$. If $0 \le 1 - \dfrac{h_{\tilde{N}}((\hat{X}_N)^*)}{\bar{h}_N} \le \varepsilon$, go to step 8; otherwise, increase the values of N, $\tilde{N}$, and K, and go to step 2 with $k = 1$. Note that, we reject the results if $h_{\tilde{N}}((\hat{X}_N)^*) - \bar{h}_N > 0$, that is, we restart the process with increased values of N, $\tilde{N}$, and K if we end up with an unbiased estimate of the lower bound of the 'true' optimal objective function value that is greater than that of the upper bound.

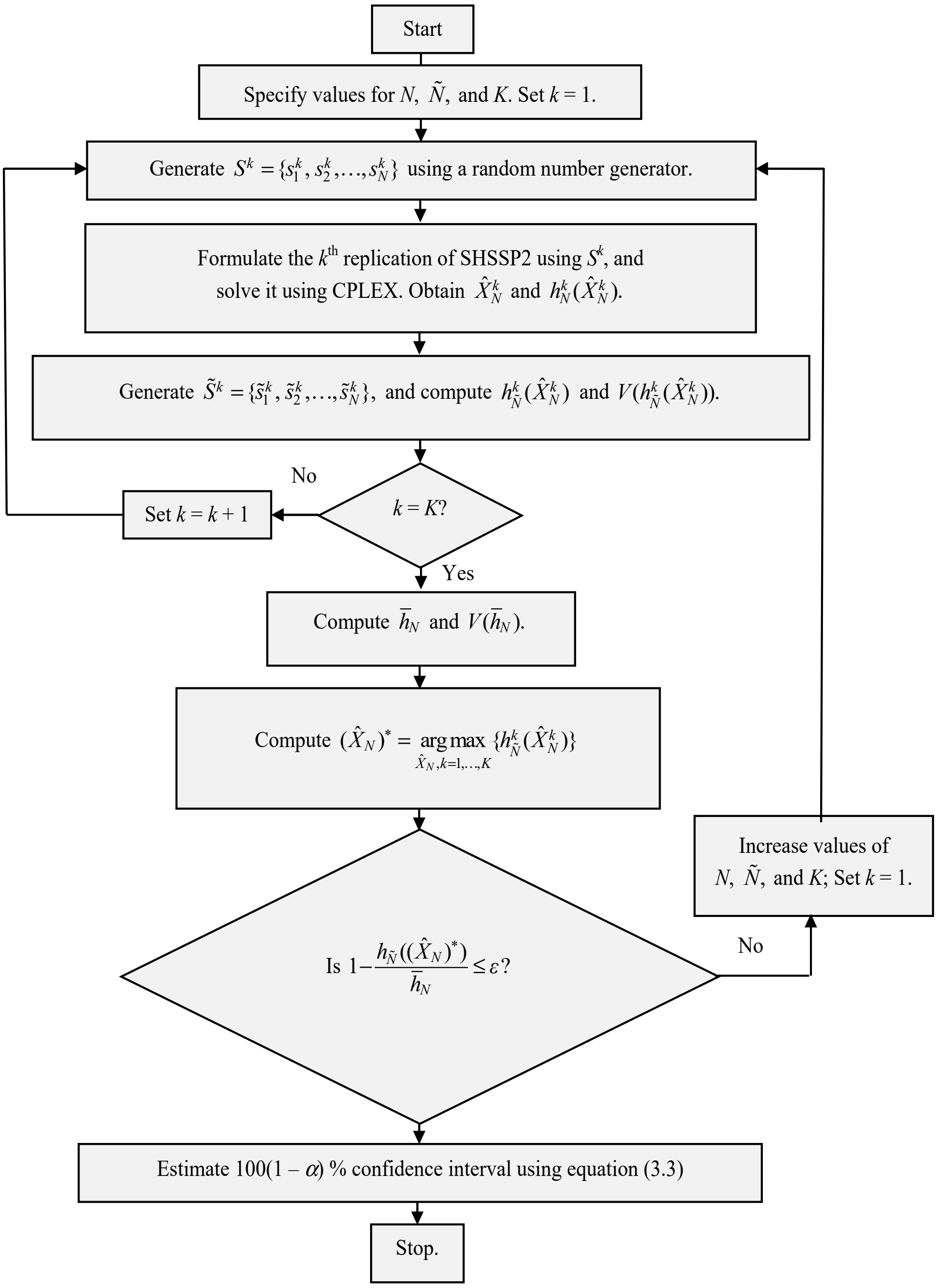
Step 8 Estimate the $100(1 - \alpha)$ % confidence interval using equation (3.3), and stop.

Note that due to the decision in step 7, this equation becomes as follows:

$$\left[0, \bar{h}_N - h_{\tilde{N}}\left(\left(\hat{X}_N\right)^*\right) + t_{1-\alpha,K-1}\sqrt{V\left[\bar{h}_N\right] + V\left[h_{\tilde{N}}\left(\left(\hat{X}_N\right)^*\right)\right]}\Big/\sqrt{K}\right] \tag{3.3}$$

In steps 1, 2, and 3, we construct and solve the k^{th} replication of SHSSP2. The statistical properties of the solutions obtained are then checked to verify convergence and to decide on whether to stop or to continue sampling further as described in the remaining steps. The solution quality improves asymptotically toward the optimal values with increments in the values of N, $\tilde{N}$, and K, but the CPU time also thus increases. Based on the results of a preliminary investigation conducted to study this trade-off, the values of N, $\tilde{N}$, and K were varied as follows: starting with the values of $(N, \tilde{N}, K) = (0.125|S|, 0.25|S|, 2)$, whenever the optimality gap at the q^{th} iteration does not satisfy the termination criterion at step 7, we increase K to $2 + q$ for the $(q + 1)^{th}$ iteration, while holding N and $\tilde{N}$ fixed. A flowchart of this integrated B&C and MXM-based algorithm is given in Figure 1.

Figure 1 Flowchart of the integrated B&C and MCM-based algorithm



4 Computational results

In this section, we report computational results for two groups of experiments. In the first group, we justify the use of the MCM for deriving near-optimal solutions to the SHSSP. In particular, we demonstrate that MCM efficiently determines solutions that exhibit a small gap between the unbiased estimates of the lower and upper bounds on the 'true' optimal objective function value. We use the B&C method of CPLEX to obtain the 'true' objective function value, $h(X^*)$, and evaluate the quality of solutions obtained by the MCM using the near-optimality criterion ratio, $\frac{h_{\tilde{N}}(\hat{X}_N)^*)}{h(X^*)}$. The second group of experiments considers problem instances that CPLEX is unable to solve because of excessive memory or CPU time requirements. For these problems, we compute $100(1 - \alpha)$ % confidence intervals on the resulting optimality gaps using the proposed MCM approach in order to evaluate solution quality. Note that we accept the results of an iteration when the tolerance ratio given by (3.4) falls within $[0, \varepsilon]$, where $0 < \varepsilon < 1$. All runs for both the groups of experiments were made on an Intel® Core™ i7-3770 CPU 3.40 GHz computer having 8.00 GB of RAM, and were implemented using CPLEX 12.5 built in Java 1.7.0_10, where the software embedding Java is NetBeans IDE 7.2.1.

Note that if the initial and final values of $(N, \tilde{N}, K)$ are identical for an instance, this means that the stopping criterion at Step 7 was satisfied after the first loop through the procedure for that instance. Also, note that $h_{\tilde{N}}((\hat{X}_N)^*)$ is an unbiased estimate of the lower bound of the 'true' optimal objective function value, but there is still a possibility that we might end up with a biased subset of scenarios even though the stopping criterion is satisfied, and in that case, $\frac{h_{\tilde{N}}((\hat{X}_N)^*)}{h(X^*)}$ can be greater than one.

Table 1 presents the results for our first group of experiments. The data for the problem instances specified in Table 1 is as follows:

Problem instance	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
|M|	2	2	2	2	2	2	2	2	2	2
|I|	5	5	5	5	5	5	5	5	5	5
|J|	5	6	7	8	9	10	11	12	13	14
|S|	32	32	32	32	32	32	32	32	32	32

Problem instance	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*
|M|	2	2	2	2	2	3	3	3	3	3	3
|I|	5	5	5	5	5	6	6	6	6	6	6
|J|	15	16	17	18	19	6	7	8	9	10	11
|S|	32	32	32	32	32	32	32	32	32	32	32

Table 1 Results on the performances of the B&C and MCM

Problem instance	*Branch-and-Cut method (B&C)*		*Monte Carlo method (MCM)*						*Near-optimality criterion*
	Optimal objective function value $h(X^*)$	*CPU time for B&C*	*Initial* $(N,\tilde{N},K)$	*Final* $(N,\tilde{N},K)$	*Estimated optimal objective* $h_{\tilde{N}}\left((\hat{X}_N)^*\right)$	*Average objective function value* $\overline{h}_N$	*Tolerance ratio* $1-\frac{h_{\tilde{N}}\left((\hat{X}_N)^*\right)}{\overline{h}_N}$	*CPU time for MCM*	$\frac{h_{\tilde{N}}\left((\hat{X}_N)^*\right)}{h(X^*)}$
1	75,800	260	(4, 8, 2)	(4, 8, 3)	75,150.0	75,300.0	0.002	5	0.991
2	101,800	328	(4, 8, 2)	(4, 8, 2)	101,350.0	105,933.3	0.043	32	0.996
3	127,800	257	(4, 8, 2)	(4, 8, 2)	127,800.0	134,400.0	0.049	12	1.000
4	153,800	163	(4, 8, 2)	(4, 8, 3)	153,800.0	158,166.7	0.028	76	1.000
5	177,800	94	(4, 8, 2)	(4, 8, 2)	185,000.0	185,133.3	0.001	466	1.040
6	193,800	832	(4, 8, 2)	(4, 8, 4)	194,025.0	194,200.0	0.001	1,601	1.001
7	219,800	1,151	(4, 8, 2)	(4, 8, 2)	219,575.0	230,833.3	0.049	157	0.999
8	245,800	462	(4, 8, 2)	(4, 8, 3)	245,600.0	245,875.0	0.001	272	0.999
9	271,600	347	(4, 8, 2)	(4, 8, 2)	271,575.0	271,800.0	0.001	11	1.000
10	297,800	404	(4, 8, 2)	(4, 8, 2)	297,575.0	297,800.0	0.001	59	0.999
11	321,800	483	(4, 8, 2)	(4, 8, 5)	331,000.0	341,500.0	0.031	242	1.029
12	345,800	720	(4, 8, 2)	(4, 8, 4)	344,900.0	346,700.0	0.005	298	0.997
13	369,800	670	(4, 8, 2)	(4, 8, 2)	370,025.0	370,475.0	0.001	3	1.001
14	393,800	863	(4, 8, 2)	(4, 8, 2)	393,800.0	394,475.0	0.002	124	1.000
15	409,600	6,017	(4, 8, 2)	(4, 8, 2)	408,450.0	409,800.0	0.003	199	0.997
16	122,700	6,161	(4, 16, 2)	(4, 16, 2)	122,700.0	123,150.0	0.004	492	1.000
17	158,700	10,230	(4, 16, 2)	(4, 16, 3)	158,362.5	159,000.0	0.004	20	0.998
18	194,700	28,776	(4, 16, 2)	(4, 16, 3)	194,025.0	199,275.0	0.026	121	0.997
19	230,700	31,220	(4, 16, 2)	(4, 16, 4)	230,812.5	231,825.0	0.004	486	1.000
20	266,700	20,891	(4, 16, 2)	(4, 16, 5)	266,700.0	266,925.0	0.001	641	1.000
21	302,700	20,371	(4, 16, 2)	(4, 16, 4)	302,700.0	312,725.0	0.032	606	1.000

Table 2 Results for large-sized problem instances by the MCM

Problem instances	$\|M\|$	$\|I\|$	$\|J\|$	$\|S\|$	*Initial* $(N, \tilde{N}, K)$	*Final* $(N, \tilde{N}, K)$	*Best objective function value* $h_{\tilde{N}}\left((\hat{X}_N)^*\right)$	*Average objective function value* $\bar{h}_N$	*Tolerance ratio* $1-\frac{h_{\tilde{N}}\left((\hat{X}_N)^*\right)}{\bar{h}_N}$	*% Confidence interval on optimality gap* $\left(\frac{\text{confidence interval}}{h_{\tilde{N}}\left((\hat{X}_N^k)^*\right)}\right)\times 100$	*CPU time for the MCM*
1	3	6	12	64	(4, 16, 2)	(4, 16, 2)	338,700	338,700	0.000	10.102	302.129
2	3	6	13	64	(4, 16, 2)	(4, 16, 2)	374,925	374,925	0.000	9.471	363.608
3	3	6	14	64	(4, 16, 2)	(4, 16, 3)	410,475	410,925	0.001	4.581	770.224
4	3	6	15	64	(4, 16, 2)	(4, 16, 3)	447,487.5	469,516.7	0.047	1.432	1,035.914
5	4	6	6	64	(4, 16, 2)	(4, 16, 2)	123,350	124,550	0.010	21.822	39.353
6	4	6	7	64	(4, 16, 2)	(4, 16, 2)	165,037.5	166,100	0.006	28.623	66.569
7	4	6	8	64	(4, 16, 2)	(4, 16, 2)	211,825	212,500	0.003	12.286	224.170
8	4	6	9	64	(4, 16, 2)	(4, 16, 3)	260,275	260,725	0.002	6.572	606.243
9	4	6	10	64	(4, 16, 2)	(4, 16, 4)	307,600	308,050	0.001	3.540	971.760
10	4	6	11	64	(4, 16, 2)	(4, 16, 3)	355,600	356,050	0.001	7.764	672.386
11	4	6	12	64	(4, 16, 2)	(4, 16, 2)	403,600	411,825	0.020	16.110	596.087
12	5	6	6	64	(4, 16, 2)	(4, 16, 2)	124,787.5	125,350	0.004	18.661	22.270
13	5	6	7	64	(4, 16, 2)	(4, 16, 2)	165,350	165,575	0.001	10.202	124.912
14	5	6	8	64	(4, 16, 2)	(4, 16, 2)	212,812.5	212,925	0.001	13.912	199.888
15	5	6	9	64	(4, 16, 2)	(4, 16, 5)	265,375	265,900	0.002	2.002	1,492.744
16	5	6	10	64	(4, 16, 2)	(4, 16, 2)	324,050	324,050	0.000	13.890	351.362

The following key observations are evident from Table 1. First, the problem instances that are highlighted (instances 4, 6, 8, 11, 12, and 17–21) required extra replications (additional loops through steps 2–7). In particular, instances 4, 8, 17, and 18 required one more loop, instances 6, 12, 19, and 21 required two more loops, and instances 11 and 20 required three more loops. Second, the worst tolerance ratio obtained was 0.049 (for instances 3 and 7). The average value of the tolerance ratio was 0.014, which indicates that the difference between the lower and upper bounds at termination was about 1.4% on average. Third, the worst near-optimality criterion obtained was 1.04 (for instance 5), where the average near-optimality criterion value was 1.002, which implies that the difference between the true optimal objective function value obtained using CPLEX and the estimated optimal objective function value obtained by MCM is about 0.5% on average. For 19 out of the 21 instances, the MCM approach consumed less CPU time than CPLEX in obtaining a solution. Not counting the case of instances 5 and 6 for which MCM required larger CPU times as compared with that required by CPLEX, the average CPU time required by MCM was 203 seconds while that required by CPLEX was 6,830 seconds.

Table 2 presents results for the second group of instances for which an optimal solution could not be obtained by the direct B&C method implemented within CPLEX for a time limit of 120,000 seconds. Note that the worst termination tolerance ratio obtained was 0.047 (for instance 4), and the average of these values over all instances was 0.006, which implies that the difference between the unbiased estimates of lower and upper bounds is 0.6% on average. The width of the 95% confidence interval on the optimality gap (see equation (3.3), i.e., the width of the confidence interval on the unknown value of $[h(X^*) - h(\hat{X}_N)^*]$ is presented in the penultimate column of Table 2 as a percentage of the lower bound estimate $(h_{\tilde{N}}(\hat{X}_N)^*)$. The average of these values is 11.311%, which indicates that the confidence interval on the optimality gap is within 12% of the lower bound, on average. The smallest width of this confidence interval is 1.432% (for instance 4) while the largest width is 28.632% (for instance 6). Six problem instances (namely, 3, 4, 8, 9, 10, and 15), which are highlighted in Table 2, needed additional loops through steps 2–7, and they also indicate smaller confidence intervals. Hence, even though the tolerance ratio obtained by our method is very small, a smaller confidence interval can be obtained by using a larger value of *K*, albeit at the cost of increased CPU times. Note that the maximum of the CPU time values shown in the last column of Table 2 is within 1,500 seconds.

5 Concluding remarks

We have addressed in this paper a two-stage stochastic programming approach for a SHSSP where the surgical operating times are probabilistic. In view of the fact that the number of underlying scenarios in a discretised representation of the joint probability distribution of operating times can grow rapidly with the size of the problem, we utilised a hybridised B&C-MCM-based approach to solve this problem, where the reduced set of scenario sub-problems generated by the MCM technique are iteratively solved using the B&C algorithm implemented within CPLEX. Unbiased estimates of lower and upper bounds on the optimal objective function value were used to design a suitable stopping criterion. In addition, a 100(1 – α)% confidence interval on the optimality gap at termination was derived. For comparative purposes, using relatively small-sized problems that could be solved to optimality by CPLEX, our computational results revealed that the proposed hybridised B&C-MCM approach can solve problems several times faster (by a factor of 33.7 on average, barring an outlier) than simply resorting to the B&C method of CPLEX alone. The relative difference between the lower and upper bounds obtained by our method was 1.4% on average, and the relative difference between the optimal objective function value and the best objective function value obtained by our method was about 0.5% on average. For larger-sized problems that could not be solved by CPLEX due to out-of-memory difficulties or excessive computational times, our method generated solutions with the relative difference between the unbiased estimates of lower and upper bounds being about 0.6% on average. Also, the average percentage confidence interval on the optimality gap obtained for these problems was 11.311%. Hence, the proposed hybridised B&C-MCM approach offers an effective solution methodology for the SHSSP addressed in this paper.

Acknowledgements

This work has been supported by the National Science Foundation under Grant No. CMMI-0856270.

References

Batun, S., Denton, B.T., Huschka, T.R. and Schaefer, A.J. (2011) 'Operating room pooling and parallel surgery processing under uncertainty', *INFORMS Journal on Computing*, Vol. 23, No. 2, pp.220–237.

Beliën, J. (2007) 'Exact and heuristic methodologies for scheduling in hospitals: problems, formulations and algorithms', *4OR: A Quarterly Journal of Operations Research*, Vol. 5, No. 2, pp.157–160.

Beliën, J. and Demeulemeester, E. (2006) 'Scheduling trainees at a hospital department using a branch-and-price approach', *European Journal of Operational Research*, Vol. 175, No. 1, pp.258–278.

Beliën, J. and Demeulemeester, E. (2007) 'Building cyclic master surgery schedules with leveled resulting bed occupancy', *European Journal of Operational Research*, Vol. 176, No. 2, pp.1185–1204.

Berrada, I., Ferlan, J. and Michelon, P. (1996) 'A multi-objective approach to nurse scheduling with both hard and soft constraints', *Socio-Economic Planning Science*, Vol. 30, No. 3, pp.183–193.

Blake, J.T. and Donald, J. (2002) 'Mount Sinai hospital uses integer programming to allocate operating room time', *Interfaces*, Vol. 32, No. 2, pp.63–73.

Bosi, F. and Milano, M. (2001) 'Enhancing constraint logic programming branch and bound techniques for scheduling problems', *Software Practice and Experience*, Vol. 31, No. 1, pp.17–42.

Burke, E.K., Causmaecker, P.D., Petrovic, S. and Berghe, G.V. (2002) 'A multi-criteria metaheuristic approach to nurse rostering', *Proceedings of the 2002 Congress on Evolutionary Computation (CEC2002)*, pp.1197–1202.

Cardoen, B., Demeulemeester, E. and Beliën, J. (2009) 'Sequencing surgical cases in a day-care environment: an exact branch-and-price approach', *Computers & Operations Research*, Vol. 36, No. 9, pp.2660–2669.

Chen, J.G. and Yeung, T.W. (1993) 'Hybrid expert-system approach to nurse scheduling', *Computers in Nursing*, Vol. 11, No. 4, pp.183–190.

Denton, B.T. (Ed.) (2013) *Handbook of Healthcare Operations Management, Methods and Applications*, Springer, New York, NY.

Denton, B.T., Miller, A.J., Balasubramanian, H.J. and Huschka, T.R. (2010) 'Optimal allocation of surgery blocks to operating rooms under uncertainty', *Operations Research*, Vol. 58, No. 4, pp.802–816.

Dexter, F., Macario, A. and Traub, R.D. (1999) 'Which algorithm for scheduling add-on elective cases maximizes operating room utilization?', *Anesthesiology*, Vol. 91, No. 5, pp.1491–1500.

Dexter, F., Traub, R.D. and Lebowitz, P. (2001) 'Scheduling a delay between different surgeons' cases in the same operating room on the same day using upper prediction bounds for case durations', *Anesthesia and Analgesia*, Vol. 92, No. 4, pp.943–946.

Guinet, A. and Chaabane, S. (2003) 'Operating theatre planning', *International Journal of Production Economics*, Vol. 85, No. 1, pp.69–81.

Jaszkiewicz, A. (1997) 'A metaheuristic approach to multiple objective nurse scheduling', *Foundations of Computing and Decision Science*, Vol. 22, No. 3, pp.169–184.

Jaumard, B., Semet, F. and Vovor, T. (1998) 'A generalized linear programming model for nurse scheduling', *European Journal of Operational Research*, Vol. 107, No. 2, pp.1–18.

Jebali, A., Alouane, A.B.H. and Ladet, P. (2006) 'Operating rooms scheduling', *International Journal of Production Economics*, Vol. 99, Nos. 1–2, pp.52–62.

Lamiri, M., Xie, X., Dolgui, A. and Grimaud, F. (2008) 'A stochastic model for operating room planning with elective and emergency demand for surgery', *European Journal of Operational Research*, Vol. 185, No. 3, pp.1026–1037.

Lapierre, S.D., Batson, C. and McCaskey, S. (1999) 'Improving on time performance in health care organizations: a case study', *Health Care Management Science*, Vol. 2, No. 1, pp.27–34.

Litvak, E. and Long, M.C. (2000) 'Cost and quality under managed care: irreconciable differences', *The American Journal of Managed Care*, Vol. 6, No. 3, pp.305–312.

Mak, W.K., Morton, D.P. and Wood, R.K. (1999) 'Monte Carlo bounding techniques for determining solution quality in stochastic programs', *Operations Research Letters*, Vol. 24, No. 1, pp.47–56.

Marcon, E., Kharraja, S. and Simonnet, G. (2003) 'The operating theatre planning by the follow-up of the risk of no realization', *International Journal of Production Economics*, Vol. 85, No. 1, pp.83–90.

Ozakarahan, I. (1989) 'Flexible nurse scheduling support systems', *Computer Methods and Programs in Biomedicine*, Vol. 30, Nos. 2–3, pp.145–153.

Ozakarahan, I. (1995) 'Allocation of surgical procedure to operating rooms', *Journal of Medical Systems*, Vol. 19, No. 4, pp.333–352.

Ozakarahan, I. (2000) 'Allocation of surgeries to operating rooms using goal programming', *Journal of Medical Systems*, Vol. 24, No. 6, pp.339–378.

Santibanez, P., Begen, M. and Atkins, D. (2005) *Managing Surgical Waitlists for a British Columbia Health Authority*, Centre for Operations Excellence, Sauder School of Business, University of British Columbia, Canada, Research Report.

Shapiro, A. (2001) 'Monte Carlo simulation approach to stochastic programming', *Proceedings of the 2001 Winter Simulation Conference*, pp.428–431.

Shapiro, A. (2003) 'Monte Carlo sampling approach to stochastic programming', *ESAIM: Proceedings*, Vol. 13, pp.65–73.

Sherali, H.D. and Smith, J.C. (2002) 'Improving discrete model representations via symmetry considerations', *Management Science*, Vol. 13, No. 2, pp.1396–1407.

Sherali, H.D., Ramahi, M.H. and Saifee, Q.J. (2002) 'Hospital resident scheduling problem', *Production Planning and Control*, Vol. 13, No. 2, pp.220–233.

Shylo, O.V., Prokopyev, O.A. and Schaefer, A.J. (2012) 'Stochastic operating room scheduling for high-volume specialties under block booking', *INFORMS Journal on Computing*, Vol. 25, No. 4, pp.682–692.

Trivedi, V.M. and Warner, D.M. (2001) 'A branch and bound algorithm for optimum allocation of float nurses', *Management Science*, Vol. 47, No. 10, pp.972–981.

Venkataraman, R. and Brusco, M.J. (1996) 'An integrated analysis of nurse staffing and scheduling policies', *Omega*, Vol. 24, No. 1, pp.57–71.

Verweij, B., Ahmed, S., Kleywegt, A.J., Nemhauser, G. and Shapiro, A. (2003) 'The sample average approximation method applied to stochastic routing problems: a computational study', *Computational Optimization and Application*, Vol. 24, No. 2, pp.289–333.

Weiss, E.N. (1990) 'Models for determining estimated start times and case orderings in hospital operating rooms', *IIE Transactions*, Vol. 22, No. 2, pp.143–150.